（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2015-2016学年2学期期末考试**

**《计量经济学》试卷（A）**

1. 判断题 （每题1分，共10分）

1. 在用工具变量（Instrumental Variables）法估计时，如果有多个内生变量，则我们应当将工具变量也分成多组，分别用作不同内生变量的工具变量。 ( 错 )

2. 时间序列数据做OLS时，如果仅满足同期外生假设（Contemporaneously Exogenous），而不满足严格外生假设（Strictly Exogenous），则OLS估计结果虽然仍可能保持无偏（Unbiased），但将是不一致的(Inconsistent)。 ( 错 )

3. 在用OLS估计模型时，如果解释变量与误差项相关，那么不仅，的估计值也会是有偏且不一致的。 ( 对 )

4. 在存在序列相关(Serial Correlation)的情况下，OLS估计值是有偏而不一致的。 （ 错 ）

5. 在大样本情况下，有大数定理和中心极限定理，即使没有同方差假定，误差项的分布也是近似服从正态分布，因而我们可以照常进行t检验、F检验等统计推断。 ( 错 ) 6. 如果y和x正相关，则如果我们用y做被解释变量，用x当解释变量，做一个一元回归，这个回归当中x的系数必然大于零。 ( 对 )

7. OLS回归所得到的残差跟任一解释变量必然不相关。 ( 对 )

8. OLS的基本假设之一是模型的线性，因而形如之类的非线性模型是无法用OLS估计的。 （ 错 ）

9. 时间序列（Time Series）数据可能会因为共同趋势等问题带来谬误回归的结果，而截面(Cross-section)或面板(Panel)数据就不会。 （ 错 ）

10. 面板数据的一阶差分（First Difference）模型和固定效应（Fixed Effects）模型都可以消掉一切随时间不变的不可观测因素，在任何情况下它们得到的结果都是相同的。 （错 ）

1. 选择题（共40分，每题2分）
2. 在没有随机抽样（Random Sampling）假定的情况下，要使得OLS的估计结果具有一致性，下列哪个假设是不必要的
3. 同期外生性（Contemporaneously Exogenous）
4. 严格外生性（Strictly Exogenous）证明一致性只需要同期外生
5. 不完全共线性（No Perfect Collinearity）
6. 平稳且弱相关（Stationary and Weak Dependence）

