计量经济学Eviews实验指导书

Lab 6 异方差模型的诊断和矫正

胡华平

2018/4/13

Table of Contents

# 异方差的诊断和矫正

## 实验目的及要求

* **目的**：掌握异方差问题的检验与处理方法。
* **要求**：在老师指导下完成计量经济模型的异方差检验，并对存在异方差的模型进行修正，最终得到正确的分析结果。

## 实验原理

* 对于不同的样本点，随机误差项的方差不再是常数，而互不相同，则认为出现了异方差性。
* 异方差的实质表现为随机误差项的方差随着解释变量（引起异方差的解释变量）观测值的变化而变化。
* 对于出现异方差的原模型主要采用校正其异方差，再对校正后的模型采用普通最小二乘法估计。

## 实验内容

1. 采用最小二乘法建立主回归模型
2. 侦查模型是否存在异方差问题
   1. 非正式检验方法（图形检验法）：
      * 残差趋势图（dot plot）
      * 残差散点图（scatter plot）
   2. 正式检验方法
      * Park检验法
      * Glejser检验法
      * BPG检验法
      * White检验法
3. 在发现存在异方差的基础上，对原模型进行异方差问题的处理：
   1. 使用加权最小二乘法校正异方差
   2. 使用White校正法解决异方差

### 实验背景——汽车油耗

**汽车油耗数据**：表1给出给出了81辆汽车在Y单位油耗的行驶里程数（英里/加仑），X2最高时速（英里/小时），X3发动机马力，X4汽车空间（立方英尺），X5车身重量（百磅）等方面的数据。

表1 汽车单位油耗里程数据(n=81)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| obs | Y | X2 | X3 | X4 | X5 |
| 1 | 65.4 | 96 | 49 | 89 | 17.5 |
| 2 | 56.0 | 97 | 55 | 92 | 20.0 |
| 3 | 55.9 | 97 | 55 | 92 | 20.0 |
| 4 | 49.0 | 105 | 70 | 92 | 20.0 |
| 5 | 46.5 | 96 | 53 | 92 | 20.0 |
| 77 | 18.1 | 165 | 322 | 50 | 45.0 |
| 78 | 17.2 | 140 | 238 | 115 | 45.0 |
| 79 | 17.0 | 147 | 263 | 50 | 45.0 |
| 80 | 16.7 | 157 | 295 | 119 | 45.0 |
| 81 | 13.2 | 130 | 236 | 107 | 55.0 |

变量说明见表2：

表2 变量定义及说明

|  |  |
| --- | --- |
| variable | label |
| obs | 汽车品牌序号 |
| Y | 单位油耗的行驶里程数（英里/加仑） |
| X2 | 最高时速（英里/小时） |
| X3 | 发动机马力 |
| X4 | 汽车空间（立方英尺） |
| X5 | 车身重量（百磅） |

请考虑如下样本回归模型：

## 主要实验步骤

### 导入数据并进行预处理

* 目标：
* 思路：
* 新建Eviews工作文件（见图1）
  + 提示：Excel数据，每个同学的Y数据都不同，找到自己学号对应下的Y
  + Eviews菜单操作：
    1. 依次操作：FileNewWorkfile
    2. 进行workfile create引导设置：
       - workfile structure type: unstructured/undatede
       - data range：81
       - workfile names(optional):
         * WF: car（**建议命名**）
         * Page: mileage（**建议命名**）
* Eviews导入数据
  + 提示：Excel数据，每个同学的Y数据都不同，找到自己学号对应下的Y数据（X数据所有同学都一样）
  + 菜单操作（Excel和Eviews）：
    1. Excel找到数据。Excel表格中仅保留自己需要的数据（obs, Y, X2, X3, X4, X5）
    2. Eviews导入数据。FileImportImport From File：d:/econometrics/data/Lab6-car.xlsx

