电子科技大学信息与软件工程学院

标 准 实 验 报 告

**（实验）课程名称 计量经济学**

**电子科技大学教务处制表**

**电 子 科 技 大 学**

**实 验 报 告**

**学生姓名：** **\*\*\*** **学 号：202109120\*\*\*** **指导教师：李亚静**

**实验地点： 信软楼西303**  **实验学时**：**2学时**

**一、实验室名称：软件工程实验室**

**二、实验项目名称：SAT互动回归练习**

**三、实验原理：**

**（一）普通最小二乘法**

普通最小二乘法（OLS）通过最小化残差平方和来估计回归方程中的参数，是计量经济学中最经典、应用最广泛的估计方法。该方法的使用需满足一系列古典假设，包括：

1. 回归模型是线性的，模型设定无误且含有误差项；
2. 误差项总体均值为零；
3. 所有解释变量与误差项都不相关；
4. 误差项观测值互不相关（不存在序列相关性）；
5. 误差项具有同方差(不存在异方差性)；
6. 任何一个解释变量都不是其他解释变量的完全线性函数（不存在完全多重共线性）

在上述古典假设成立的情况下，普通最小二乘估计量是最佳线性无偏估计量。所谓无偏估计量，是指OLS参数估计值的均值等于总体参数的真实值。

**（二）假设检验**

假设检验是统计推断的一个重要方法，用于对一个或多个关于参数的假设进行验证。**可以用于验证经济理论模型的参数值的显著性水平高低**。

在假设检验中，我们提出一个假设和一个备择假设。原假设通常是和我们预期相反否定的假设，而备择假设则是对原假设的否定或替代。通过对收集到的样本数据进行统计分析，我们根据统计量的结果来判断是支持原假设还是支持备择假设。

其中，**t检验**是一种常用的假设检验方法，它基于样本均值、样本标准差和样本大小，计算得到一个t统计量。通过比较该统计量与t分布的临界值，可以判断两组样本均值是否存在显著差异。如果计算得到的t统计量超过了临界值，就可以拒绝零假设，认为样本均值之间存在显著差异。

**（三）遗漏变量与不相干变量**

在计量经济学中，遗漏变量和不相干变量是两个常见的问题。

遗漏变量是指在建立经济模型时，未能将所有对因变量产生影响的重要因素考虑进去的情况。当遗漏变量存在时，经济模型的估计结果可能会出现偏差。

不相干变量是指在经济模型中加入了对因变量无实际影响的变量。这些变量可能与自变量或因变量相关，但并不能提供关于因变量的额外解释能力。

研究者需要仔细选择并纳入与因变量相关且具有经济意义的变量，避免遗漏变量，也尽量避免引入不相干变量。

**（四）多重共线性**

多重共线性是指线性回归模型中的解释变量之间由于存在精确相关关系或高度相关关系而使模型估计失真或难以估计准确。如果存在多重共线性的情况下，标准误非常大，估计的精度很小，使得模型的预测能力降低，同时也会增加误判的风险。

常见地多重共线性地检验方法有——相关性检验和方差膨胀因子（VIF）。

在发现多重共线性出现后，可以通过增大样本容量、剔除支配变量、剔除多余变量、变换解释变量地方式进行弥补。

**四、实验目的：**

* 掌握使用EViews软件进行普通最小二乘估计；
* 加深对普通最小二乘估计方法的认识。
* 在实验中认识模型中遗漏变量、不相干变量带来地影响，并做出相应地处理。
* 在实验探究中识别多重共线性，并做出诊断和处理。

**五、实验内容：**

本次实验研究了关于SAT成绩的单方程线性回归模型。实验要求从0到1完成影响因素的收集、符号的预期、模型的设定。

要求我们根据自己的判断，构建8个不同的理论模型，通过8个不同的模型设定来探究不同模型设定所带来的不同情况。

在设定过程中，我们重视考虑多重共线性、遗漏变量和不相干变量等问题。我们使用假设检验、拟合优度和经济理论等方法对模型进行诊断和检验。假设检验帮助我们评估每个自变量对SAT成绩的显著性，以判断其是否对因变量产生显著影响。拟合优度指标如R-squared值则可评估模型对实际数据的拟合程度，值越高表示模型解释数据的能力越强。此外，我们还将结合经济理论对模型结果进行解释和验证，以确保模型的经济意义和实际可行性。

最终探究一个符合四个模型设定标准的最佳模型。该模型将为我们提供关于SAT成绩的解释和预测，有助于对相关问题做出更全面和准确的分析。

**六、实验器材（设备、元器件）：**

计算机、EViews软件

1. **实验步骤：**
2. **分析影响因素：**需要广泛搜集与SAT成绩相关的变量，并根据领域知识和文献综述进行筛选。收集到的变量应该涵盖不同方面，并考虑其可能的理论联系。具体应该考虑如下内容——
   1. 这些变量可能包括学生的背景特征（如性别、年龄、家庭背景）、学术能力（如出勤率、GPA）。
   2. 分析这些变量的预期符号，并给出理论依据
3. **分析数据：**从课程网站中下载相关的数据样本，并将其导入到Eviews软件中，利用Eviews查询其基本的数据特征，具体包括——
4. 各个变量参数之间的相关系数
5. 各个变量参数均值标准差、均值等
6. **选定方程的形式：**在第一步和第二部的基础上，根据数值和相关经验、理论考虑判断哪些因素应该被纳入回归模型之中。不要简单的包含全部的变量，考虑一些不相干变量的去除。
7. **执行回归：**根据选定方程的形式在Eviews软件中执行回归
8. **结果评判：**评价回归结果的经济意义、整体拟合优度以及单个参数的符号和显著性。
9. **问题分析：**评价回归结果可能存在什么样的计量经济问题，具体可以包括——
   * 1. 遗漏变量
     2. 存在不相干变量
     3. 多重共线性
10. **更改方程设定，重新执行回归：**根据第六步的问题分析，重新更改方程的设定，并完成新一轮的结果分析（即重复3-7步），共完成8项回归的分析。
11. **总结：**选择一个回归中最优的模型设定，给出解释和理由。

**八、实验数据及结果分析：**

1. 分析影响因素

书中为我们提供了一些可能的影响因素，包括两类变量，第一类是测度学生研究水平的（包括加权绩点、出勤率等），第二类是其他类型的（如种族、性别等）。相关变量的设定如下——

表 1：解释变量的设定

|  |  |
| --- | --- |
| **变量** | **含义** |
|  | 为第i位学生的加权平均绩点 |
|  | 代表第i位学生是否修读了数学高级课程，为虚拟变量，如果修读了为1，否则为0 |
|  | 代表第i位学生是否修读了英语高级课程，为虚拟变量，如果修读了为1，否则为0 |
|  | 代表第i位学生是否修读了数学高级课程或英语高级课程，为虚拟变量，如果是为1，否则为0 |
|  | 代表第i位学生的母语是不是英语，为虚拟变量，如果不是为1，否则为0 |
|  | 代表第i位学生是不是亚洲人，为虚拟变量，如果是为1，否则为0 |
|  | 代表第i位学生是不是男性，为虚拟变量，如果是为1，否则为0 |
|  | 代表第i位学生是否参加过SAT准备课程，为虚拟变量，如果参加过为1，否则为0 |

接下来，我们对这些变量进行符号的预期和估计，相关结果和理由如下表所示——

表 2：对各个变量的符号预期和理由

|  |  |  |
| --- | --- | --- |
| **变量** | **预期符号** | **理由** |
|  | + | 加权平均绩点越高，学生成绩越好，在SAT考试中取得高分的可能性更大 |
|  | + | 修读了数学高级课程那么在SAT的数学考试中可能更胜一筹，其逻辑水平较高 |
|  | + | 修读了英语高级课程那么在SAT的英文理解中可能更胜一筹，其语言能力强 |
|  | + | 修读了数学高级课程或英语高级课程那么在SAT的相关科目中可能更胜一筹，语言或者逻辑水平更优。 |
|  | - | 如果母语是英语，那么他在SAT的考试中对语言的理解可能更胜一筹，其语言能力强；反之能力弱。 |
|  | - | 根据文献，白人的SAT成绩更高 |
|  | + | 根据文献，男性的SAT成绩更高 |
|  | + | 如果一个学生准备了SAT考试，那么他可能会复习得更充分，相比不复习得取得更高成绩 |

在以上变量中，我认为GPA、APMATH、APEND、ESL、PREP应该包含在模型中。个人认为种族和性别与其他无关，AP变量和APMATH、APEND表达意义相似，如果纳入可能出现多重共线性的问题。

1. 分析数据

我们打开Eviews软件，依次选择“File->Import ->Import From File”将相关表格数据导入Eviews软件中。

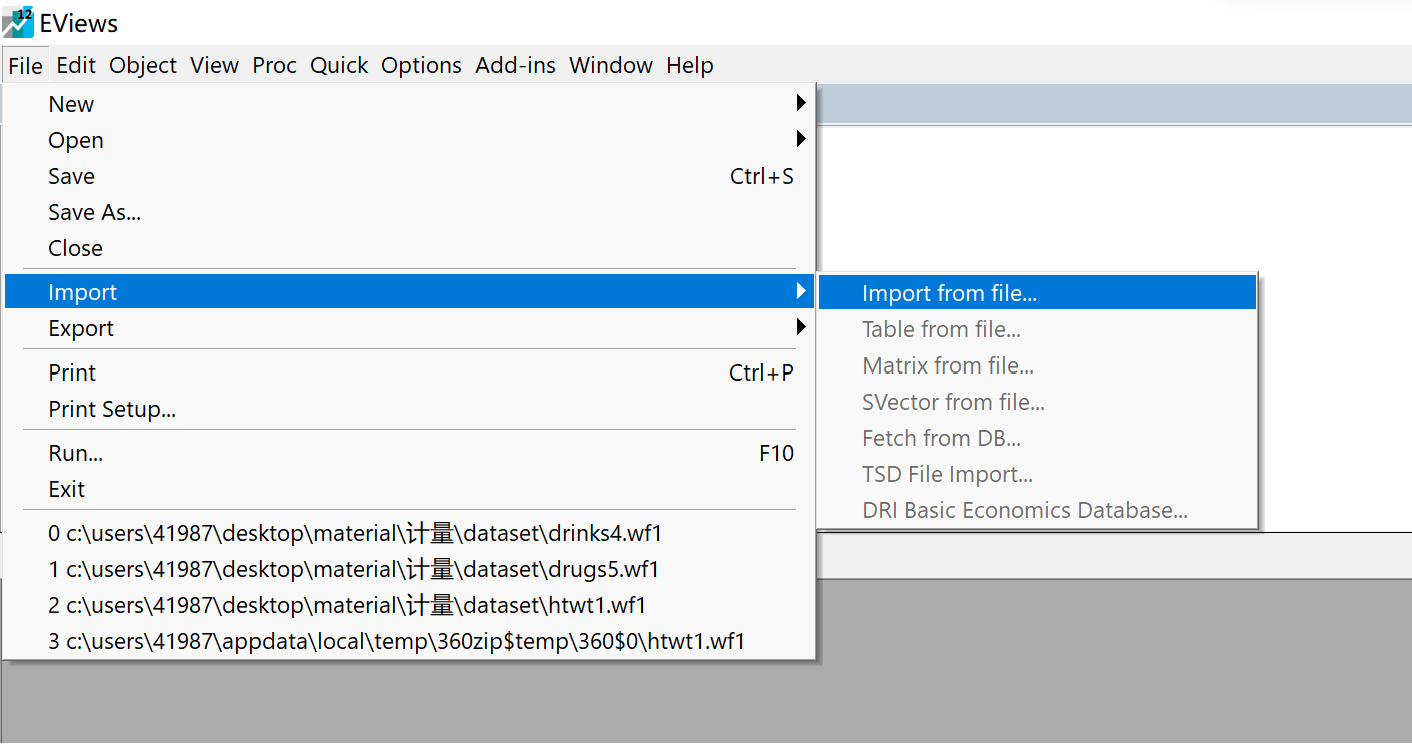


图 1: 将数据导入表格

首先将要查阅的数据选中，右击选择“OPEN->as GROUP”，即可展示导入的完整数据。在Group视窗中，选择“View->Descriptive Stats->Common Sample”即可查看数据的详细特征（包括最值、标准差、方差等），相关数据特性如下所示——

