（说明： A3双面印刷）

班级： 学号： 姓名：

装 订 线

**杭州师范大学经济与管理学院2017-2018学年第2学期期末考试**

**《计量经济学》试卷（A）**

1. 判断（共10分，每题1分）

1．所谓无偏性（Unbiased），指的是一个参数的估计值的期望值等于其真实值。 （ R ）

2. 序列相关（Serial Correlation）是时间序列(Time Series)数据特有的问题，而异方差（Heteroskedasticity）则是截面(Cross-Section)数据特有的问题。 （ W ）

3. 用存在单位根（Unit Roots）的数据进行回归，得到的结果必然是有偏的。 （ W ）

**4. 零条件均值假设（）要求核心解释变量与误差项无关，对控制变量无要求。**  （ W ）

5. 在回归时，放入越多的解释变量，R2就会越大，但Adjust-R2不一定。 （ R ）

6. 工具变量估计会加剧多重共线性（Multicollinearity）。 （ R ）

7. 如果存在处理效应异质性（Treatment-effect Heterogeneity），那OLS获得的结果，其实就是一个平均处理效应。 （ R ）

8. 线性概率模型（Linear Probability Model, LPM）必然存在异方差。 （ R ）

9. 即使分组不是随机的，如果我们能控制这非随机因素，那双重差分模型（Difference-in-difference, DID）模型的估计结果依然可能是无偏（Unbiased）而一致（Consistent）的。 （ R ）

10. 在二阶段最小二乘法中，我们可以直接检验二阶段估计的残差与工具变量的相关性，来检验工具变量的外生性。 （ W ）

1. 选择（共40分，每题2分）

根据如下命令和执行结果，完成1-3题

use http://fmwww.bc.edu/ec-p/data/wooldridge/hprice1  
qui: reg price lotsize sqrft bdrms  
predict u,res

g usq=u\*u

reg usq price lotsize sqrft bdrms

图片包含 屏幕截图

已生成高可信度的说明

1. 这组命令的目的是：B
2. 检验序列相关
3. **检验异方差**
4. 检验单位根
5. 检验模型误设
6. 执行的结果说明它检验的问题存在吗 A
7. 存在
8. 不存在
9. 都有可能
10. 不知道
11. 这个结果说明以上的回归存在什么问题？ D
12. 有偏
13. 有偏而不一致
14. 不一致
15. 可能依然是无偏一致的，但不再是最优的
16. 在回归中，以下哪个部分，是与残差项相关的 C
17. 解释变量
18. 拟合值
19. 被解释变量
20. 自变量

下面这组回归的结果缺失了一些信息。根据已有信息推断，完成5-7题

图片包含 屏幕截图

已生成高可信度的说明

1. 样本容量缺失了。从F统计量的自由度推断，这个回归的样本容量是多少？ D
2. 48
3. 49
4. 50
5. 51
6. 以5%为显著性标注的话，两个解释变量中，哪个可能是显著的？ A
7. copen
8. cadmn
9. 都显著
10. 都不显著
11. 这个回归的R2大约是 B
12. 0.0819
13. 0.1187
14. 不确定
15. 不知道
16. 关于时间序列中的序列相关，以下论断中错误的是： B
17. 它不会导致结果的有偏不一致
18. 我们可以在reg命令当中添加robust这个option，来获得序列相关稳健统计量
19. 动态完备模型不存在序列相关
20. 对变量做差分，可能可以缓解序列相关问题
21. 如果将回归中的所有解释变量分成两组，分别用和表示**，**那回归式就变成了 **，**此时**，**这说明：B
22. 既可以通过多元回归获得，也可以通过多次一元回归获得
23. 如果，那即使在回归中被遗漏，也不会影响的无偏和一致性
24. 如果，那一旦在回归中被遗漏，就不可能是无偏一致的。
25. 以上都是错的
26. 如果我们想研究受教育程度对收入的影响，我们在也在回归中加入了智商作为控制变量，但因为数据缺失，我们未能在回归中控制家庭背景。如果保持其他条件不变，家庭背景、智商和受教育程度都会提高收入，而家庭背景与智商无关，但智商和家庭背景的改善都会提高受教育程度，那家庭背景变量的遗漏会导致：C
27. 智商变量的系数不受影响
28. 受教育程度变量的系数依然无偏一致
29. 智商变量的系数会被低估
30. 受教育程度变量的系数会被低估
31. 关于无偏性与一致性，以下论断中正确的是 C
32. 不管何种情况，两者所需要的假设基本是一样的
33. 无偏性是一种大样本性质
34. 两者都需要某种形式的零条件均值假设
35. 即使这两个性质都不存在，回归结果也没什么问题
36. 以下关于测量误差的论断中，正确的是？ B
37. 测量误差必然会导致结果有偏而不一致
38. 如果误差与真实值不相关，那只有解释变量上的测量误差才会导致结果有偏而不一致
39. 如果误差与真实值不相关，那解释变量上的测量误差会导致系数被高估
40. 以上都不正确
41. 以下关于工具变量法的论述中，正确的是： A
42. 工具变量的数目至少要与内生变量数目一样多
43. 在回归时需要严格区分不同的内生变量与其对应的工具变量，
44. 如果工具变量本身是外生，且与内生变量相关的，那工具变量法的估计结果总是优于普通OLS
45. 工具变量回归中不存在异方差问题
46. 考虑以下联立方程组，,是内生变量，而是外生变量。