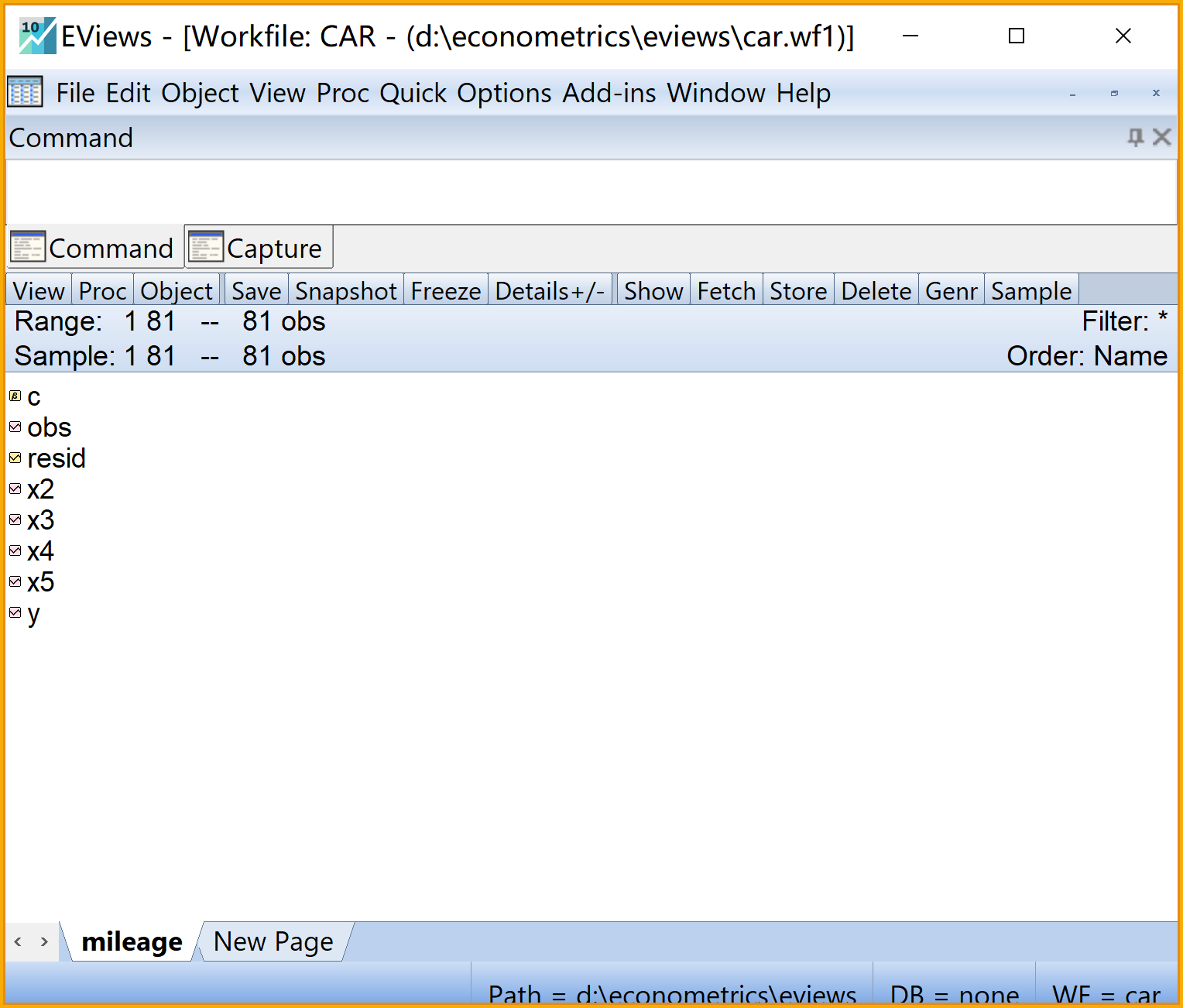


图1 导入数据的Eviews视窗

### 采用最小二乘法建立主回归模型

* 目标：
* 思路：
* 提示：主回归模型为
* Eviews菜单操作（见图2）：
  1. 依次选择QuickEstimation Equation
  2. 引导设置Equation Estimationspecification
     1. Equation specification：输入命令 Y c X2 X3 X4 X5
     2. Estimation settings：
        + Method: 下拉选择LS - Least Squares (NLS and ARMA)
        + Sample: **默认设置**
     3. 点击OK
  3. 模型命名：建议为eq\_m0
  + 主回归分析结果见图3：