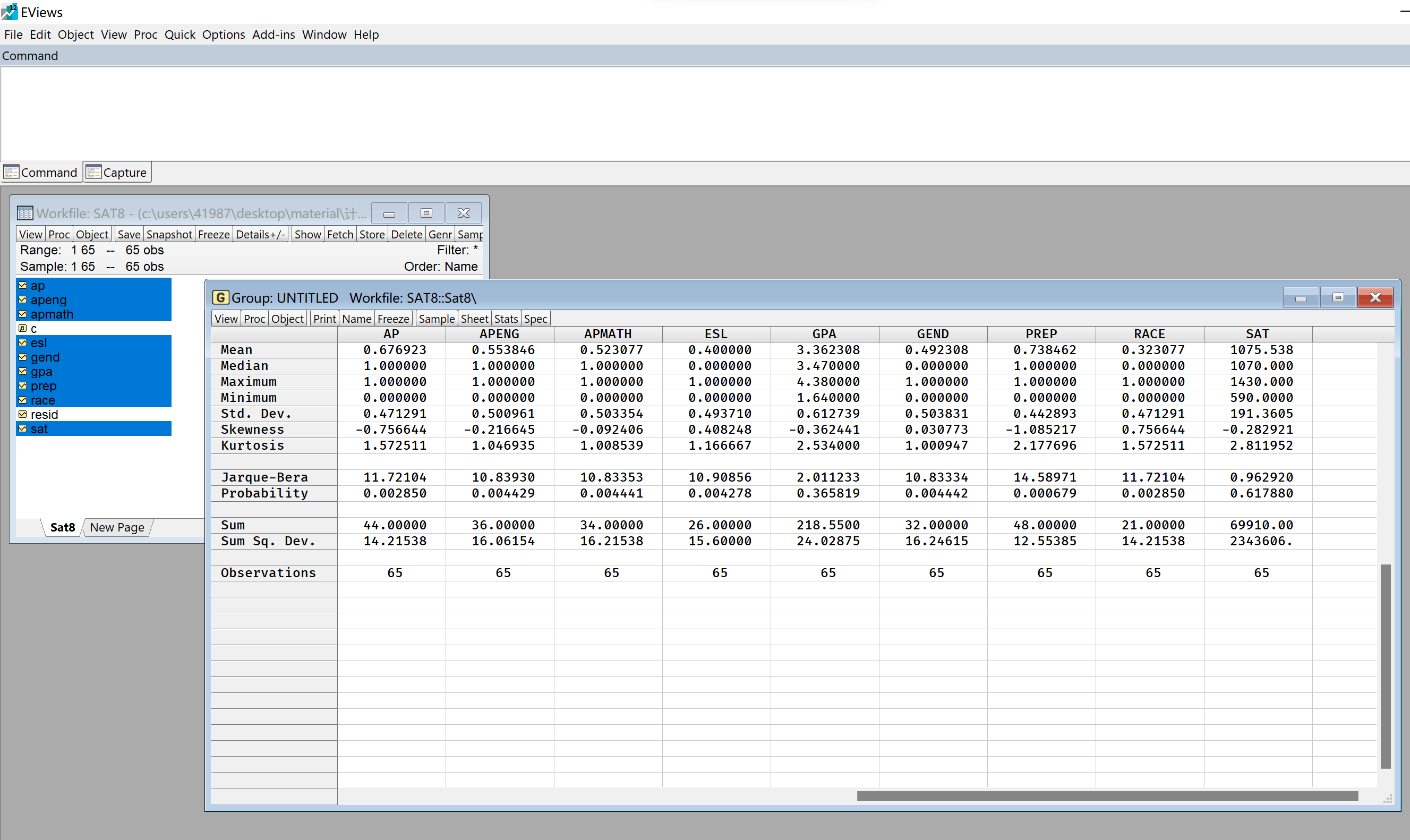


图 2：数据分析

可以看到各个变量的方差、均值范围均在合理区间内

随后，选择“views->Covariance Analysis…”进行相关系数分析。

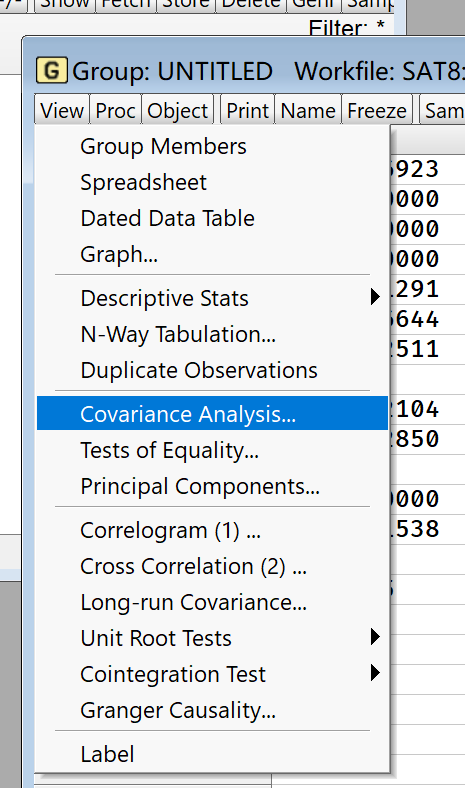


图 3：进行相关系数分析

相关系数结果

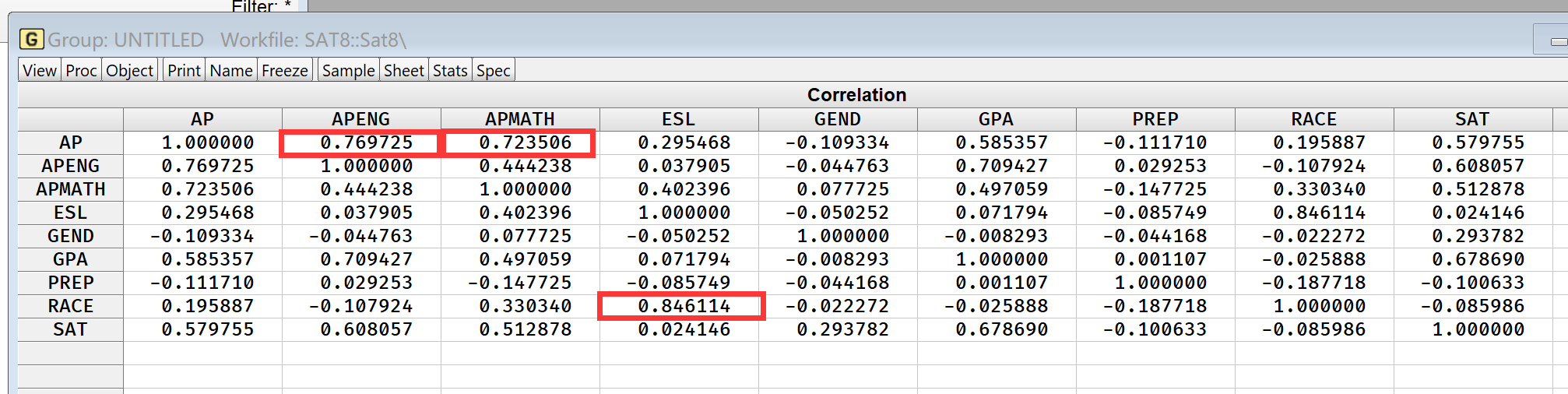


图 4：相关系数分析结果

可以看到在红圈部分，AP对APENG和APMATH相关系数比较高，因为AP（是否修读英语或数学）包含了APENG和APMATH，如果三者同时引入会引发一定的多重共线性，造成错误情况。也可以看到RACE和ESL相关度很高，因为白种人以英语为母语的居多，亚洲人一般不会以英语为母语，如果把这两者同时纳入模型中考虑，也可能出现多重共线性。

综上所述，我们首先探究GPA、APMATH、APEND、ESL、PREP纳入模型的情况

1. 选择ESL、PREP纳入模型（回归8-7）

输入命令“**LS SAT C GPA APENG APMATH ESL PREP**”进行回归，回归的结果如下图所示。

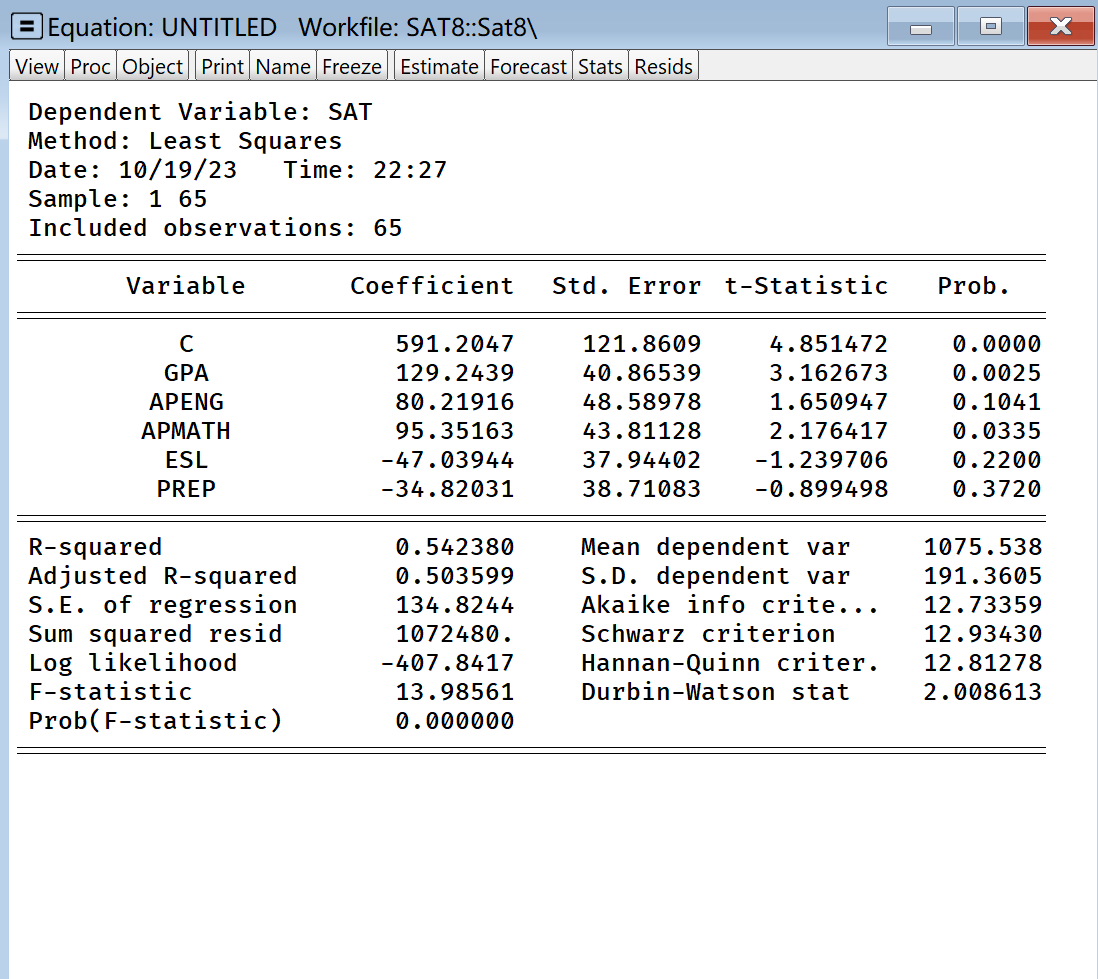


图 5：回归8-7的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

总体而言，这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、母语是否为英语以及是否参加过SAT准备课程之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。例如，参与数学课程和英语课程、英语母语学生以及较高的GPA预期会提高SAT成绩，而参加过SAT准备课程则可能对SAT成绩产生负面影响。这似乎和常识相违背。

1. **方程的拟合优度**

调整的判定系数大小为0.503599，也就是说，约50%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、是否为ESL学生、是否参加过SAT准备课程）解释。这个数值可以被视为具有中等水平的解释能力。

1. **参数的显著性**

自由度为59，假设显著性水平为5%。进行假设检验如下：

表 3：回归8-7的假设检验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | ESL | PREP |
| 预期符号 | + | + | + | - | + |
| t值 | 3.16 | 1.65 | 2.17 | -1.23 | -0.89 |
| 是否拒绝 | 拒绝 | 不拒绝 | 拒绝 | 不拒绝 | 不拒绝 |