2. 要使得OLS的结果满足无偏性和一致性，解释变量和误差项必须不相关，以下几种情况中，哪一情况不会导致和相关？

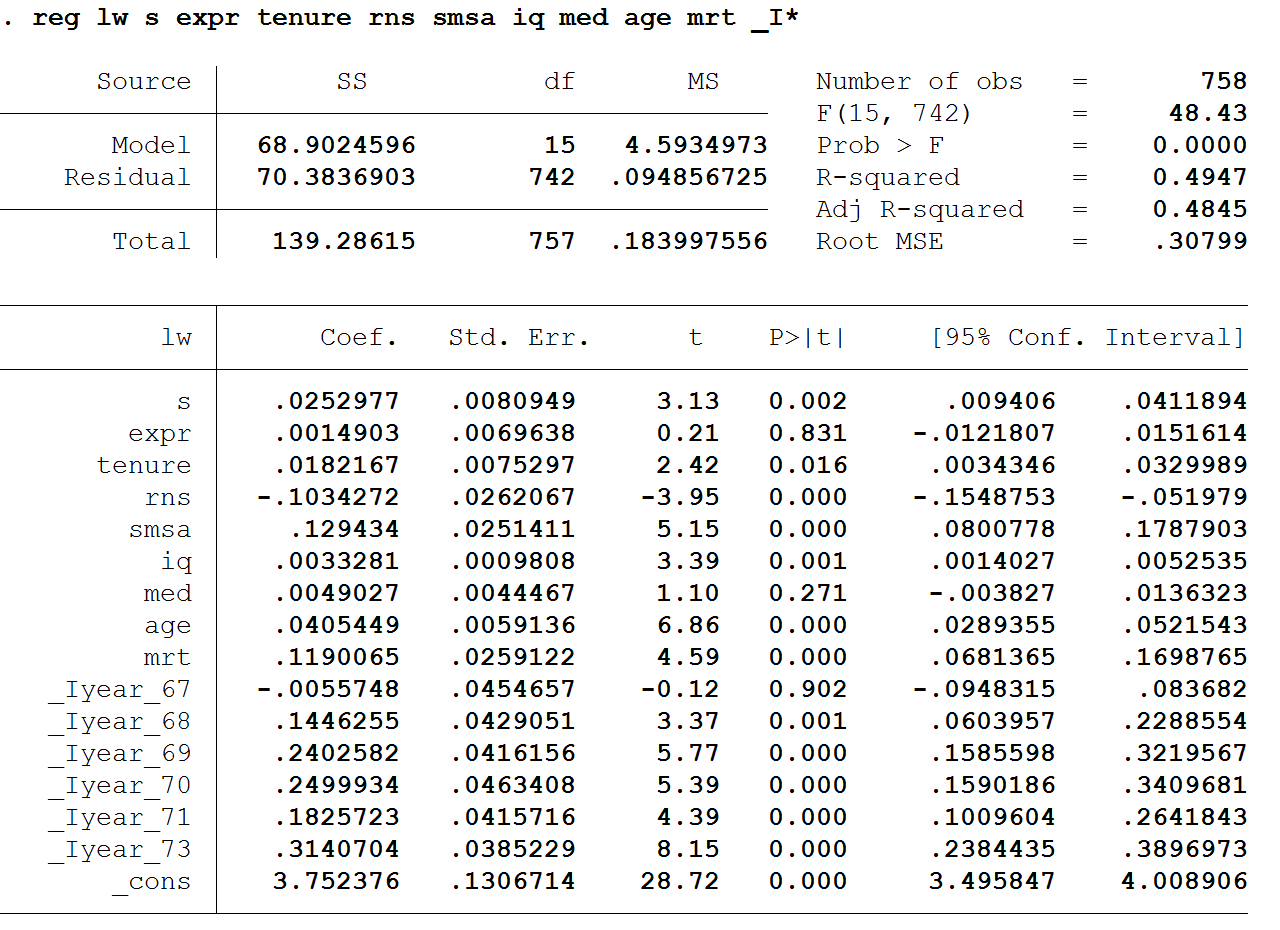
A、某一与相关的解释变量被遗漏，并且这一变量与另一解释变量相关

B、某一解释变量存在严重测量误差，并且测量误差大小与的真实值无关

C、误差方差与某一解释变量相关 这是异方差。

D、被解释变量反过来会影响某一解释变量

根据下图所示回归结果，完成3-5题。





3、这个回归中的被解释变量lw是工资的对数值。如果以5%为显著标准，则这个结果表明，以下哪个变量对工资没有显著影响？

1. 受教育程度s
2. 工作经验expr
3. 年龄age
4. 婚姻状况mrt

4、\_I开头的变量分别是各个年份相应的虚拟变量。比如\_Iyear\_67代表67年，如果观测值是67年的，则这一变量等于1，否则它等于0。样本中包含的年份是从66年一直到73年，其中72年数据缺失。从回归结果看，这些年份当中，哪一年的工资水平是最低的？

A、66年 其他选项年份虚拟变量均显著大于0.

B、68年

C、71年

D、73年

5、以下关于上图回归结果的描述中，错误的是

A、母亲受教育程度med变量对收入的影响不显著

B、从F检验结果看，各个变量是联合显著的

C、被解释变量lw的总离差SST是139.28615

D、这个回归不包含常数项 里面显然有常数项

6、 要证明OLS估计量是最优线性无偏估计量（Best Linear Unbiased Estimator, BLUE），我们需要同方差(Homoskedasticity)假定：,其中是n阶单位矩阵。当这个假定不成立时，OLS估计量是

A、有偏的

B、不一致的

C、依然无偏且一致的 异方差不影响无偏和一致性

D、以上都错

7、假设GDP变量当中是各个季度的GDP数据，未经调整时，这一数据显然兼有季节性和趋势性。假设我们还有一个时间变量t和季节变量Quarter，则以下哪组Stata命令可以实现去除这个季节性和趋势性，并将去除后的数据存入新变量SAG中？

A、tssmooth ma SAG=GDP, window(4)

B、xi: reg GDP i.Quarter t  
predict SAG

C、xi: reg GDP i.Quarter t  
predict SAG,res

D、xi: reg GDP i.Quarter  
predict SAG,res

8、我们想研究收入上的性别歧视，于是以收入为被解释变量，我们做了一个回归。以下相关论点中，正确的是：

A、通过在回归中加入性别虚拟变量，我们可以研究不同性别在收入上的差异

B、通过在回归中加入性别和时期虚拟变量，我们可以研究性别歧视在不同时期的变化情况

C、通过在回归中加入性别和教育水平变量，我们可以研究不同性别在教育回报上的差异

D、通过在回归中加入性别和国家虚拟变量，我们可以研究不同国家在性别歧视上的差异

9、William Baumol 1986年在American Economic Review上发表了名为Productivity Growth, Convergence and Welfare的论文，研究了各国之间在增长上的收敛情况。他用16个工业化国家的1870到1979年的人均GDP数据估计了如下模型：