，其中哪个方程是不可识别的？ A

1. 方程(1)
2. 方程(2)
3. 都可识别
4. 都不可识别
5. 以下关于各种检验的说法，正确的是： B (这道题有问题,AB都算对)
6. Dickey-Fuller检验可用于检验单位根，其初始假设是存在单位根
7. Hausman检验可用于对比各种假设有差异的回归结果是否存在显著差异
8. 即使存在异方差， t检验和F检验也依然是有效的
9. 大样本情况下，LM检验是无效的
10. 时间序列数据中经常存在趋势性，这可能会导致伪回归，因而需要设法消除其影响。消除时间序列数据趋势性的方法包含： A
11. 在回归中直接控制趋势变量
12. 对所有变量，都将其对趋势变量做回归获得拟合值，再用拟合值来进行回归
13. 为每个时期都设置一个虚拟变量放入回归
14. 对所有变量分别做移动平均，再将移动平均后的变量放入回归
15. **根据如下指令，回答以下问题：  
    reg price lotsize sqrft bdrms  
    predict ybar, xb  
    g y2=ybar\*ybar  
    g y3=y2\*ybar  
    reg price lotsize sqrft bdrms y2 y3  
    test y2 y3**

**这组命令是在？A**

1. **检验模型是否存在未被解释的非线性部分**
2. **检验异方差**
3. **检验内生性**
4. **检验序列相关**
5. **第17题的最后两条指令执行结果如下  
   图片包含 屏幕截图

   已生成极高可信度的说明  
   这说明？B**
6. **17题所检验的问题不存在**
7. **最后一步回归存在多重共线性**
8. **y2和y3不可以被同时放入回归**
9. **以上都不对**
10. 为什么对时间序列，OLS的无偏和一致性并不要求随机抽样，而对于截面数据我们就需要这个假设？ D
11. 因为时间序列数据可能存在自相关
12. 因为时间序列数据可能存在单位根
13. 因为时间序列数据可能存在趋势性和季节性
14. 因为时间序列数据是对同一个体的连续观测，不可能满足随机抽样
15. 以下关于在回归所用的模型中添加一个解释变量的论断中错误的是：D
16. 如果新添加的变量与其他解释变量相关，则会加剧多重共线性问题
17. 是否添加一个变量，其背后应当有经济学理论做支撑
18. 无论新添加什么变量，R2都至少不会减少
19. 我们可以使用**逐步回归法**，先在回归中加入所有变量，再逐步剔除不显著的，只保留显著的变量。
20. 计算及简答题（40分，每题20分）
21. 我们想知道组织工人参与培训，是否能降低生产中的废品率。因而我们做了两个回归。其中lscrap是相应企业的废品率，grant则是一个虚拟变量，如果这个企业组织过培训，它就等于1，否则等于0。grant\_1代表的则是相应企业上一年份是否组织过培训。根据如下图所示回归命令和结果，回答此后问题。  
    图片包含 屏幕截图

    已生成极高可信度的说明  
    图片包含 屏幕截图

    已生成极高可信度的说明
22. 完成如下回归结果报告表格（5分）

|  |  |  |
| --- | --- | --- |
|  | 随机效应回归 | 固定效应回归 |
| grant | -0.214  ( 0.148) | -0.252  ( 0.151 ) |
| grant\_1 | -0.373  ( 0.205) | -0.422  ( 0.210) |
| 时间虚拟变量 | 控制 | 控制 |
| 样本容量 | 162 | 162 |

1. 我们回归所用的两种方法有什么差异？

通过减去均值，固定效应可以控制随时间不变的不可观测因素，不会因为这些因素与解释变量的相关性而导致结果的有偏而不一致。代价是所有随时间不变的因素都会在回归中被约去而无法估计。此外，固定效应还会导致变量的总离差减少，使估计误差增加，有效性下降。固定效应也不能与滞后因变量同时出现。

随机效应不再控制随时间不变的不可观测因素，因而不具有固定效应的各项缺点。但它的要求这些随时间不变的不可观测因素与解释变量不相关，假设比固定效应严格。此外，这因素还使得随机效应必然存在序列相关，因而需要使用GLS来估计。