可以看到，有三个参数并不显著（APENG和ESL、PREP），其余参数均显著。不显著的参数较多，说明方程的模型设定可能欠佳。

1. **方程的问题**

* **遗漏变量：**方程明显地出现了预期符号和实际不一致的情形，这很可能是遗漏变量而产生偏误。或许没有加入考虑的GEND的确对SAT有一定的影响。
* **不相干变量：**PREP可能是一个不相干变量，因为他在这五个变量中最不显著，可能其对SAT的成绩影响比较小。PREP的理论依据也不够充分，因为有可能是基础比较差的人才进行准备。

因为遗漏变量比不相干变量更加严重，因此我们考虑加入GEND变量。

1. 选择ESL、PREP、GEND纳入模型（回归8-12）

输入命令“**LS SAT C GPA APENG APMATH ESL PREP GEND**”进行回归，回归的结果如下图所示。

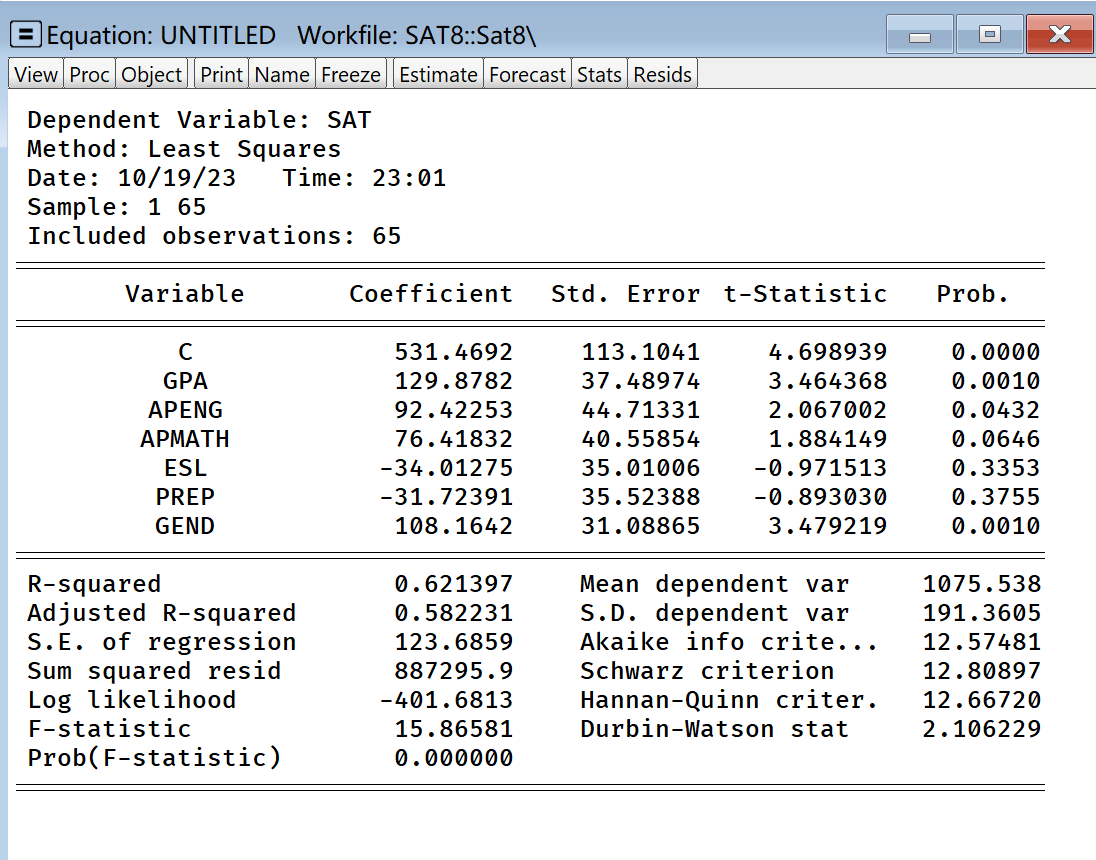


图 6：回归8-12结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、母语是否为英语、性别以及是否参加过SAT准备课程之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。例如，参与数学课程和英语课程以及较高的GPA预期、性别为男生、英语母语学生会提高SAT成绩，而参加过SAT准备课程则可能对SAT成绩产生负面影响。这似乎和常识相违背。

这个方程中，我们增加了性别这一变量，可以看到性别对SAT的影响很大。其他变量不变的情况下，性别为男生取得的SAT分数比性别为女生取得的SAT分数高108个单位。与文献符合。

1. **方程的拟合优度**

调整的判定系数大小为0.582231，也就是说，约58%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、是否为ESL学生、是否参加过SAT准备课程、性别）解释。这个数值可以被视为具有中等水平的解释能力。

增加了性别的模型相比原来没有增加性别的模型拟合程度更优。

1. **参数的显著性**

自由度为58，假设显著性水平为5%。进行假设检验如下：

表 4：回归8-12的假设检验

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | ESL | PREP | GEND |
| 预期符号 | + | + | + | - | + | + |
| t值 | 3.46 | 2.06 | 1.88 | -0.97 | -0.89 | 3.479 |
| 是否拒绝 | 拒绝 | 拒绝 | 不拒绝 | 不拒绝 | 不拒绝 | 拒绝 |

可以看到，有3个参数并不显著（ESL、PREP、APMATH）。其余参数均显著。但是相比之前的模型，显著性有所提高。

1. **方程的问题与诊断**

* **不相干变量：**PREP可能是一个不相干变量，因为他在这6个变量中仍然是最不显著，可能其对SAT的成绩影响比较小。PREP的理论依据也不够充分，因为有可能是基础比较差的人才进行准备。

我们使用Eviews软件判定PREP是否为不相干变量。

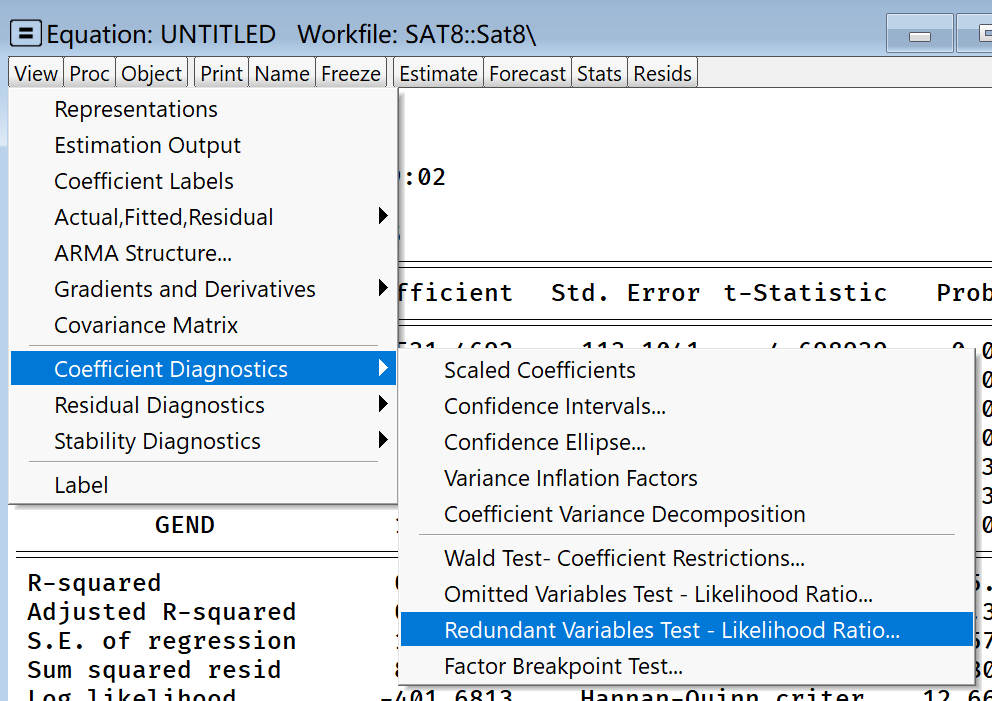


图 7：进行冗余变量检验

在弹出的提示框中输入“PREP”，表明我们检验的是PREP变量。

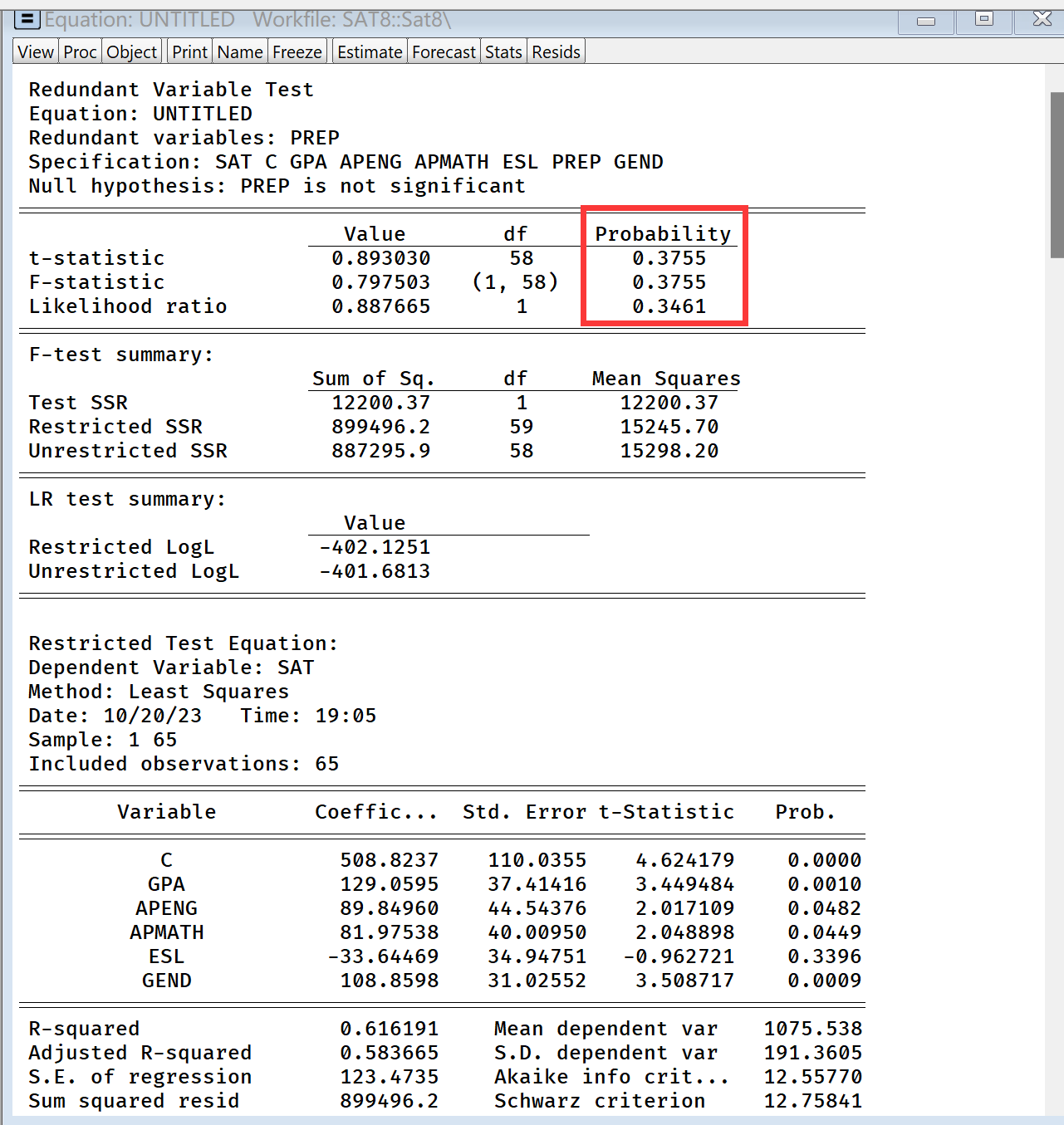


图 8：对PREP的冗余检验结果

从结果中，可以明显的看到，结果并不显著，因此我们接收假设——接受原假设，PREP是冗余变量。在下方删除了PREP的回归中，无论是单个参数的显著性还是整体拟合优度都有了或多或少的提升，我们将在下一个回归中进行具体地分析。

实际从理论分析，最初并不应该选择PREP作为变量。因为许多学生是由于之前参加SAT成绩较差，或者是由于预期会考的很差才去参加prep课程的。因此，即便perp课程很可能提高SAT分数，但是认为需要参加课程的学生本来就比其他同学考得差（假设方程中其他变量不变）。这两个效果会相互抵消，使得PREP成为不相关变量。