他估计得到的接近1，这是收敛存在的强烈证据。不过，这个研究是有缺陷的，它的主要问题是：

A、没有控制其他跟增长相关的变量

B、1870年的到1979年时间太长，数据不可比

C、16个国家的样本容量太小

D、只用了工业化国家，相当于只选择了增长获得成功的国家，这导致严重的样本选择偏误（Sample Selection bias） 上课讲过

10、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：

A、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

B、如果新添加的变量与被解释变量相关，则会减少残差方差

C、是否添加一个变量，其背后应当有经济学理论做支撑

D、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。 上课讲过，这是有些老的教科书里仍然有的错误说法，应当绝对避免。

11、以下关于虚拟变量（Dummy Variable）的论断中，错误的是：

A、虚拟变量就是只能取0和1两个值的变量

B、作为~~被~~解释变量，虚拟变量在使用上与其他变量并没有什么不同

C、通常，对于年级、种族等不连续又有n>2个取值的变量，我们不会把他们直接放入回归，而是根据其不同取值，设置n个相应的虚拟变量后再加入回归 应当是n-1个

D、如果把虚拟变量当成被解释变量直接做回归，那这就是线性概率模型（Linear Probability Model, LPM）

12、假设我们有两年共计24个月若干个人收入情况的观测值，Year变量(取值为1或2)、Month变量（取值为1-12）和Male变量（取值为0或1）分别储存了数据的年份、月份和个人性别信息，数据没有任何缺失，而我们希望向回归的范围限制在第二年下半年6个月的男性收入当中。在Stata当中，如下if语句，哪个可以实现我们的目标？

A、if Year=2 & Month>6 & Male=1

B、if Male & (Year-1) & Month>6 ==才是判断，=是赋值。

C、if Year=1 & Month>6 & Male>0

D、if !Year=1 & Male & Month>6

13、如果我们想研究高等教育与个人收入之间的关系，但我们担心它们直接可能存在内生性，为了控制这种内生性，以下方法哪个是不可行的？

A、通过断点回归（Regression Discontinuity, RD）的方法，比较高考录取线上下分数几乎无差异，而是否接受高等教育有差异的人群

B、找一个外生的变量做工具变量（Instrumental Variable），比如所在地区的高考录取率，这个变量会影响个人获取高等教育的可能性。

C、跟踪观察一个群体在获取高等教育前后个人收入的变化情况，对比类似但未受高等教育的人群在同一时期的个人收入变化

D、直接对比受教育程度不同的人群的收入差异

14、为什么对截面数据，OLS的一致性并不要求数据的平稳和弱相关性，而对于时间序列数据我们就需要这个假设？

A、因为时间序列数据可能存在自相关

B、因为时间序列数据可能存在单位根

C、因为时间序列数据可能存在趋势性和季节性

D、因为时间序列数据不满足随机抽样假定

15、关于限值因变量（Limited Dependent Variables），以下论述中错误的是：

A、如果被解释变量是虚拟变量，那么我们可以用Probit/Logit模型

B、如果被解释变量等于0的数据占样本不小的比重，那我们应当使用Poisson模型 这里应当用Tobit

C、如果解释变量包含持续期之类可能存在截取（Censored）的变量，那我们应当使用Censored模型

D、Logit和Probit模型假设有所差异，因而回归得到的参数是不可直接比较的

16、为什么我们在检验异方差或序列相关时，可以用估计得到的残差方差作为误差方差的估计量，而在检验解释变量的内生性时，不可以直接检验残差和解释变量之间的相关性？以下对这个问题的解释当中，错误的是：