1. 我们将固定和随机效应回归的结果分别存到FE和RE当中，然后做一个hausman检验。其结果如下：  
   图片包含 屏幕截图

   已生成极高可信度的说明这个检验的命令应当如何输入？其结果说明什么？

hasuman FE RE

不存在显著差异。随机效应可以用。

1. 根据以下Stata指令及其执行结果，回答此后问题（20分，每小题5分）

use http://fmwww.bc.edu/ec-p/data/wooldridge/hprice2

g rooms2=rooms\*rooms

reg price lnox ldist rooms rooms2 stratio

图片包含 屏幕截图

已生成高可信度的说明

1. 在回归中，price和rooms变量分别表示房屋总价和房间数目，根据这个结果，保持其他条件不变，一个三个房间的房子，和两个房间的房子相比，价格差多少？

5\*2364.375-23390.36=-11568.485

1. 保持其他条件不变，房间数目和房屋总价是什么关系？是否存在某个房屋数目，会使得房价最高，或者最低？存在的话，是多少？

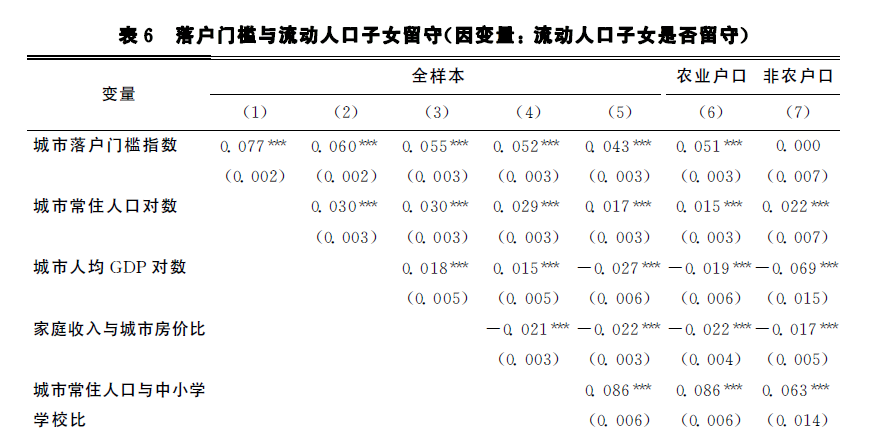
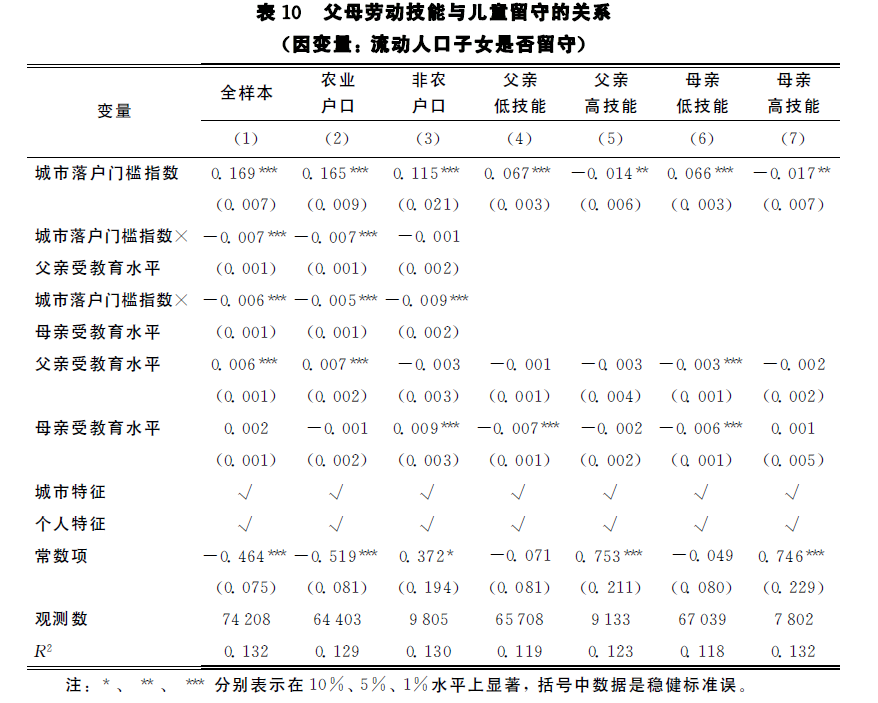
U型曲线，最低，大约5

1. 如果输入test rooms rooms2，得到的结果应该是显著的吗？为什么？这个结果说明什么？  
   这是F test，其初始假设是两个变量系数都为0，根据已有t test结果，两个变量系数都显著异于零。因而F test结果应该显著。这一结果说明房间数目存在显著影响。
2. 在回归中，lnox和ldist分别是房屋所在地区空气污染指数和离城市中心区距离的对数值，如果我想知道在市区和郊区，空气污染对房价的影响是否存在差异，我应该如何设计回归？如何用这个回归的结果，来回答这个问题？

设置lnox与ldist的交叉项放入回归。看这个交叉项系数是否显著。

1. 问答题（10分）

中国存在规模庞大的留守儿童。他们在成长的关键期，不得不与父母分离，这一状况会对他们的成长产生严重的负面影响。2018年1月份的经济学季刊上，魏东霞与谌新民的论文《落户门槛、技能偏向与儿童留守——基于2014年全国流动人口监测数据的实证研究》讨论了这个话题。以下两张表格是文中部分实证结果。

这些结果说明，落户门槛的提高会如何影响儿童留守？它对父母户籍和受教育程度不同的儿童的影响又分别有什么差异？

落户门槛的提高，增加了儿童留守的可能性。只对农村户籍有效，并且父母的受教育程度会缓解这一影响。