因此我们重新考虑，将方程中的PREP去掉，剔除这一冗余变量，保留其他变量。

1. 选择ESL、GEND纳入模型（回归8-6）

输入命令“**LS SAT C GPA APENG APMATH ESL GEND**”进行回归，回归的结果如下图所示。

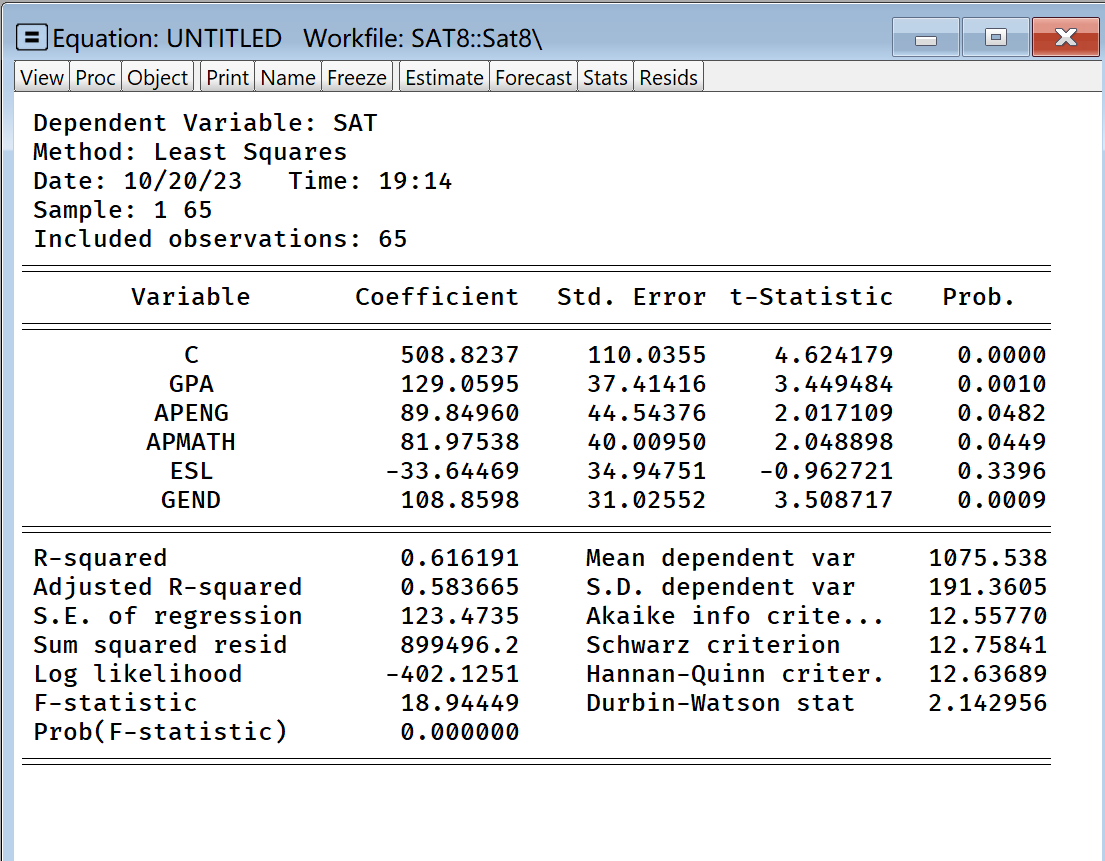


图 9：回归8-6的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、母语是否为英语、性别之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。

例如，参与数学课程和英语课程以及较高的GPA预期、性别为男生会提高SAT成绩。非英语母语学生产生负面影响。其他变量不变情况下，一个非英语为母语的人会比母语的人取得更低（低于约33个单位的成绩）的SAT成绩。

方程总体的符号预期与预估相一致。

1. **方程的拟合优度**

调整的判定系数大小为0.583665，也就是说，约58%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、是否为ESL学生）解释。这个数值可以被视为具有中等水平的解释能力。

剔除了是否准备SAT考试这一变量的模型相比原来纳入了是否准备SAT考试的模型拟合程度基本保持一致，两者具备相近的解释能力。

1. **参数的显著性**

自由度为59，假设显著性水平为5%。进行假设检验如下：

表 5：回归8-6的假设检验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | ESL | GEND |
| 预期符号 | + | + | + | - | + |
| t值 | 3.449 | 2.017 | 2.048 | -0.96 | 3.058 |
| 是否拒绝 | 拒绝 | 拒绝 | 拒绝 | 不拒绝 | 拒绝 |

可以看到，有1个参数并不显著（ESL）。其余参数均显著。

但是相比之前的模型，显著性有了较大程度上的提高。APMATH第一次由不显著变得显著，而且从整体来看F值较高（达到了18），说明去掉PREP后在方程的显著性提升上有一定效果。

1. **方程的合理性分析**

这个回归模型在多个方面都是合理的，理由如下——

* 首先，该回归模型的方程具有明确的理论意义，而且方程中各项参数的符号和预期完全一致。这意味着我们的假设与实际情况相符，从理论上支持了模型的合理性。
* 其次，该回归模型的拟合优度处于中等偏上的水平，具有一定的解释能力。调整的判定系数表明约50%的SAT成绩变异性可以由模型中的自变量解释，这说明我们的模型对数据拟合得较好。
* 此外，尽管其中一个变量不具有显著性，但是整体显著性水平较高。这意味着模型中的自变量作为一个整体对因变量的解释能力是显著的。

我们没有理由去担心理论上合理的变量（ESL），即使它们在预期的方向上存在轻微的不显著。

GEND参数估计值的绝对值似乎比文献中所记载的要大，在其他条件不变的情况下，男生比女生多取得108个单位的SAT成绩是个比较难以信服的数据。但是经过前面几个回归的分析，这个回归到目前为止是比较合理的。

我们还是继续探索，为了检验RACE的相关性，我们尝试纳入最后一个未被考虑的变量RACE。

1. 选择ESL、GEND、RACE纳入模型（回归8-13）

输入命令“**LS SAT C GPA APENG APMATH ESL GEND RACE**”进行回归，回归的结果如下图所示。

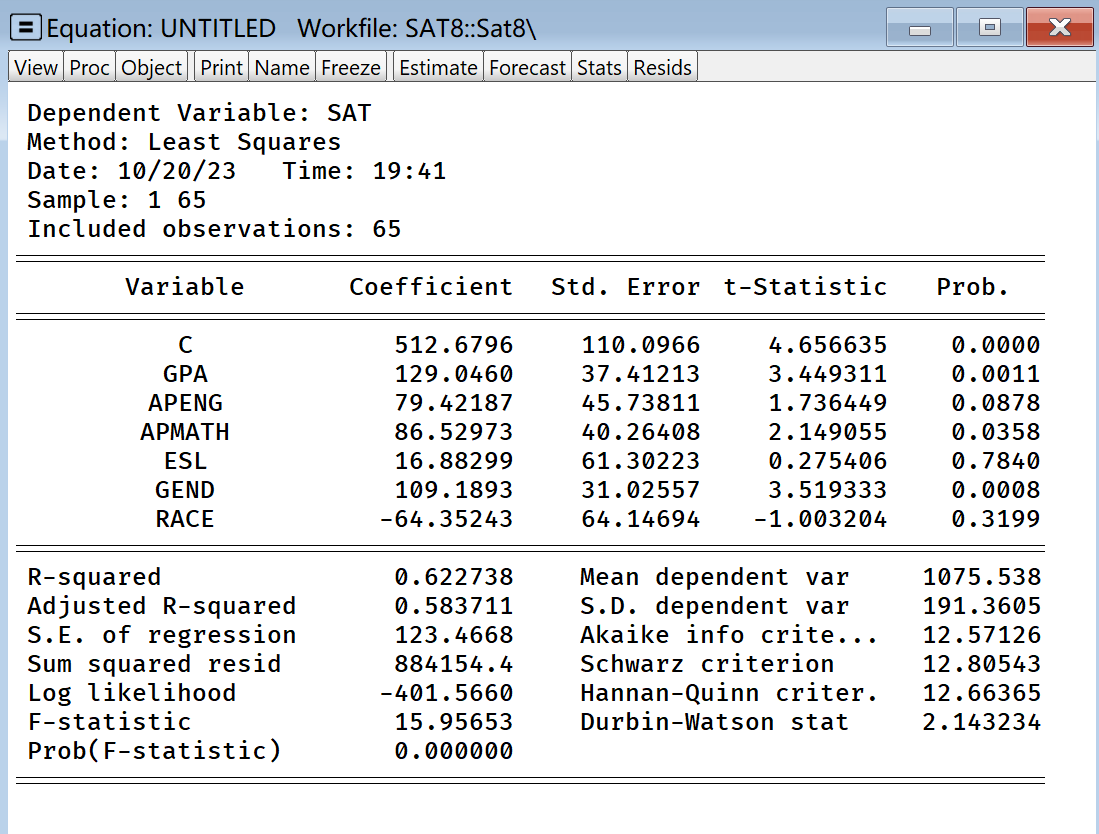


图 10：回归8-13的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、母语是否为英语、性别以及种族之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期的影响。例如，参与数学课程和英语课程以及较高的GPA预期、性别为男生学生均会提高SAT成绩，而白人学生的身份会对SAT成绩产生负面的影响。在其他变量不变的情况下，亚洲学生会比白人学生多考取64个单位的SAT成绩。但英语母语学生取得成绩比非英语母语学生更少，这与预期不符合。

除ESL外，以上和我们的预期基本一致，符合我们的初始预期。

这个方程中，我们增加了种族这一变量，其影响和文献描述一致（为负相关）。

1. **方程的拟合优度**

调整的判定系数大小为0.583711，也就是说，约58%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、是否为ESL学生、性别、种族）解释。这个数值可以被视为具有中等水平的解释能力。

增加了种族ESL变量的模型相比原来没有增加种族的模型拟合程度更优。

1. **参数的显著性**

自由度为58，假设显著性水平为5%。进行假设检验如下：

表 6：回归8-13的假设检验

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | ESL | GEND | RACE |
| 预期符号 | + | + | + | - | + | - |
| t值 | 3.449 | 1.736 | 2.149 | 0.275 | 3.519 | -1.003 |
| 是否拒绝 | 拒绝 | 不拒绝 | 拒绝 | 不拒绝 | 拒绝 | 不拒绝 |

可以看到，有3个参数并不显著（APENG、ESL、RACE）。其余参数均显著。但是相比之前的模型，显著性有所下降。说明方程可能存在一些问题。

1. **方程的问题与诊断**

* **多重共线性：**这个方程可能存在较为不完全的多重共线性。首先从相关系数图（图 3）可以看出，ESL和RACE的相关系数比较大（达到了0.8）；从理论上来讲，如果一个人是亚洲人（RACE=1）那么他非常有可能母语不是英文（ESL=1），这就导致了ESL和RACE之间存在不完全的共线性问题。

我们从几个方面来具体判定这个多重共线性问题——

首先，回归数据上，如图可以看到：APENG、ESL、RACE三个变量均表现为不显著，且ESL变量的预期符号与实际不符合。但是整体方程的显著性（F值）高，拟合优度（调整的判定系数）较高。这些特征符合多重共线性的问题