图2 主回归模型Eviews操作

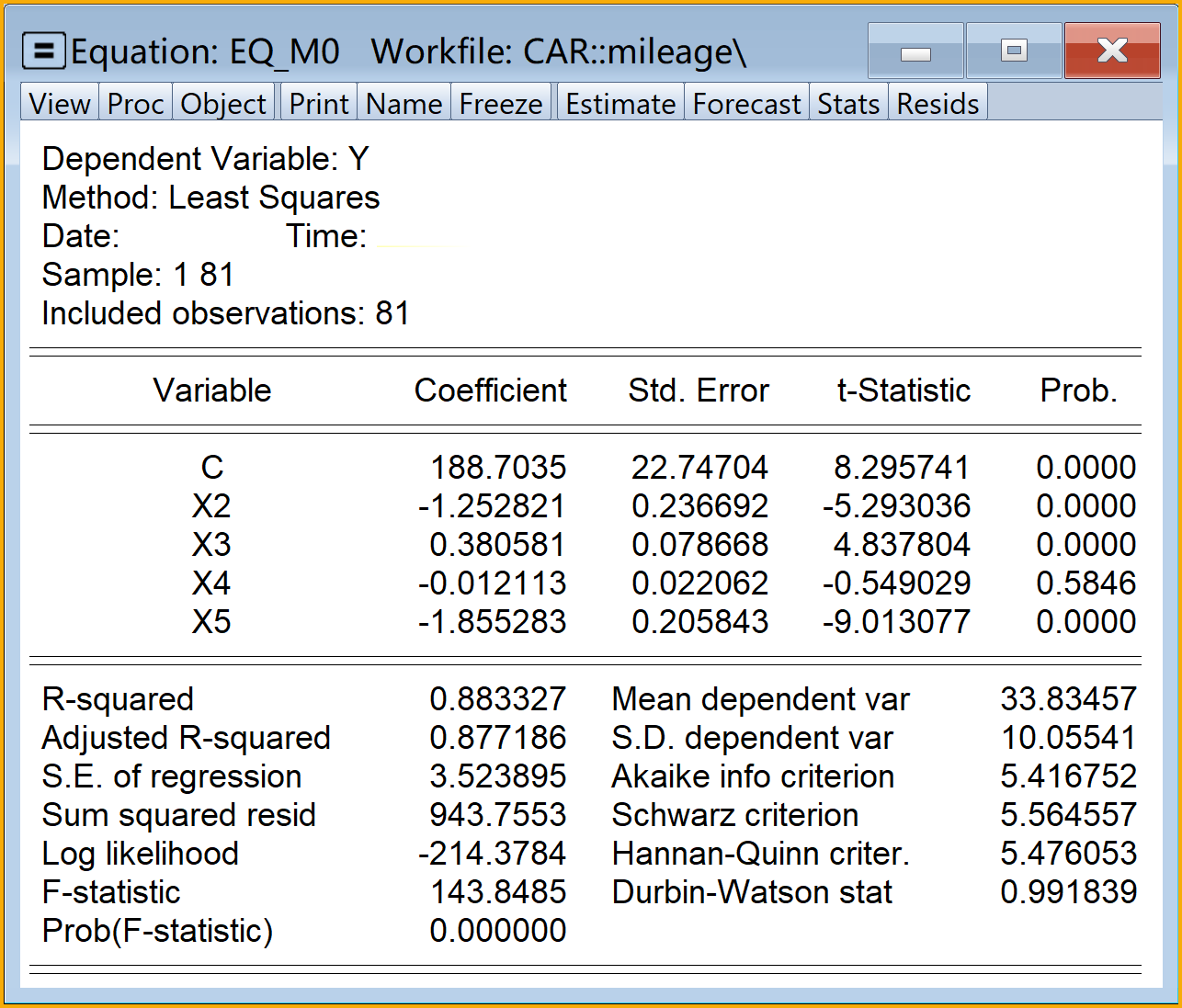


图3 主回归模型Eviews报告

### 侦查模型是否存在异方差

#### 初步观察法（观察主回归方程）

* 目标：观察主回归方程分析报告，分析回归报告结果，得出初步结论
* 思路：观察检验，判定系数，检验的关系
* 提示：
  + 模型使用的数据是否是截面数据
  + 主回归分析报告的值
  + 模型整体检验结果
  + 斜率系数的检验结果
* 分析结论： 根据主回归报告（见图3），表明模型可能存在**严重**的异方差问题。

#### 非正式检验法（图示法）

* 目标：观察或的图形模式
* 思路：判定或与、、、、等的图形关系
* 提示：
  + 描点图（dot plot）是分析一个变量的图形模式。例如或（做纵轴）相对于（做横轴）的图形关系
  + 散点图（scatter plot）是分析两个变量之间的图形模式。例如或（做纵轴），相对于或或或的图形关系。
* Eviews菜单操作：
  1. 分别生成新序列、、和（见图4）
     1. 生成残差和序列（建议分别命名为ei和ei\_sqr）
        + 命令视窗（Command）输入命令 ：series ei=resid
        + 命令视窗（Command）输入命令 ：series ei\_sqr=resid^2
        + 运行命令：命令行中按Enter键
     2. 生成残差序列（建议命名为Y\_sqr）
        + 命令视窗（Command）输入命令 ：series Y\_sqr=Y^2
        + 运行命令：命令行中按Enter键
     3. 生成残差序列（建议分别命名为X2\_sqr,X3\_sqr,X4\_sqr,X5\_sqr）
        + 命令视窗（Command）输入命令 ：series X2\_sqr=X2^2
        + 命令视窗（Command）输入命令 ：series X3\_sqr=X3^2
        + 命令视窗（Command）输入命令 ：series X4\_sqr=X4^2
        + 命令视窗（Command）输入命令 ：series X5\_sqr=X5^2
        + 运行命令：上述命令行中依次按Enter键
     4. 查看结果：
        + 双击ei
        + 双击ei\_sqr
        + 双击Y\_sqr
        + 双击X2\_sqr
        + 双击X3\_sqr
        + 双击X4\_sqr
        + 双击X5\_sqr
  2. 绘制和序列的描点图（dot plot）（见图5）
     1. 选择序列对象：键盘Ctrl键+依次单击选择序列ei和ei\_sqr
     2. 进入引导菜单： Quick Graph
        + 选择绘图类型(Graph type)：Dot plot
        + 选择绘图细节(Detail)： Multiple series 下拉选择 Multiple graphs
     3. 点击完成：OK
     4. 命名并保存绘图（graph）对象：建议命名为dot\_resid
     5. 查看结果：双击dot\_resid（见图6）
  3. 绘制序列对; 的散点图图（scatter plot）（见图7）
     1. 选择序列对象：键盘Ctrl键+依次单击选择序列ei；Y；X2；X3；X4；X5
     2. 进入引导菜单： Quick Graph
        + 选择绘图类型(Graph type)：Scatter
        + 选择绘图细节(Detail)： Multiple series 下拉选择 Multiple graphs - First vs. all
     3. 点击完成：OK
     4. 命名并保存绘图（graph）对象：建议命名为scatter\_ei
     5. 查看结果：双击scatter\_ei（见图7）
  4. 绘制序列对; 的散点图图（scatter plot）（见图8）
     1. 选择序列对象：键盘Ctrl键+依次单击选择序列ei\_sqr；Y；X2；X3；X4；X5
     2. 进入引导菜单： Quick Graph
        + 选择绘图类型(Graph type)：Scatter
        + 选择绘图细节(Detail)： Multiple series 下拉选择 Multiple graphs - First vs. all
     3. 点击完成：OK
     4. 命名并保存绘图（graph）对象：建议命名为scatter\_ei\_sqr
     5. 查看结果：双击scatter\_ei\_sqr（见图8）