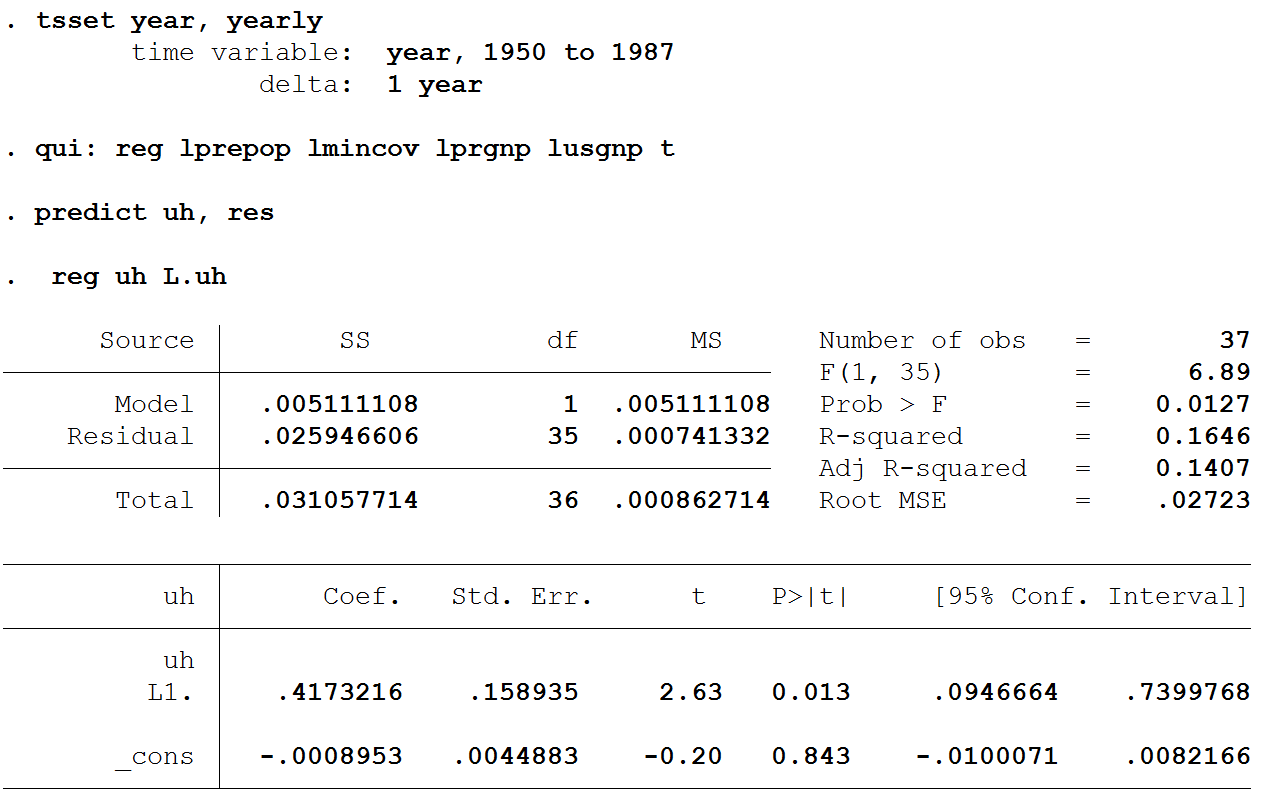
A、即使存在异方差或序列相关，残差依然是误差的一致估计量

B、当和相关时，不是的一致估计量

C、无论和是否相关，OLS的方法决定了和都必然是不相关的，因此这个检验没有意义

D、以上都是错的 ABC都是对的，所以“以上都是错的”这句话是错的。

根据如下Stata命令及执行结果，回答17-18题



17、这组命令的目的在于

A、检验单位根

B、检验模型误设

C、检验异方差

D、检验序列相关

18、命令的执行结果表明，它所检验的问题存在吗？

A、存在

B、不存在

C、都有可能

D、以上都不对

19、以下关于Hausman检验的描述中，错误的是

A、Hausman检验所直接检验的其实是两组假设不同的回归结果是否存在显著差异

B、如果Hausman检验的结果是显著的，则通常可以说明假设更严格的那组回归结果是不一致的

C、在Stata里使用hausman命令检验是否可以使用面板数据的随机效应（Random Effects）模型时，需要对比随机效应和固定效应（Fixed Effects）回归结果。此时应当先输入固定效应结果，再输入随机效应结果

D、如果选项C中的hausman检验结果不显著，则我们应当使用随机效应模型，不能使用固定效应模型 hausman只能判断能不能用随机，固定效应假设比随机效应松，不需要判断。