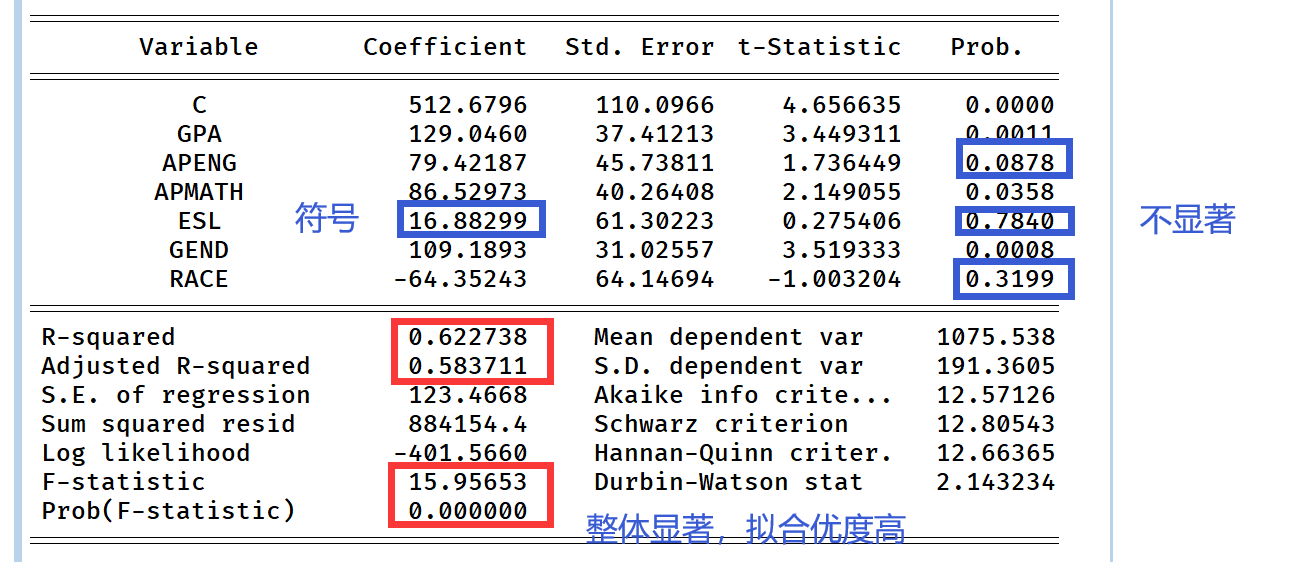


图 11：从回归数据分析回归8-13的多重共线性

其次，使用Eviews软件的相关系数判定ESL和RACE的共线性问题。

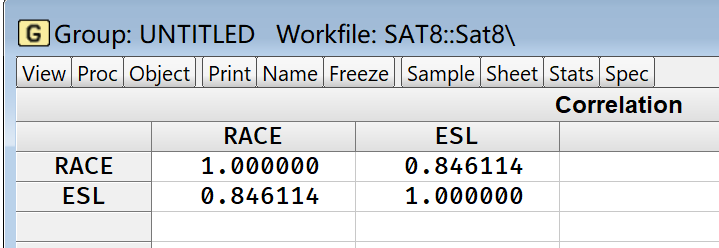


图 12：RACE和ESL的相关系数

可以看到ESL和RACE的相关系数高达0.846，其具有高度的正相关性。可以认为具有多重共线性。

再者，从VIF角度看。我们使用Eviews软件计算方差膨胀因子。依次点击View-> Coefficient Diagnostics-> Variance Inflation Factors.

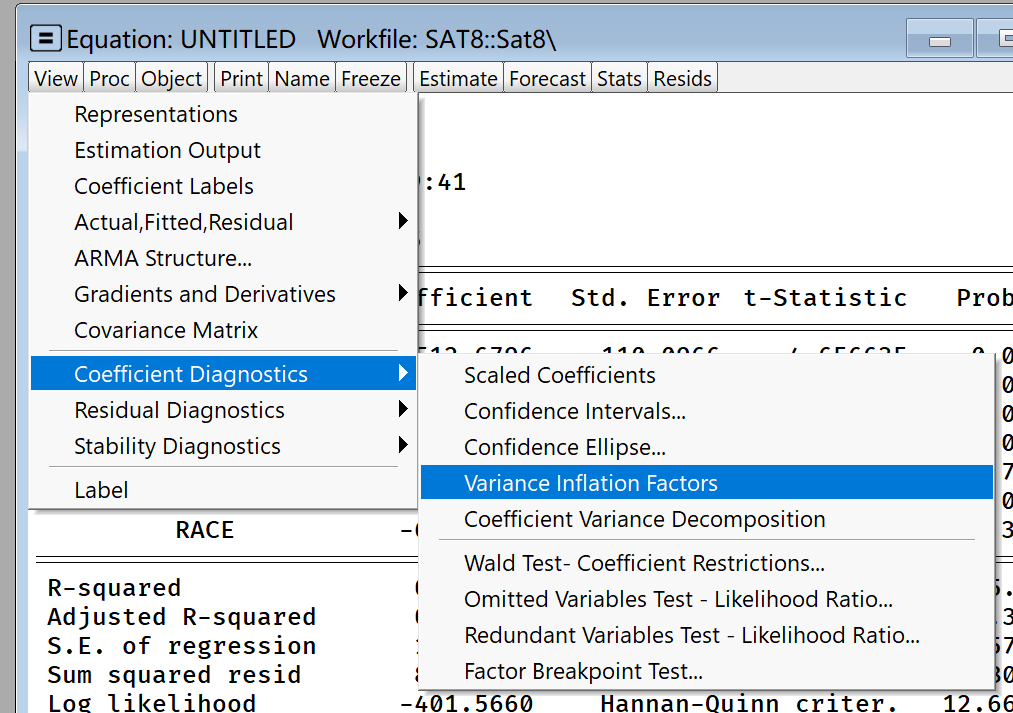


图 13：进行VIF分析的步骤

得到结果——

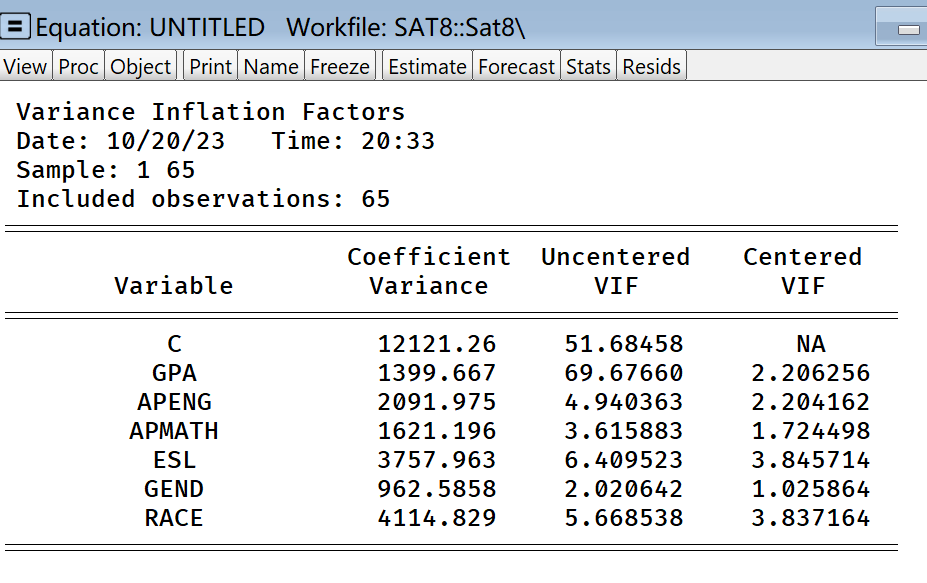


图 14：VIF分析结果

可以看到各项VIF的值处于3左右，最高未3.8，并非严重的多重共线性，但是仍然有一定的不完全多重共线性。

这种情况从理论角度来讲是因为ESL和RACE的定义相关（并且两者的样本相关参数很高），使得二者看起来像是重复变量。

不完全的多重共线性虽然不会造成OLS估计值的有偏，但是会使得估计量的标准误非常大，即估计的精度很小。

对于出现共线性的情况，我们可以选择剔除一个变量。我认为最好剔除RACE变量，因为ESL（是否以英语为母语）相比种族可能对SAT的成绩影响更大。其更能直观反应堆SAT的影响，剔除RACE的情况已经在回归8-6（第5步）进行了演示分析，因此接下来我们尝试去掉ESL变量。

1. 选择GEND、RACE纳入模型（回归8-10）

输入命令“**LS SAT C GPA APENG APMATH GEND RACE**”进行回归，回归的结果如下图所示。

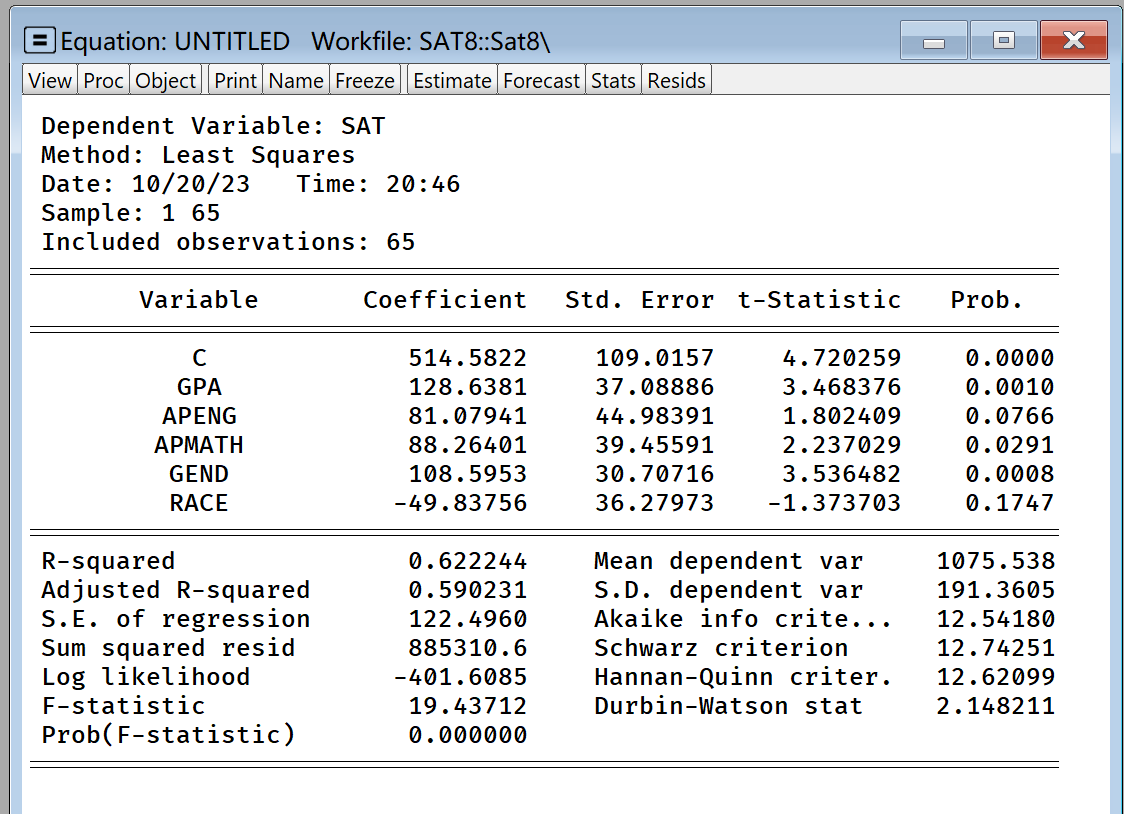


图 15：回归8-10的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、性别、种族之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。

例如，参与数学课程和英语课程以及较高的GPA预期、性别为男生会提高SAT成绩。亚洲种族产生负面影响。其他变量不变情况下，一个亚洲人会比白人取得更低（低于约49个单位的成绩）的SAT成绩。

方程总体的符号预期与预估和文献相一致。

1. **方程的拟合优度**

调整的判定系数大小为0.590231，也就是说，约59%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、性别、种族）解释。这个数值可以被视为具有中等水平的解释能力。

这个方程的拟合优度首次达到了59%，是目前为止拟合优度最高的一个方程。

1. **参数的显著性**

自由度为59，假设显著性水平为5%。进行假设检验如下：

表 7：回归8-10的假设检验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | GEND | RACE |
| 预期符号 | + | + | + | + | - |
| t值 | 3.46 | 1.80 | 2.23 | 3.56 | -1.37 |
| 是否拒绝 | 拒绝 | 不拒绝 | 拒绝 | 拒绝 | 不拒绝 |

可以看到，有2个参数并不显著（APENG、RACE）。其余参数均显著。

但整体显著性水平（F值）达到了19，可以认为具备一定的显著性。

1. **方程的合理性和问题分析**

遗漏变量和存在不相干变量这两个问题都有可能，因为方程仍然有不少参数不具备显著性，但是也不是特别严重。并没有征兆表明出现偏误，到目前为止，这个方程可以认为是合理的。理由如下——

首先，该回归模型的方程具有明确的理论基础，并且模型中各项参数的符号和预期一致。这表明我们的假设与实际情况相符，从理论角度上支持了模型的合理性。

其次，该回归模型的拟合优度属于中等偏上水平，具有一定的解释能力。通过调整后的判定系数可以看出，约59%的SAT成绩的变异性可以由模型中的自变量进行解释，这意味着模型对数据的拟合程度较为良好。