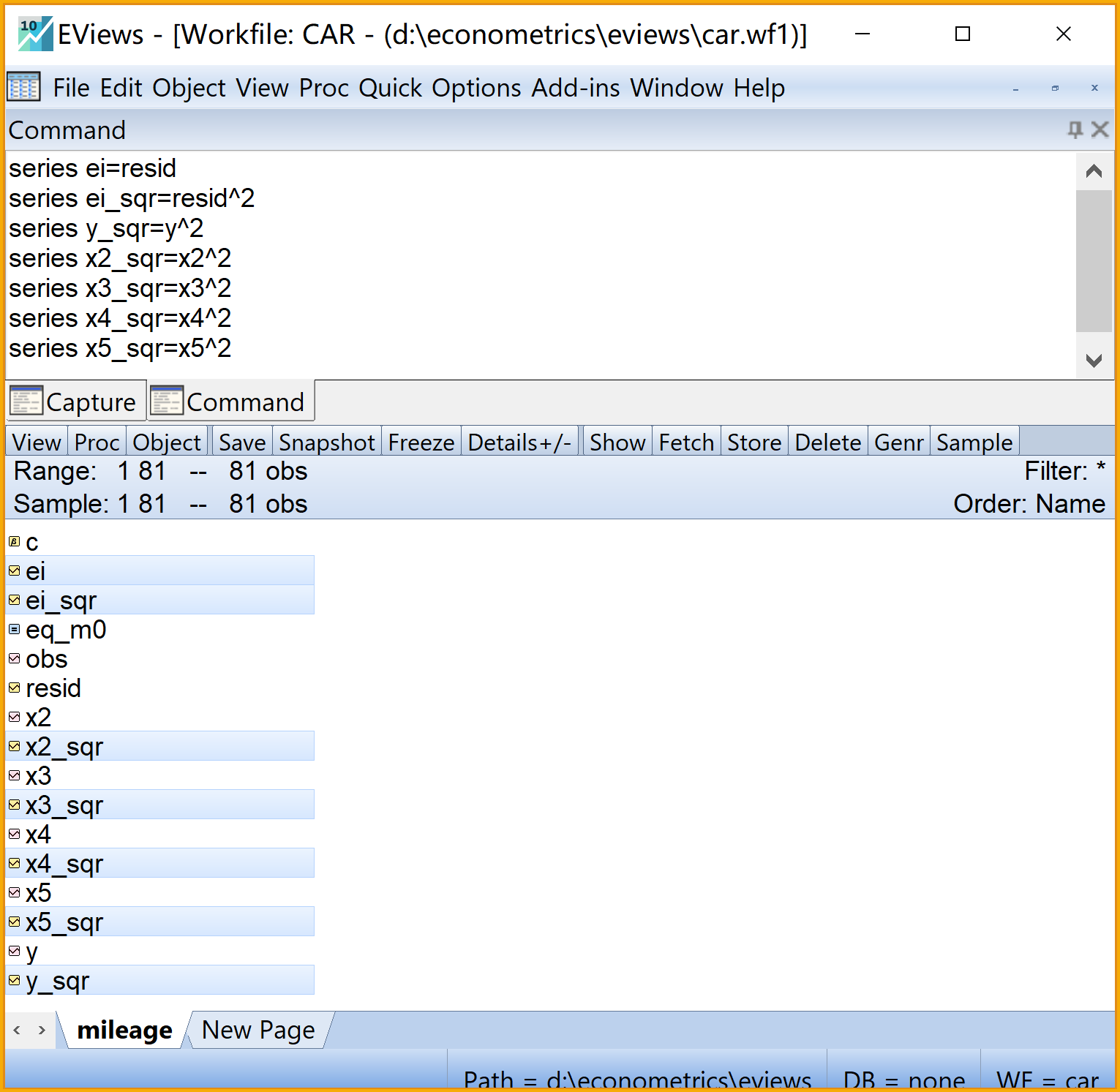


图4 生成相关变量

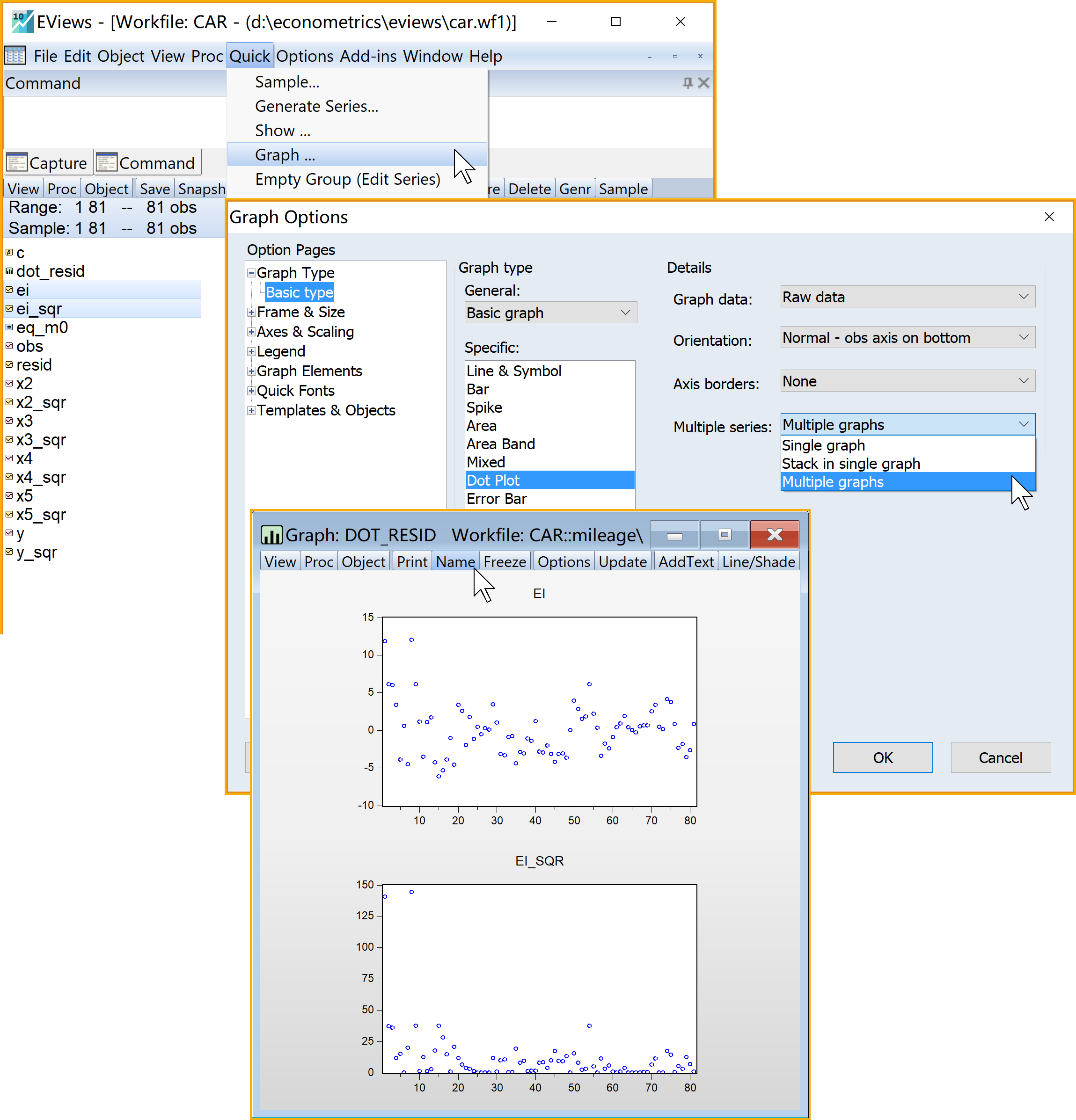


图5 残差及残差平方的描点图

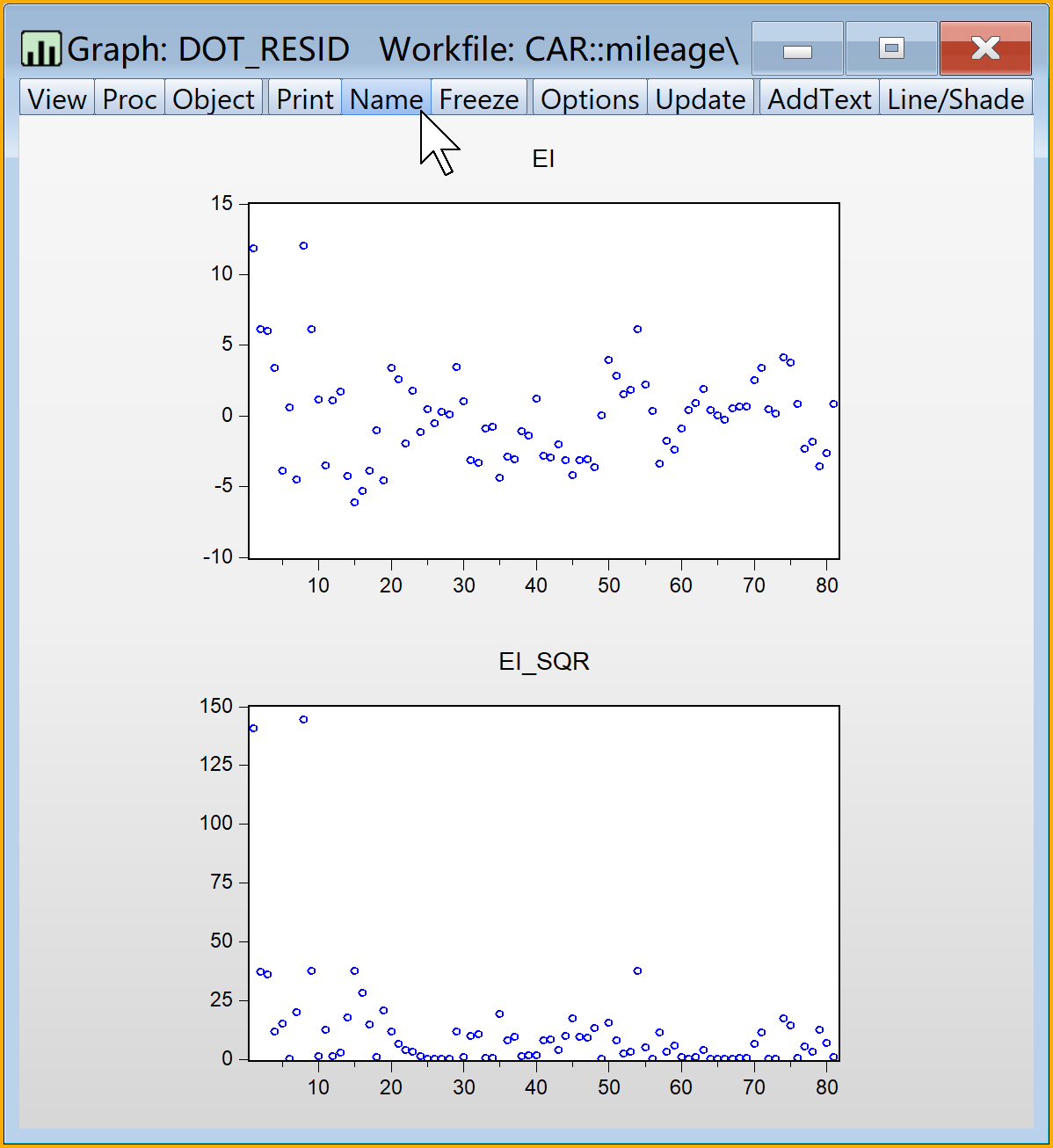


图6 残差及残差平方的描点图报告

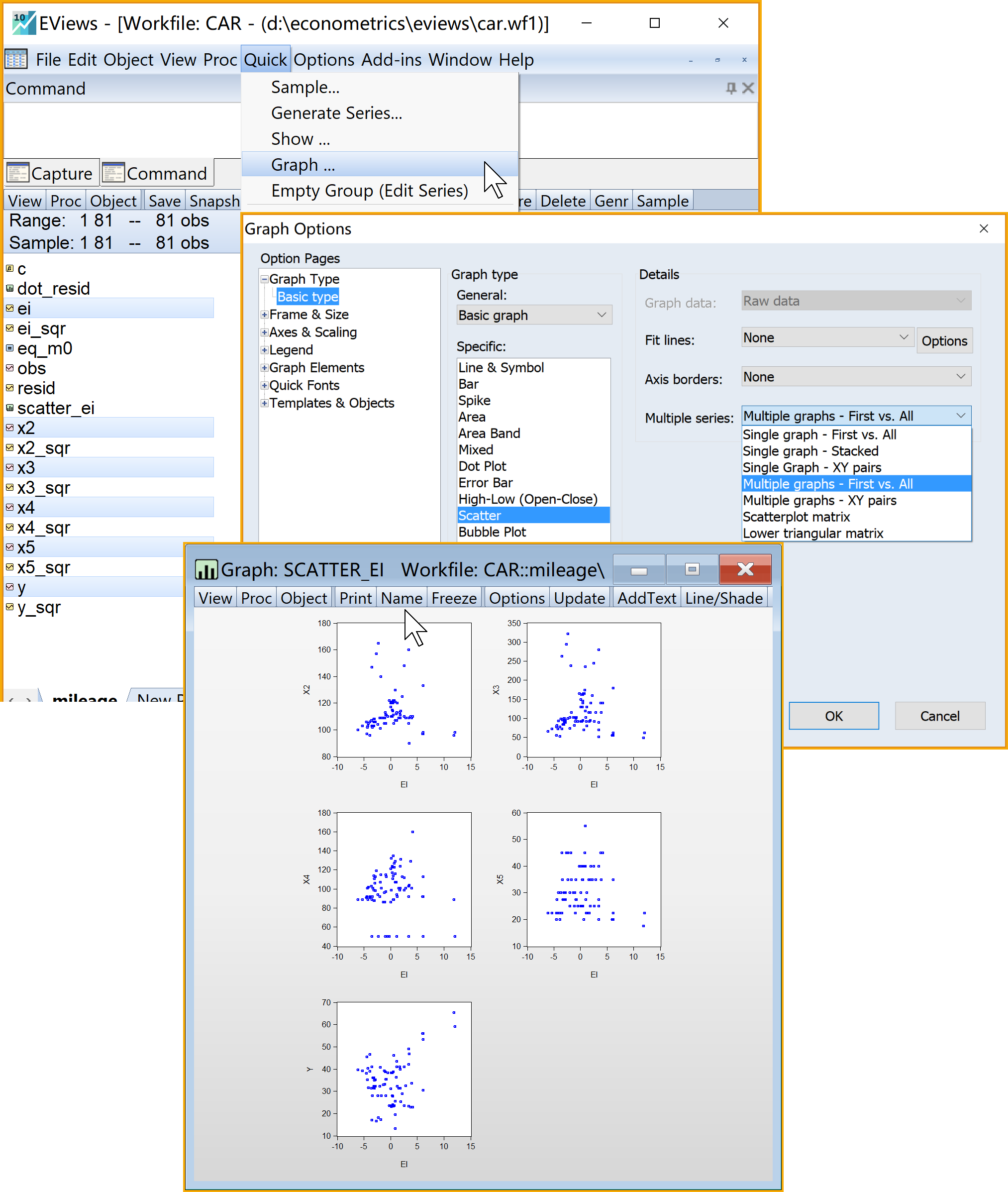


图7 残差与模型变量的散点图

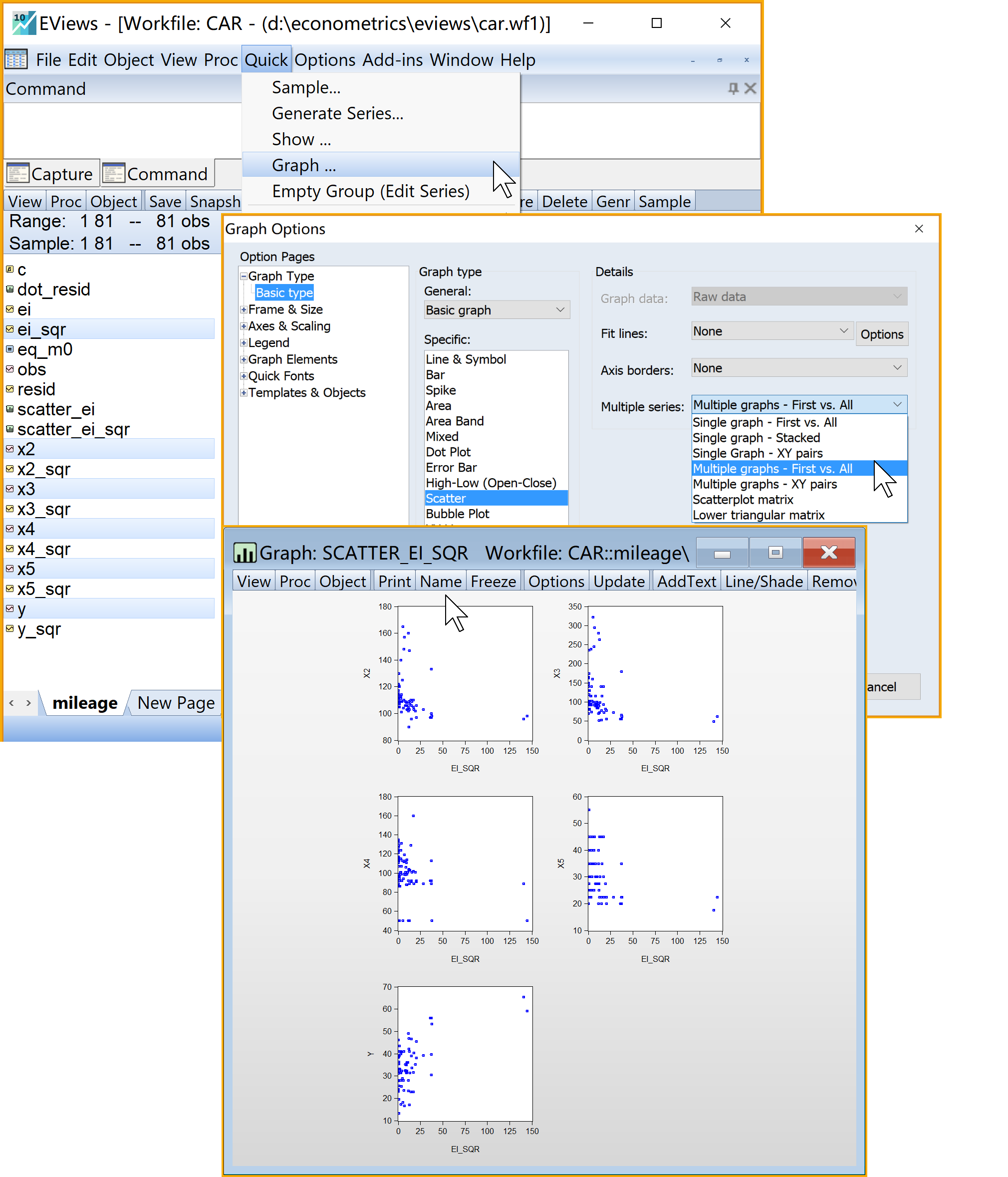


图8 残差平方与模型变量的散点图