20、假设存在联立方程模型 ，其中任一变量的系数都不等于0，则其中可识别的方程是：

A、方程（1）

B、方程（2） 排除了

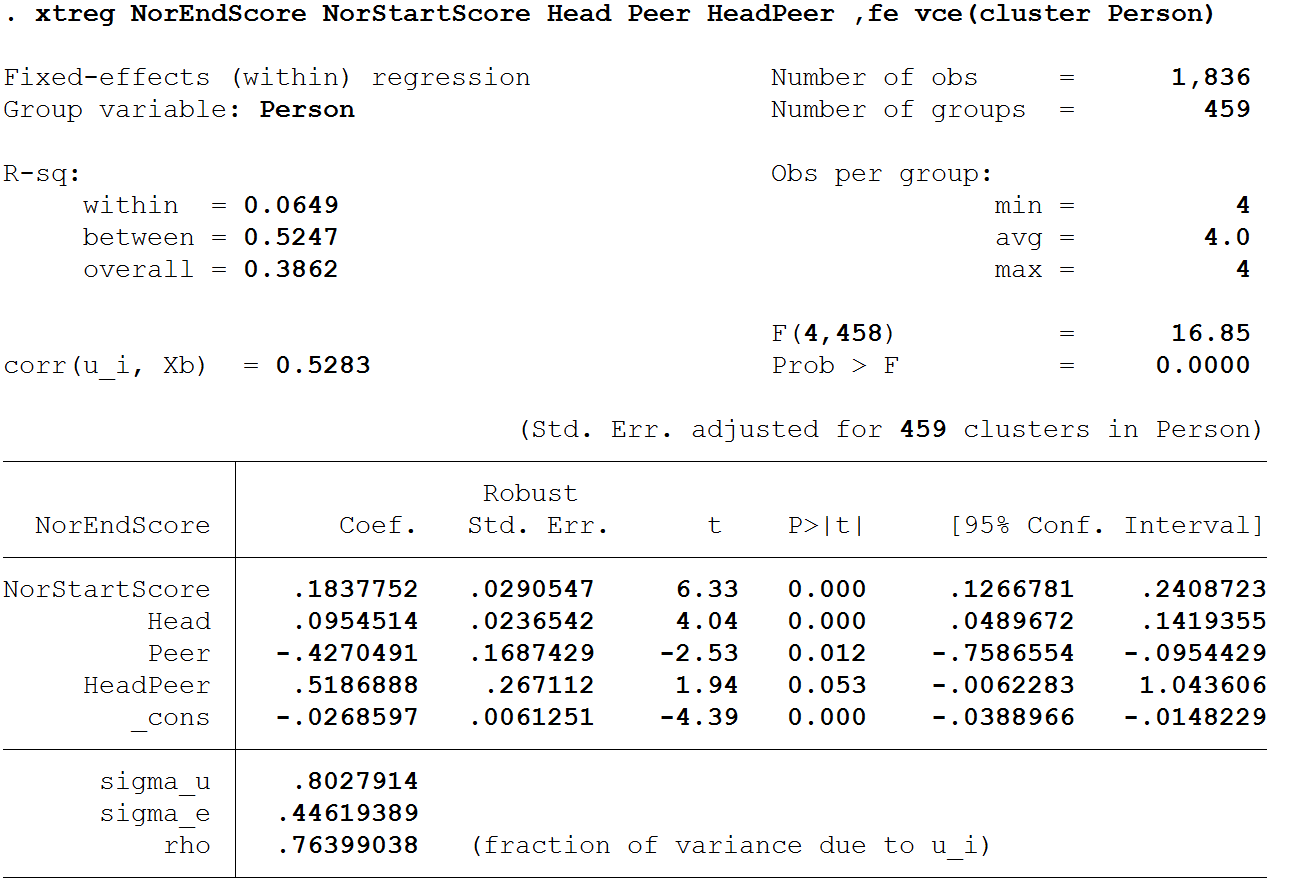
C、都可以识别

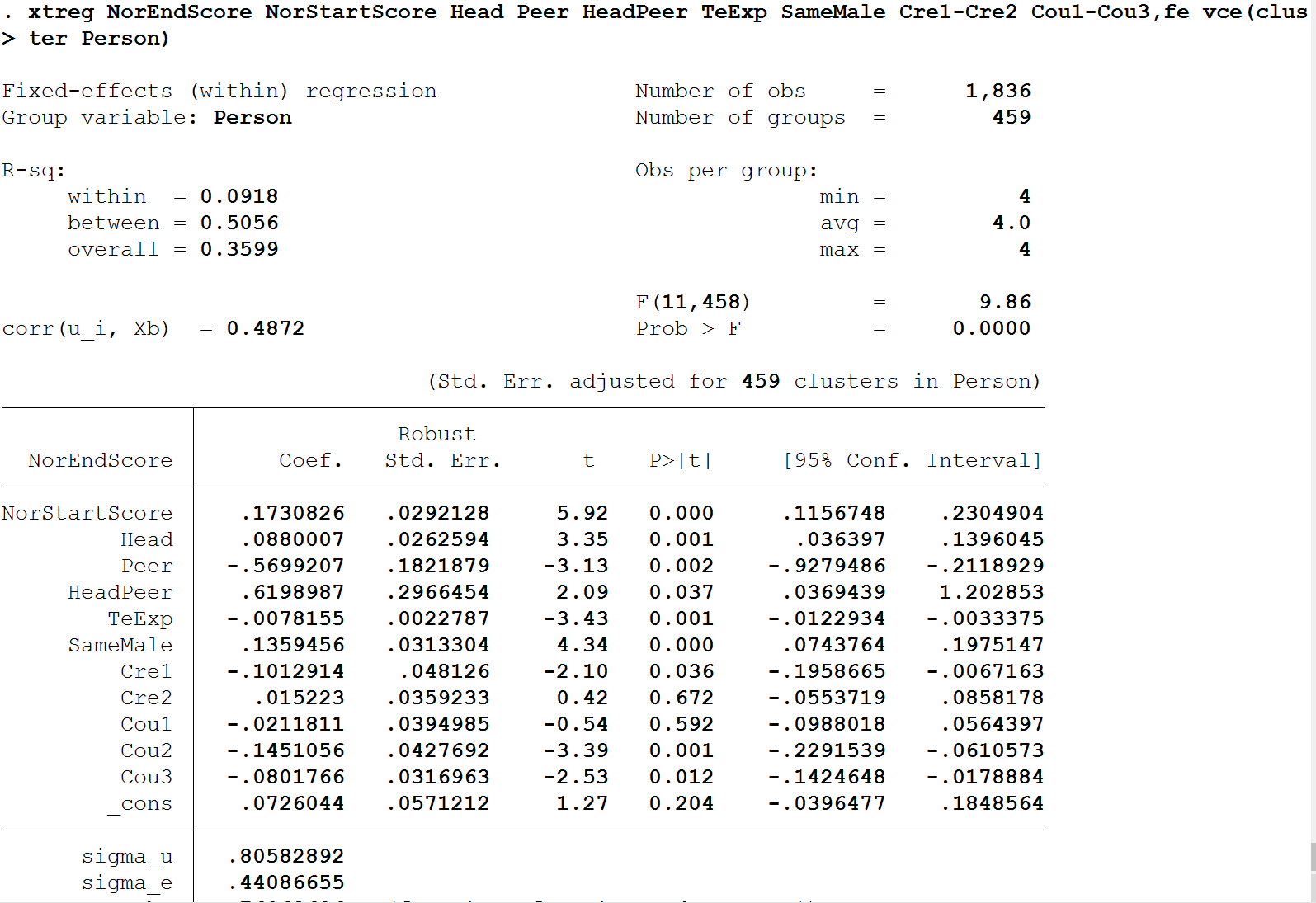
D、都不可以识别

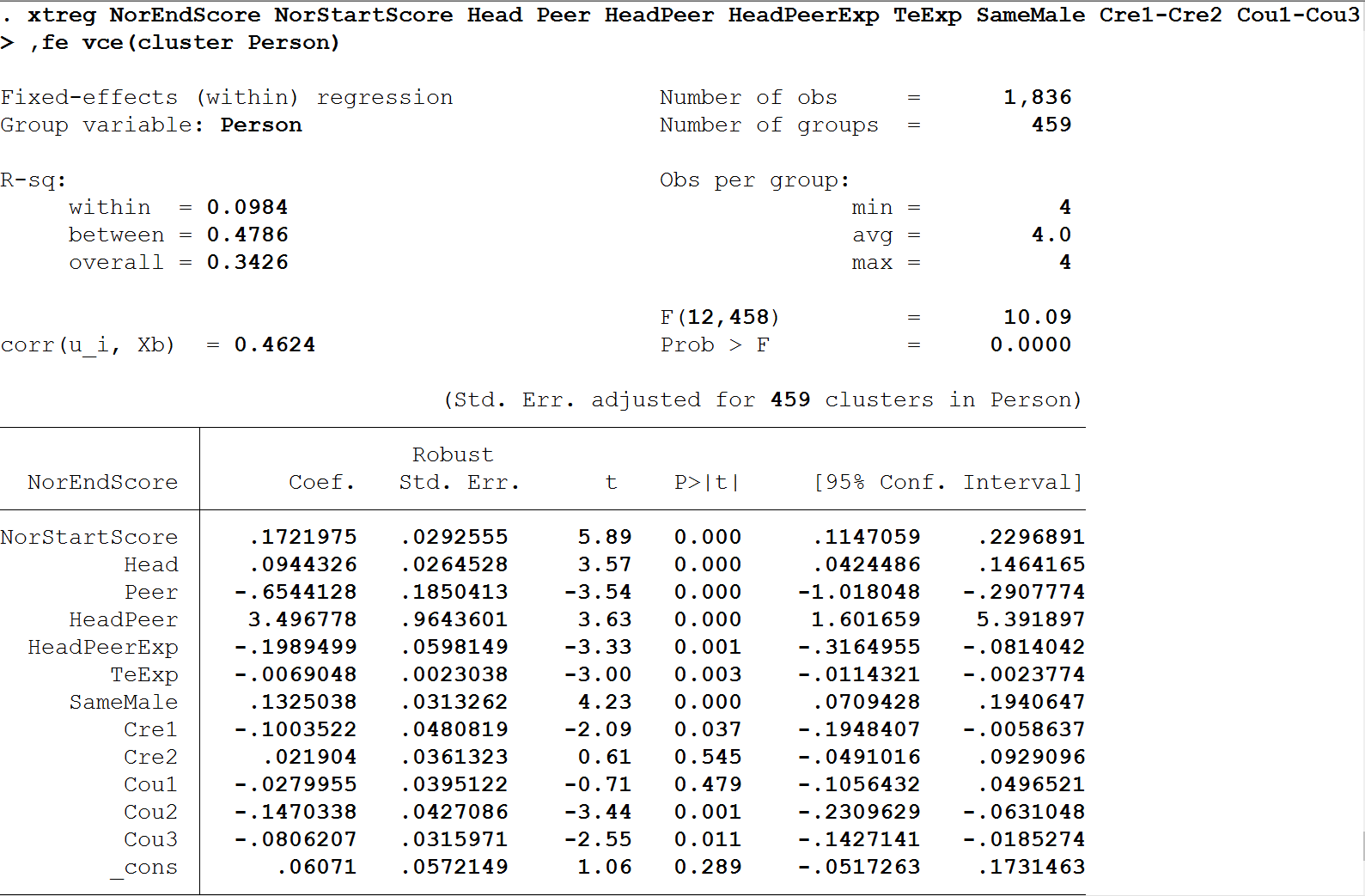
三、问答题（40分）

根据如下图所示命令及其执行结果，完成1、2两题。

图后的结果报告表格。所有小数均保留到小数点后三位。（共计15分）







1.根据图中回归结果，完成答题卷上的报告表格。所有小数均保留到小数点后三位。（共计14分）

|  |  |  |  |
| --- | --- | --- | --- |
|  | NorEndScore | NorEndScore | NorEndScore |
| 同伴效应Peer | –0.427\*\*  (0.169) | -0.570\*\*\*  (0.182) | -0.654\*\*\*  (0.185) |