此外，尽管其中两个变量在统计上不显著，但整体的显著性水平较高。我们无需对理论上合理的变量（如RACE和APENG）产生过多担忧，即使它们在预期方向上存在轻微的非显著性。

综上所述，该回归模型在理论基础、拟合优度以及整体显著性方面具备一定的价值，具备一定的合理性。

因为RACE不显著，虽然比较轻微，不太严重，但本着完备的原则，我们继续探究去掉RACE，只保留GEND的情形，或许会有不一样的收获。

1. 选择GEND纳入模型（回归8-3）

输入命令“**LS SAT C GPA APENG APMATH GEND** ”进行回归，回归的结果如下图所示。

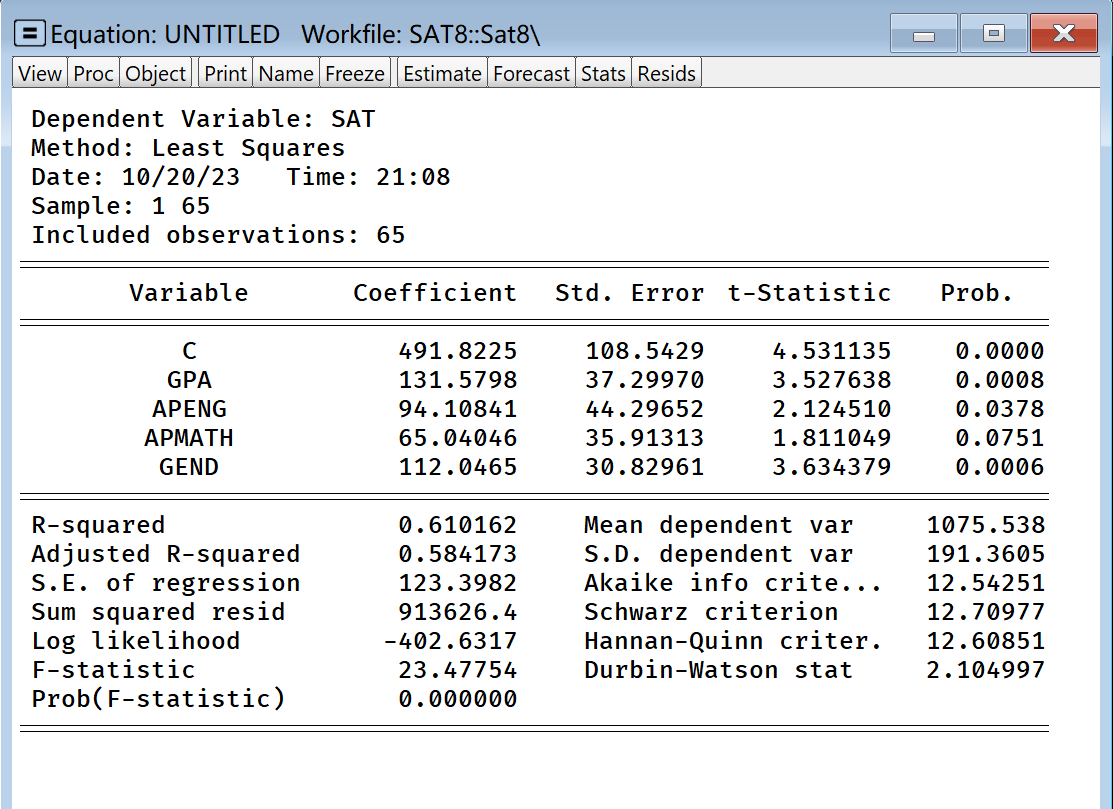


图 16:回归8-3的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、是否参与AP数学课程、性别之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。

例如，参与数学课程和英语课程以及较高的GPA预期、性别为男生会提高SAT成绩。其他变量不变的情况下，性别为男生取得的SAT分数比性别为女生取得的SAT分数高112个单位。与文献符合。

方程总体的符号预期与预估和文献相一致。

1. **方程的拟合优度**

调整的判定系数大小为0.58417，也就是说，约58%的SAT成绩的变异性可以由模型中的自变量（平均绩点、是否参与英语课、是否参与数学课、性别）解释。这个数值可以被视为具有中等水平的解释能力。

相比上一回归，这个方程的拟合优度略有下降。

1. **参数的显著性**

自由度为60，假设显著性水平为5%。进行假设检验如下：

表 8：回归8-3的假设检验

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | GEND |
| 预期符号 | + | + | + | + |
| t值 | 3.52 | 2.12 | 1.81 | 3.63 |
| 是否拒绝 | 拒绝 | 不拒绝 | 拒绝 | 拒绝 |

可以看到，有1个参数并不显著（APMATH）。其余参数均显著。

但整体显著性水平（F值）达到了23，为目前最高的F值，可以认为具备较强的显著性。

1. **方程的合理性分析**

遗漏变量仍然是一种可能的情况，毕竟这个方程包含的参数个数比其他回归少。虽然文献表明RACE有相当程度的理论意义，但是其就是是不是应该纳入方程？下面，我们来探究RACE到底是不是一个遗漏变量。

在Eviews中依次选择view-> Coefficient Diagnostics-> Redundant Variables Test -Likelihood Ratio...进入遗漏变量检测。

在弹出的窗口输入RACE，表明我们检测RACE（是否亚洲学生）是否是一个遗漏变量。相关结果如下。

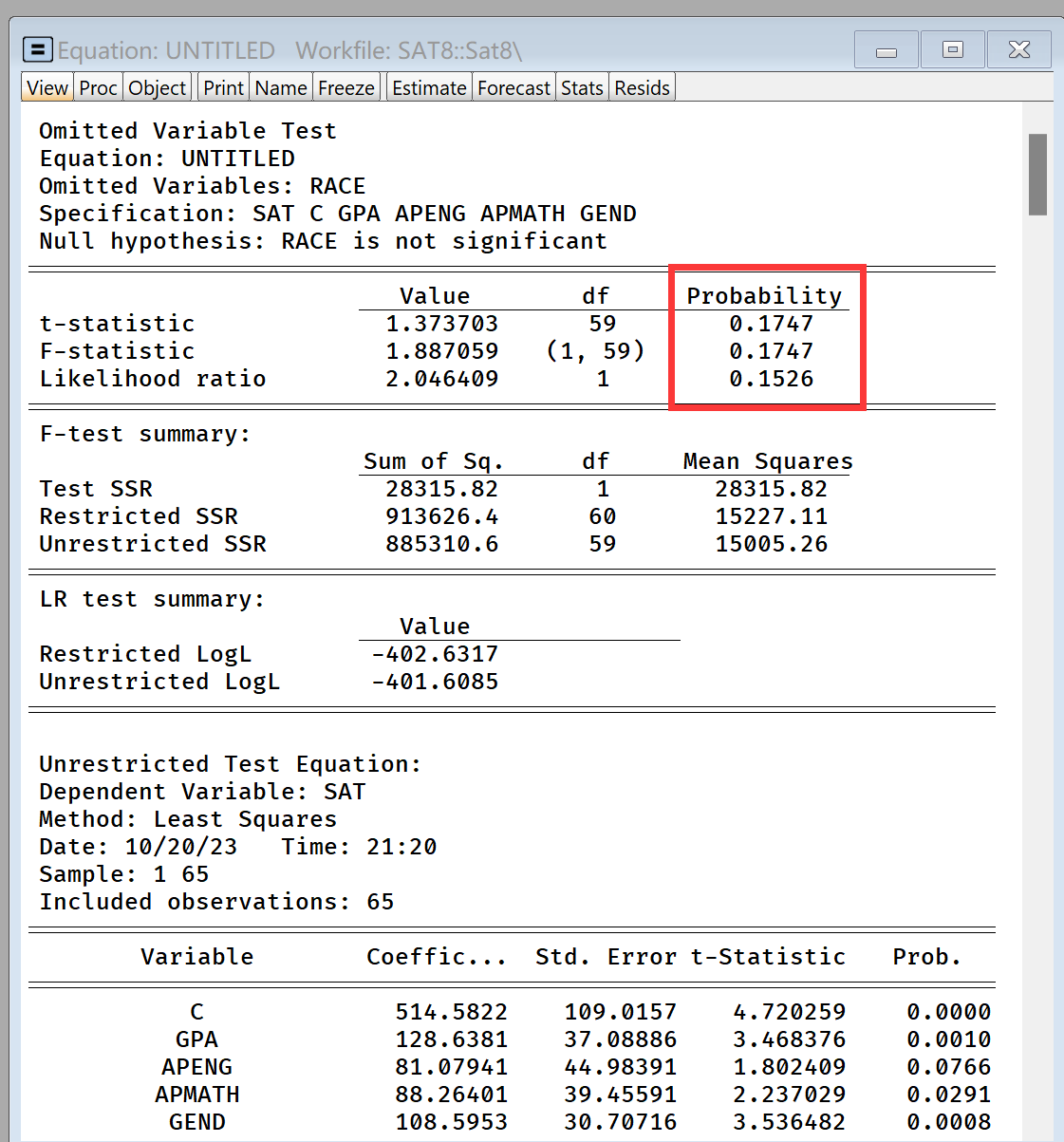


图 17：检验RACE是否是遗漏变量的结果

我们看到，t统计量已经为1.37，我们无法拒绝“RACE不是遗漏变量”的原假设。说明RACE并非一个典型的遗漏变量。尽管他存在一定的理论意义和文献依据。

从其他角度看，这都是一个比较合理的方程——

* 该回归模型的方程具有明确的理论基础，并且模型中各项参数的符号和预期一致。我们的假设与实际情况相符，从理论角度上支持了模型的合理性。
* 模型表现出一定的解释能力，拟合优度高。通过调整后的判定系数的分析，我们得知约有58%的SAT成绩的变异性可以被模型中的自变量进行解释，这意味着该模型对数据的拟合程度较为优良。
* 此外，尽管在统计上存在一个变量不显著的情况，然而在整体显著性检验中，该回归模型展现出较高的显著性水平。特别值得注意的是，整体的F值为23，这是一个相当可观的数值，显示出模型的显著性较为突出。

综上所述，该回归模型在理论基础、拟合优度以及整体显著性方面具备一定的价值，具备一定的合理性。

1. 选择AP、PREP、RACE纳入模型（回归8-18）

我们还有一个变量没有纳入考虑之中，那就是AP，AP可以看作APMTAH和APENG两个变量的线性替代。其反应学生参与英语或数学课堂的情况。接下来，我们以AP代替APMTAH和APENG，来探究其对模型的影响。

输入命令“**LS SAT C GPA AP PREP RACE** ”进行回归，回归的结果如下图所示。

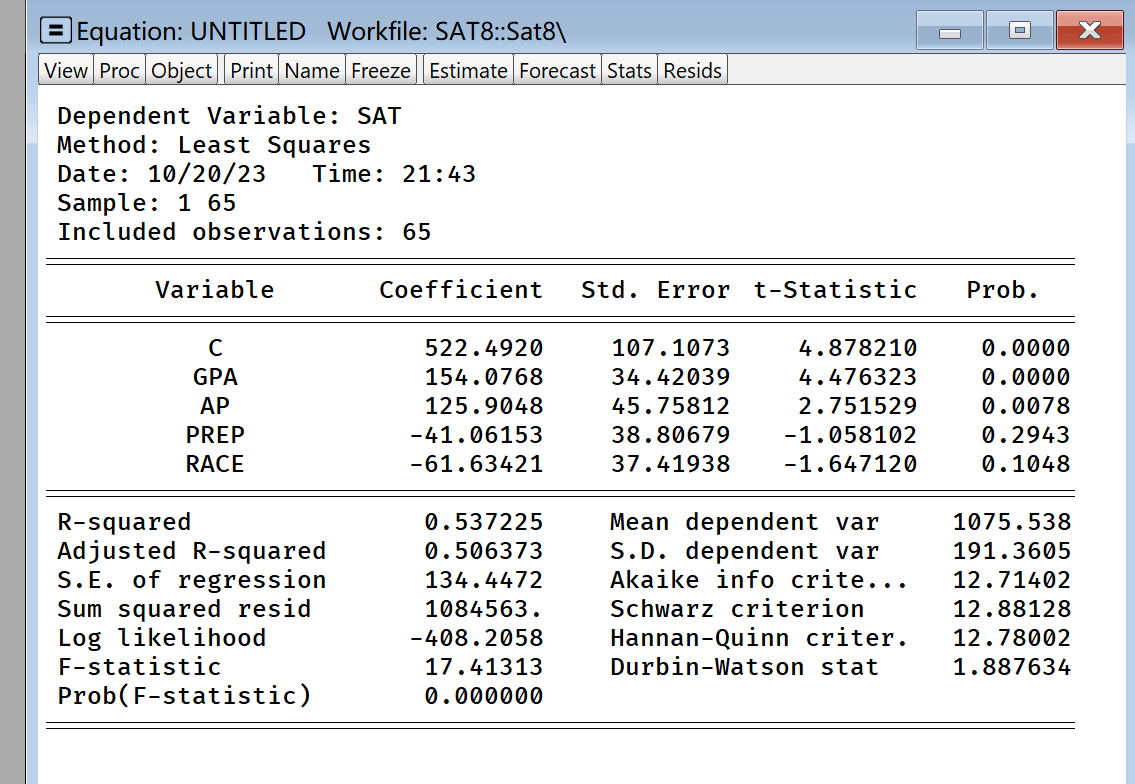


图 18:回归8-18的结果

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程或数学课程、是否准备和是否亚洲人之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。