#### 正式检验法

* 目标：利用Eviews的异方差诊断菜单，分别对主回归模型(1)进行 进行异方差诊断。
* 思路：诊断方法包括Park检验、Glejser检验、BPG检验和White检验等。根据辅助诊断方程的理论假设，分析Eviews诊断报告，与相关参考标准进行比较，得到相关结论
* 定义：
  + 主回归模型(Main Model)是指Y变量对全部X变量的线性回归（如主模型(1)）
  + 辅助诊断模型(Auxiliary Model)是指利用主回归模型的残差序列对X变量进行特定的线性回归（具体的辅助诊断模型有多个）
* 提示：
  + 操作提示：用Eviews的异方差诊断菜单 View Residual Diagnostics Heteroskedasticity Test
  + 诊断提示：若发现假设检验结果不显著（视具体检验方法而不同）则表明主模型(1)为同方差；否则就表明主模型(1)为异方差
  + 理论提示：你也可以自己根据实际情况设定个性化的辅助诊断方程

##### Park检验法

* 诊断辅助方程：
* 诊断标准：
  + 如果诊断辅助方程(2)的F检验**不显著**（对应的概率值P>0.1），则表明主模型(1)是同方差
  + 如果诊断辅助方程(2)的F检验**显著**（对应的概率值P<0.1），则表明主模型(1)是异方差。
* Eviews说明：Eviews没有Park异方差检验的诊断菜单，但可以通过选择Harvey异方差菜单，并修改相关变量设置来得到Park检验报告。
* Eviews操作（菜单操作实现，具体见图9）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择Harvey
     + 设置诊断方程（Regressors）：输入c log(X2) log(X3) log(X4) log(X5)
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（talbe）对象：点击name(建议为tab\_park)
     + 查看结果：双击tab\_park