| 是否班主任虚拟变量Head | 0.095\*\*  (0.024) | 0.088\*\*\*  (0.026) | 0.094\*\*\*  (0.026) |
| Peer与Head交叉项HeadPeer | 0.519\*  (0.267) | 0.620\*\*  ~~(0.026)~~  (0.297) | 3.497\*\*\*  (0.964) |
| Peer、Head和TeExp交叉项 HeadPeerExp | - | - | -0.199\*\*\*  (0.060) |
| 标准化入学成绩NorStartScore | 0.184\*\*\*  (0.029) | 0.173\*\*\*  (0.029) | 0.172\*\*\*  (0.029) |
| 师生是否同一性别SameMale | - | 0.136\*\*\*  (0.031) | 0.133\*\*\*  (0.031) |
| 教师工作经验TeExp | - | -0.008\*\*\*  (0.002) | -0.007\*\*\*  (0.002) |
| 教师职称虚拟变量 | 未控制 | 控制 | 控制 |
| 课程虚拟变量 | 未控制 | 控制 | 控制 |
| 固定效应 | 控制 | 控制 | 控制 |
| F检验结果 | 16.85\*\*\* | 9.86\*\*\* | 10.09\*\*\* |
| 样本容量 | 1836 | 1836 | 1836 |

2.上面这个回归旨在研究教育中的同伴效应。所谓同伴效应，指的是学生对其他人学习成绩的影响。被解释变量是标准化后的某中学一届每个学生英语、语文、自然、数学四门课的毕业成绩，样本内共有459个学生每人四门课的成绩。Peer变量是同班所有其他人这门课的平均分，显然，如果Peer变量对成绩的边际影响为正，那说明存在正的同伴效应。Head是个虚拟变量，它=1时表示这门课是由本班班主任上的。这个回归还通过固定效应模型控制了在不同课程间保持不变的变量，以此来避免各种无法观测的个人和班级特征可能造成的影响。根据以上信息和图中内容，回答以下几个问题（26分）。

(1)回归结果表明，存在显著的同伴效应吗？根据第二张图的回归结果，保持其他条件不变，在非班主任的课上，班级里其他人的平均分对学生成绩的边际影响是多少？在班主任的课上呢？这说明班主任对同伴效应有什么影响？（6分）