例如，参与数学课程和英语课程以及较高的GPA预期、白人会提高SAT成绩。其他变量不变的情况下，白人取得的SAT分数比性别为亚洲人取得的SAT分数高61个单位。与文献符合。但是，其他变量不变的情况下，准备了SAT考试的人会比不准备的SAT成绩低41个单位。这与日常经验不符合。

1. **方程的拟合优度**

调整的判定系数大小为0.506374，也就是说，约50%的SAT成绩的变异性可以由模型中的自变量解释。这个数值可以被视为这个模型具有中等偏下水平的解释能力。

这一模型的解释能力远远低于其他回归（其他的基本稳定在58%左右），因此，这个模型的解释效力是值得商榷和考量的，这表明可能是模型存在了一些问题。

1. **参数的显著性**

自由度为60，假设显著性水平为5%。进行假设检验如下：

表 9：回归8-18的假设检验

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 变量 | GPA | AP | RACE | PREP |
| 预期符号 | + | + | - | + |
| t值 | 4.47 | 2.75 | -1.64 | -1.05 |
| 是否拒绝 | 拒绝 | 拒绝 | 不拒绝 | 不拒绝 |

可以看到，有2个参数并不显著（PREP和RACE）。其余参数均显著。整体显著性水平（F值）为17。从以上两个角度看，方程的显著性都比不上其他合理的方程。但是AP和GPA的显著性水平比其他模型都要略高一些。

1. **方程的合理性分析**

* **变换的不合理性**

从理论和结果来看，把APMATH和APENG用AP替代似乎有一点画蛇添足，原始的情况下将APMATH和APENG同时纳入考量的情况下，方程的拟合优度、显著性水平都比用AP替代要好。

且APMATH和APENG并不存在严重的共线性（他们的相关系数才达到0.4），因此，同时纳入APMATH和APENG是可以的。用AP替代反而多余了。

综上所述，把APMATH和APENG用AP替代这种变换是一个不太好的模型变换方案。

* **不相干变量**

同时，PREP的符合和预期相反，这警示我们PREP也可能是个多余变量。因为许多学生是由于之前参加SAT成绩较差，或者是由于预期会考的很差才去参加prep课程的。

因此，即便prep课程很可能提高SAT分数，但是认为需要参加课程的学生本来就比其他同学考得差（假设方程中其他变量不变）。这两个效果会相互抵消，使得PREP成为不相关变量。

综上，方程在解释变量选择、不相干变量上均存在问题。不是一个合理方程。

1. 选择ESL、GEND、PREP、RACE纳入模型（回归8-16）

最后，让我们检验一下将所有参数纳入模型的效果。

输入命令“**LS SAT C GPA APENG APMATH ESL GEND PREP RACE**”进行回归，回归的结果如下图所示。

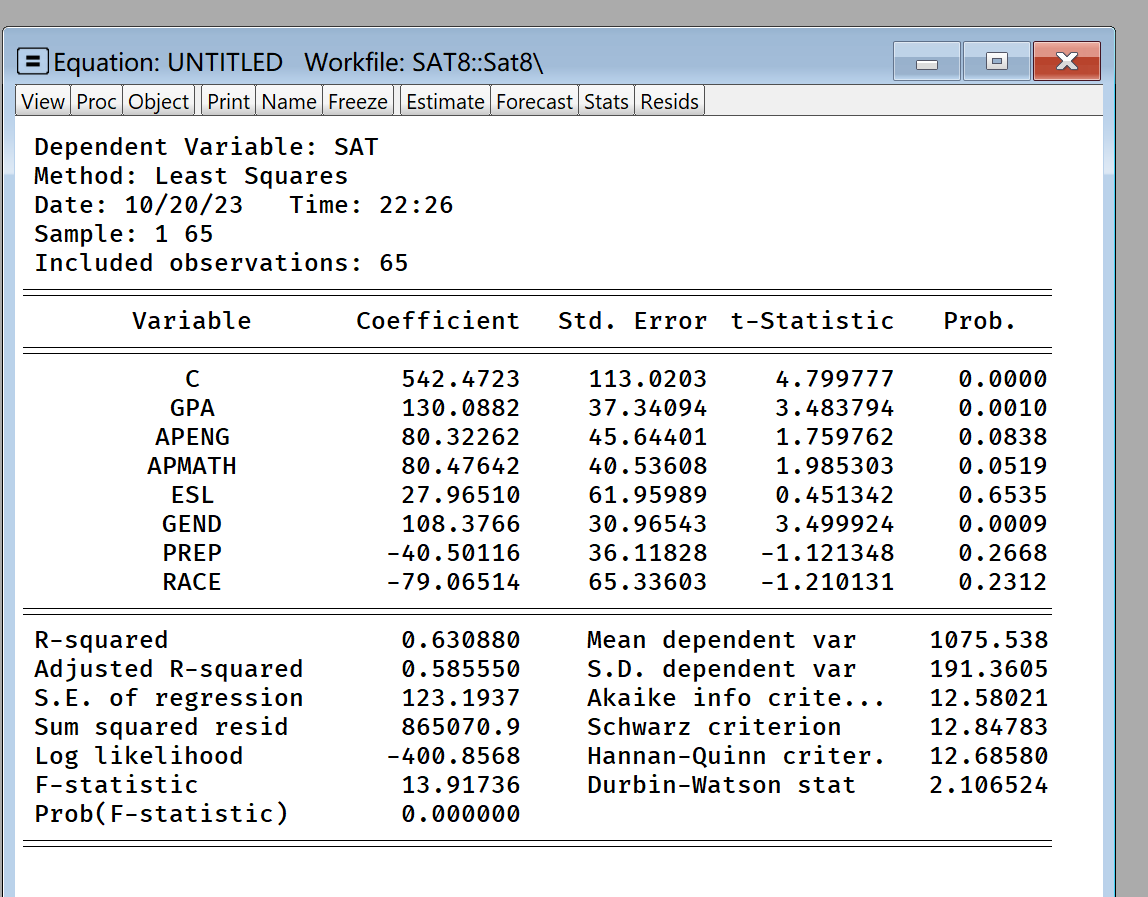


图 19：回归8-16的结果分析

即得出的拟合回归方程为：

1. **方程的经济意义**

这个回归方程描述了SAT成绩与GPA、是否参与AP英语课程、AP数学课程、是否英语母语、性别、是否准备和是否亚洲人之间的关系。

系数告诉我们在其他变量保持不变的情况下，这些因素对于SAT成绩的预期影响。

例如，参与数学课程和英语课程以及较高的GPA预期、性别为男会提高SAT成绩，但是准备了SAT考试和为亚洲人会对成绩有负影响。

值得注意的是ESL和PREP变量的符号与我们预期的不一致。在其他变量不变的情况下，英语非母语的学生的SAT分数比英语母语的学生取得的SAT分数高61个单位。这与日常经验不符合。

1. **方程的拟合优度**

调整的判定系数大小为0.5855，也就是说，约58%的SAT成绩的变异性可以由模型中的自变量解释。这个数值可以被视为这个模型具有中等水平的解释能力。

这一模型的解释能力在其他回归中处于平均水平。

1. **参数的显著性**

自由度为57，假设显著性水平为5%。进行假设检验如下：

表 10：回归8-16的假设检验

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 变量 | GPA | APENG | APMATH | ESL | GEND | PREP | RACE |
| 预期符号 | + | + | + | - | + | + | - |
| t值 | 3.48 | 1.75 | 1.98 | 0.45 | 3.49 | -1.12 | -1.21 |
| 是否拒绝 | 拒绝 | 不拒绝 | 不拒绝 | 不拒绝 | 拒绝 | 不拒绝 | 不拒绝 |

可以看到，有5个参数并不显著。整体显著性水平（F值）为13，处于较低水平。从以上两个角度看，方程的显著性都比不上其他合理的方程。是这8个回归中显著性最低的。可能存在较大问题

1. **方程的问题与诊断**

* **多重共线性**

RACE和ESL可能存在一定的共线性，因为他们表达的意义近乎相同（这在之前的回归8-13中分析过），如果一个人为亚洲人，那么他大概率并非母语为英语。我们使用Eviews软件来分析其VIF(方差膨胀因子)。依次点击View-> Coefficient Diagnostics-> Variance Inflation Factors.

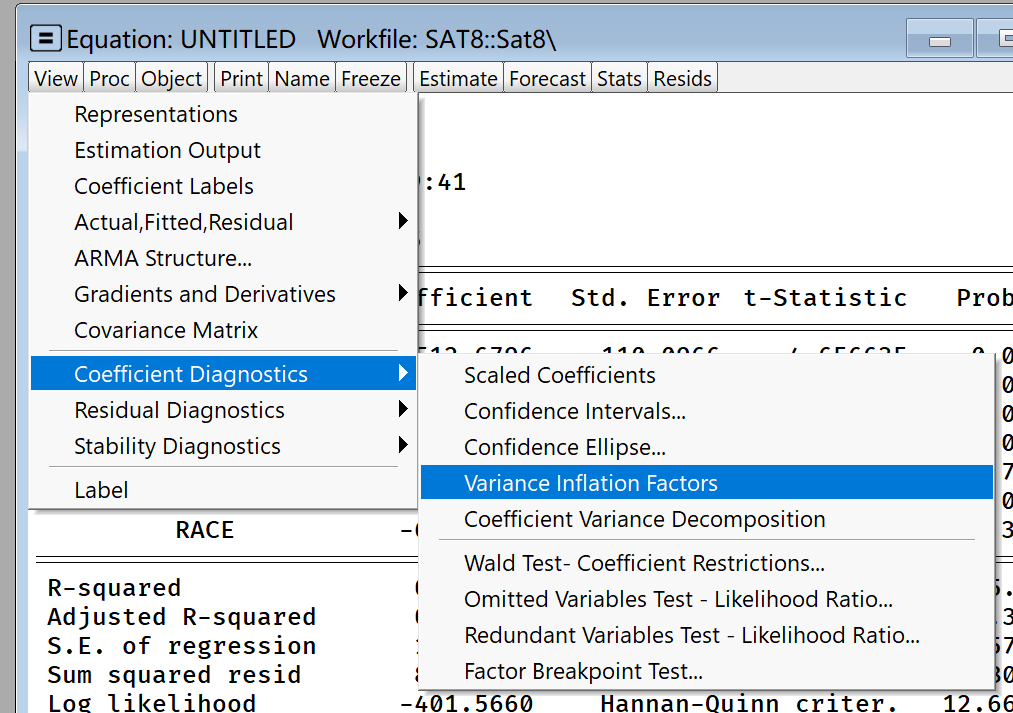


图 20：VIF分析的步骤

得到结果——

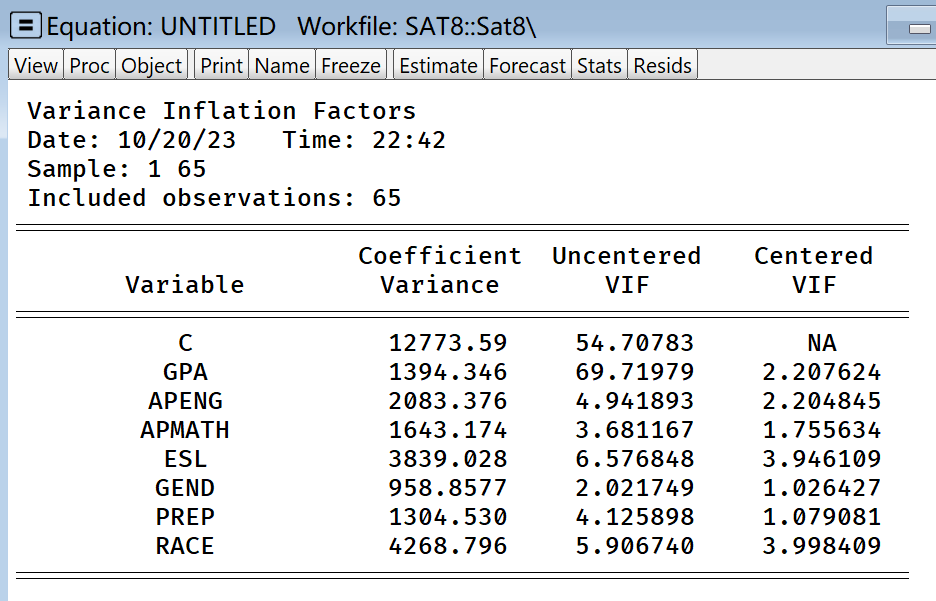


图 21:VIF结果

可以看到，最大的VIF值约为3.9984，小于5.说明多重共线性并不严重。RACE和ESL只是具备不完全的多重共线性。

在一般情况下，这种多重共线性是可以省略的

* **不相干变量**

方程中因为包含了全部的参数，所以也有可能存在不相干变量的问题。在之前的回归中，我们发现PREP是一个不相干变量，我们在这里不再赘述。我们探究RACE是否是一个不相干变量。

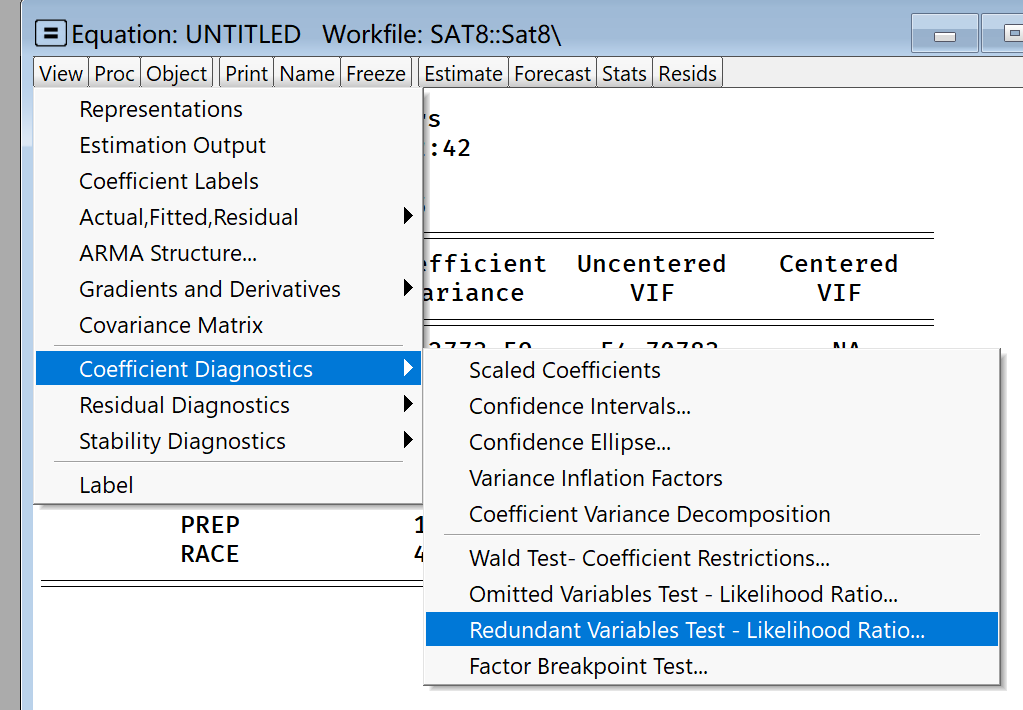


图 22：探寻是否不相干变量的步骤

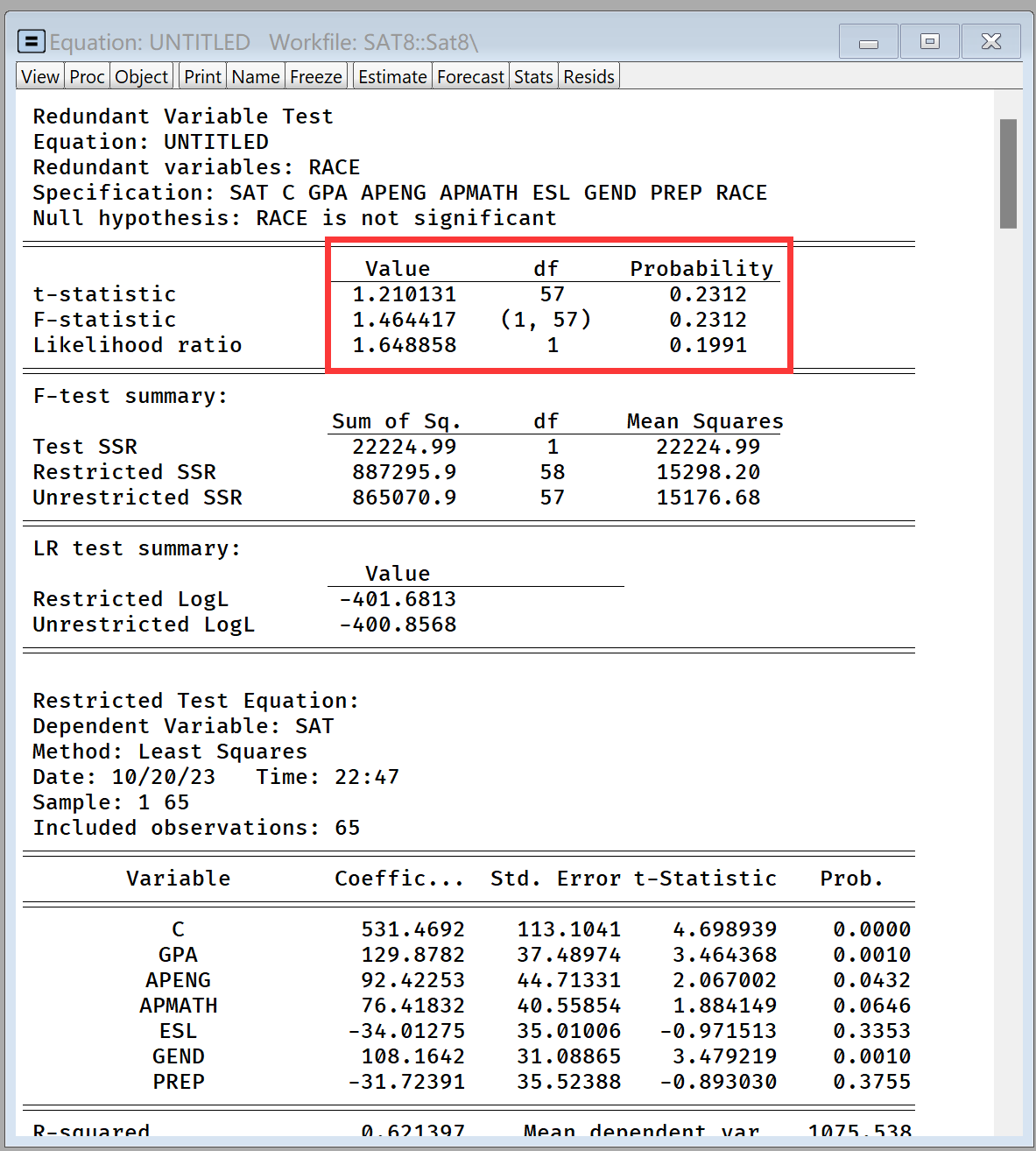


图 23:探究RACE是否不相干变量的结果

可以看到，t统计量为1.21，因此，我们不能拒绝原假设“RACE是不相干变量”。同时，RACE的确具备理论意义，那么可以得出结论，RACE应该保留在模型中。

综上所述，这个模型具备一定的多重共线性，同时，存在不相干变量（如PREP）。因此也不是一个合理的模型。

**九、实验结论：**

在本次实验中，我们对SAT考试成绩的7个主要影响因素（包括GPA成绩、AP数学课程参与、AP英语课程参与、是否英语母语、性别、种族、是否准备SAT）进行模型假定、数据分析、模型诊断。逐个回归，探究各个模型存在的问题和合理性。

* **最佳的模型**

表 11：8个模型的回归结果分析总结

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **编号** | **符号预期** | **显著性** | **拟合优度** | **存在问题** | **合理性** |
| 8-7 | 1个违背 | F值13.98，较低 | 调整判定系数50.35%，较低 | * 遗漏变量 * 不相干变量 | NO |
| 8-12 | 1个违背 | F值15.86，一般 | 调整判定系数58.22%，中等 | * 不相干变量 | NO |
| 8-6 | 0个违背 | F值18.94，良好 | 调整判定系数58.36%，中等 | / | YES |
| 8-13 | 1个违背 | F值15.95，一般 | 调整判定系数58.37%，中等 | * 不完全的多重共线性 | NO |
| 8-10 | 0个违背 | F值19.43，优良 | 调整判定系数59.02%，中等偏上 | / | YES |
| 8-3 | 0个违背 | F值23.47，优秀 | 调整判定系数58.41%，中等 | / | YES |
| 8-18 | 1个违背 | F值17.41，一般 | 调整判定系数50.06%，较差 | * 参数变换不合理（AP代替APMATH和APENG多余） * 不相干变量 | NO |
| 8-16 | 2个违背 | F值13.91，较差 | 调整判定系数58.55%，中等 | * 不完全的多重共线性 * 不相干变量 | NO |

8个模型的结果如上表所示，可以看到，方程8-6、8-10、8-3均可以视作一个合理的对SAT的回归模型。他们不存在较为严重的多重线性，在遗漏变量和不相干变量上也没有特别的体现，同时，在方程的显著性、符号的理论意义、拟合优度上面均具备较高水平。

在这三个合理方程中，综合理论、拟合优度、显著性水平等多方面来看，我认为8-10的回归结果是最好的。

其回归结果如下——

从理论来讲，GPA、数学课程和英语课程参与程度这类学术性指标确实和GPA相关；而性别、种族这类其他变量在文献研究中也有提及其与SAT的相关性。具备理论意义。

从显著性水平看，这个方程的显著性水平比较高，整体显著性水平达到了19，比较可观。

从拟合优度来看，调整的判定系数大小为0.590231，具备一定的解释能力。

在这个回归中，有意规避了RACE和ESL两个可能存在多重共线性的问题，同时，去掉了PREP这个不相干变量，因为即便prep课程很可能提高SAT分数，但是认为需要参加课程的学生本来就比其他同学考得差（假设方程中其他变量不变）。这两个效果会相互抵消，使得PREP成为不相关变量。

* **启示与思考**

在本次实验中，通过对8个回归模型的逐个分析，我得到了如下三点启示：

1. **不能因为变量的t统计量较低，就简单地认为此变量是不相干变量。**

例如：在我们看来，变量 ESL或RACE包含在方程中有充分的理由，预期的t统计量略微不显著，并不足以使我们重新思考隐含理论是否合适。

1. **注意寻找潜在的多重共线性，注意从变量的实际数据出发。**

一般来说，变量ESL与RACE不是多余的，但在本次实际情况下，我们的种族简单拆分为亚洲和非亚洲，很显然，亚洲的学生以英语母语的情况较少。因此这就存在了多重共线性。这样的话，一旦其中一个变量包含在了方程中，那么就不应该考虑另一变量。

1. **不能简单根据过往经验判断变量是否具备较大意义，应周全考虑**

从表面上看，变量PREP是一个有影响力的变量，因为根据常识PREP课程可以提高修读了该课程学生的SAT测试成绩。

但是问题是学生修读PREP课程的决定与他以前的（或者预期的）SAT测试成绩有关。往往成绩差的学生更倾向修读PREP课程，这些课程会使他们的成绩达到与那些没有选修 PREP课程的学生的成绩不相上下。因而，该变量对SAT成绩没有显著的影响。

这似乎与我们的过往经验相违背，但确实有一定道理，并且在实际的回归中有所体现。因此，过往的经验只能做简单判断，我们需要深入考虑、周全考虑。

**报告评分：**

**指导教师签字：**