显著

当Head=0，从图二中可得，Peer对成绩的边际影响是-0.570，而当Head=1，它的边际影响是-0.570+0.620=0.05

(2)第三张图的回归添加了一个三重交叉项班主任虚拟变量Head\*同班其他人平均分\*任课教师工龄，这个交叉项的结果说明不同工龄的班主任在影响同伴效应上有什么差异？试解释这一差异。根据第三张图的回归结果，工龄为多少年的班主任，任课的班内同伴效应为0？（6分）

工龄越长的班主任影响能力越弱。

当Head=1，要使得Peer的边际影响为0，则Exp应当=(3.497-0.654)/0.199=14.286

(3)回归命令有一个Option是vce(cluster Person)，其中Person是代表不同学生的变量，添加这个Option的理由和目的分别是什么？如果不添加，可能会有什么后果？（7分）

同一个学生的不同课程成绩，和不同学生的成绩之间可能存在分组异方差，这个Option是为了计算分组异方差稳健统计量。不添加的后果是统计推断结果不可信。

(4)这是个固定效应回归模型，而解释变量之一的NorStartScore就是同一个学生标准化后的入学成绩，相当于滞后的被解释变量。而我们知道在固定效应模型中如果包含滞后一阶的被解释变量，会导致回归结果有偏而不一致，那么这个回归也是如此吗？为什么？（7分）

不会。这个回归虽然也是固定效应，但它的“面板数据”里“时间维度”是不同的课程，而不是不同的时间，与被解释变量的滞后维度不同。

五、附加题(共10分)

1. 在解释变量的外生性得不到保证的情况下，的假设不再成立，OLS的结果将是有偏而不一致的。此时，如果存在外生变量Z，使得, 且满秩，则我们可以用Z做X的工具变量(Instrumental Variable)，从而获得一致的估计量。

如果工具变量数目与内生的X数目一样，我们说此时刚好识别，此时，

如果工具变量数目比内生的X多，此时矩阵的阶数高于**,** 没有逆矩阵。这时候，我们通过二阶段最小二乘法（2SLS）来进行估计，此时，

除此之外，我们还能通过名为GMM（广义矩估计）的方法，获得一个一致估计量。

设**，**则在工具变量外生假定下，

如果和同阶，恰好识别，则使得 的就是GMM估计量。如果的阶数高于**,**此时过度识别，则是使得尽可能接近0的**。**

1. 试证明，在的情况下，

**这就是**，当Z=X，显然

粗暴一点直接把式子代入那个表达式证明等号成立，并且解唯一也行

1. 在过度识别情形下，我们可以使用最小二乘法求出所需要的，此时最小化的其实是**，**试写出此时的式子。

可以直接求导与OLS一样的思路得到估计值，或者注意到OLS估计值最小化了，因而把**当成被解释变量，把当成是解释变量代入OLS估计值的式子就能得到**

1. 如果**,,**试证明，是的一致估计量

（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2015-2016学年第2学期期末考试**

**《计量经济学》答题卷（A）**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 题号 | 一 | 二 | 三 | 四 | 五 | 六 | 总分 |
| 得分 |  |  |  |  |  |  |  |

|  |  |
| --- | --- |
| 得分 |  |

1. 判断题 （每题1分，共10分）

|  |  |
| --- | --- |
| 得分 |  |

二、选择题（共40分，每题2分）

|  |  |
| --- | --- |
| 得分 |  |

三、问答题 (共40分)

1.完成如下表格（14分）

Table 1. Basic results.

|  |  |  |  |
| --- | --- | --- | --- |
|  | NorEndScore | NorEndScore | NorEndScore |
| 同伴效应Peer | –0.427\*\*  (0.169) |  |  |
| 是否班主任虚拟变量Head |  | 0.088\*\*\*  (0.026) |  |
| Peer与Head交叉项HeadPeer |  | 0.620\*\*  (0.026) |  |
| Peer、Head和TeExp交叉项 HeadPeerExp | - | - |  |
| 标准化入学成绩NorStartScore | 0.184\*\*\*  (0.029) |  |  |
| 师生是否同一性别SameMale | - |  | 0.133\*\*\*  (0.031) |
| 教师工作经验TeExp | - | -0.008\*\*\*  (0.002) |  |
| 教师职称虚拟变量 | 未控制 | 控制 | 控制 |
| 课程虚拟变量 | 未控制 | 控制 | 控制 |
| 固定效应 | 控制 | 控制 | 控制 |
| F检验结果 |  |  |  |
| 样本容量 | 1836 | 1836 | 1836 |

注: 括号中是异方差稳健标准差统计量. \*, \*\*和\*\*\* 分别表示在 10%, 5%和1%程度上显著。虽然未在表中报告，但所有回归均包含常数项。

2. 回答各个问题（共26分）

四、附加题(共10分)