具体Eviews报告见10：

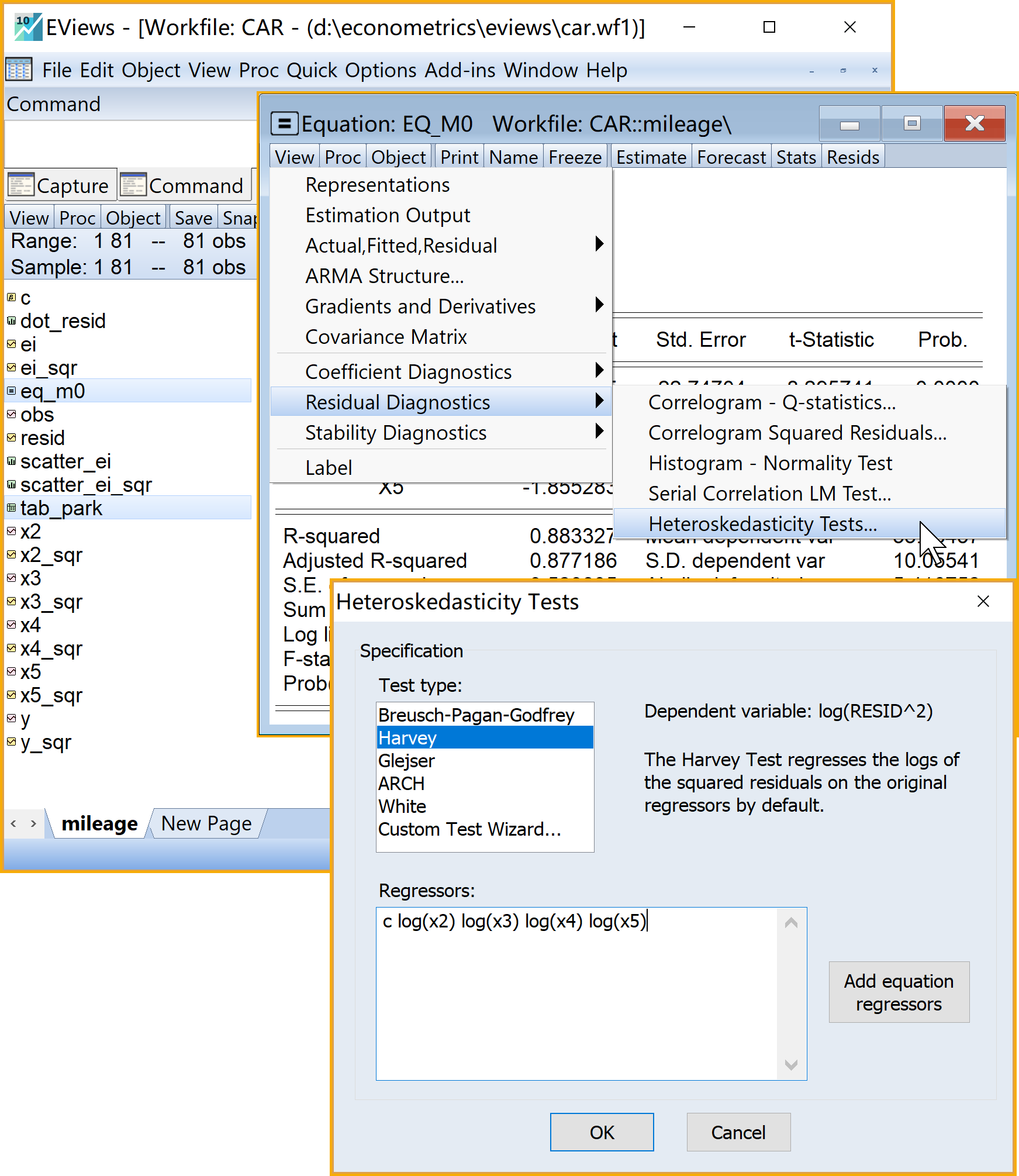


图9 Park异方差检验操作

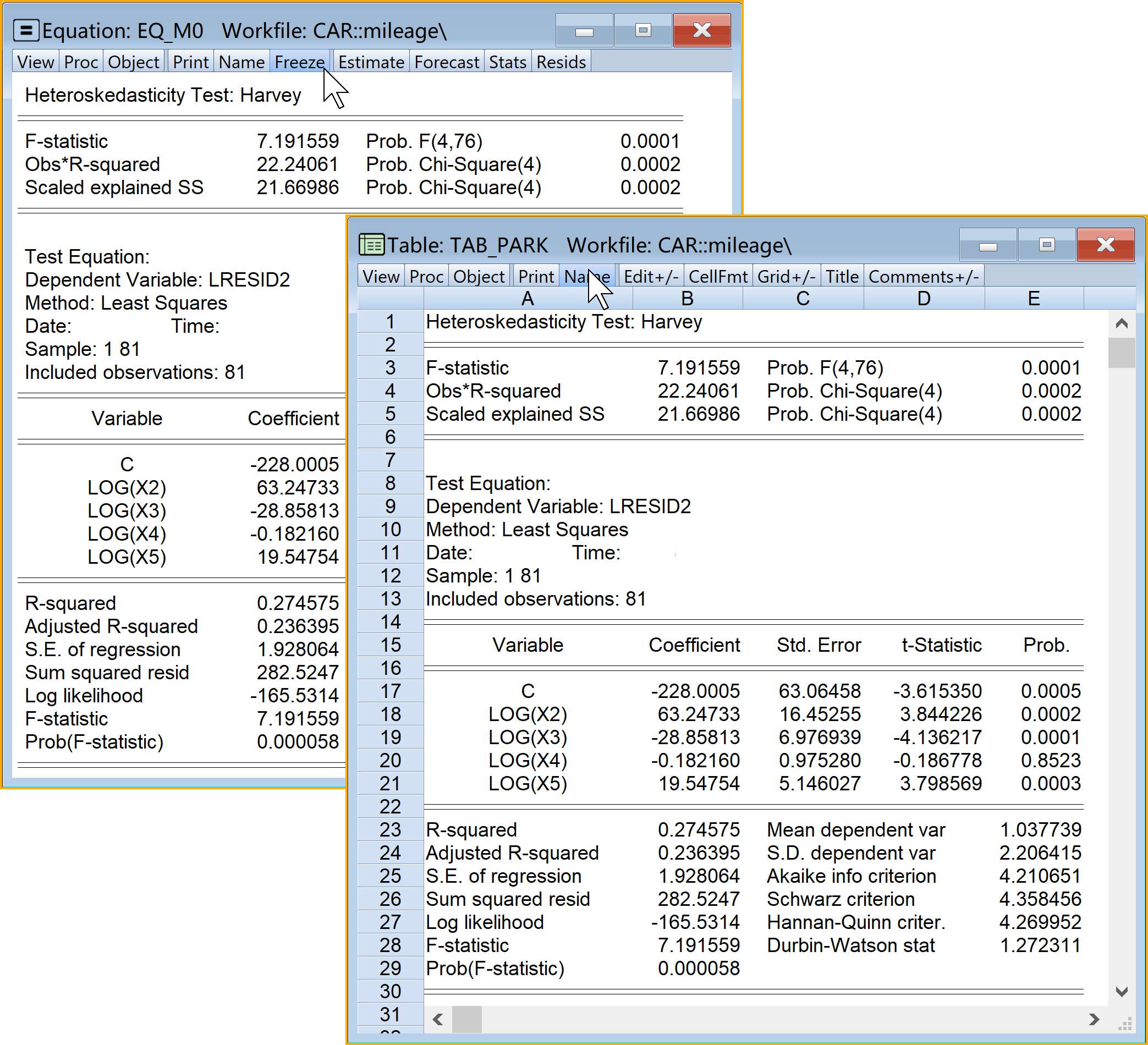


图10 Park异方差检验报告

##### Glejser检验法

* 诊断辅助方程：
* 诊断标准：
  + 如果诊断辅助方程(2)的F检验**不显著**（对应的概率值P>0.1），则表明主模型(1)是同方差
  + 如果诊断辅助方程(2)的F检验**显著**（对应的概率值P<0.1），则表明主模型(1)是异方差。
* Eviews操作1（Glejser辅助方程(3)，具体见图11）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择Glejser
     + 设置诊断方程（Regressors）：输入c X2 c X3 c X4 c X5
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（table）对象：点击name(建议为tab\_G1)
     + 查看结果：双击tab\_G1

具体Eviews报告见12：

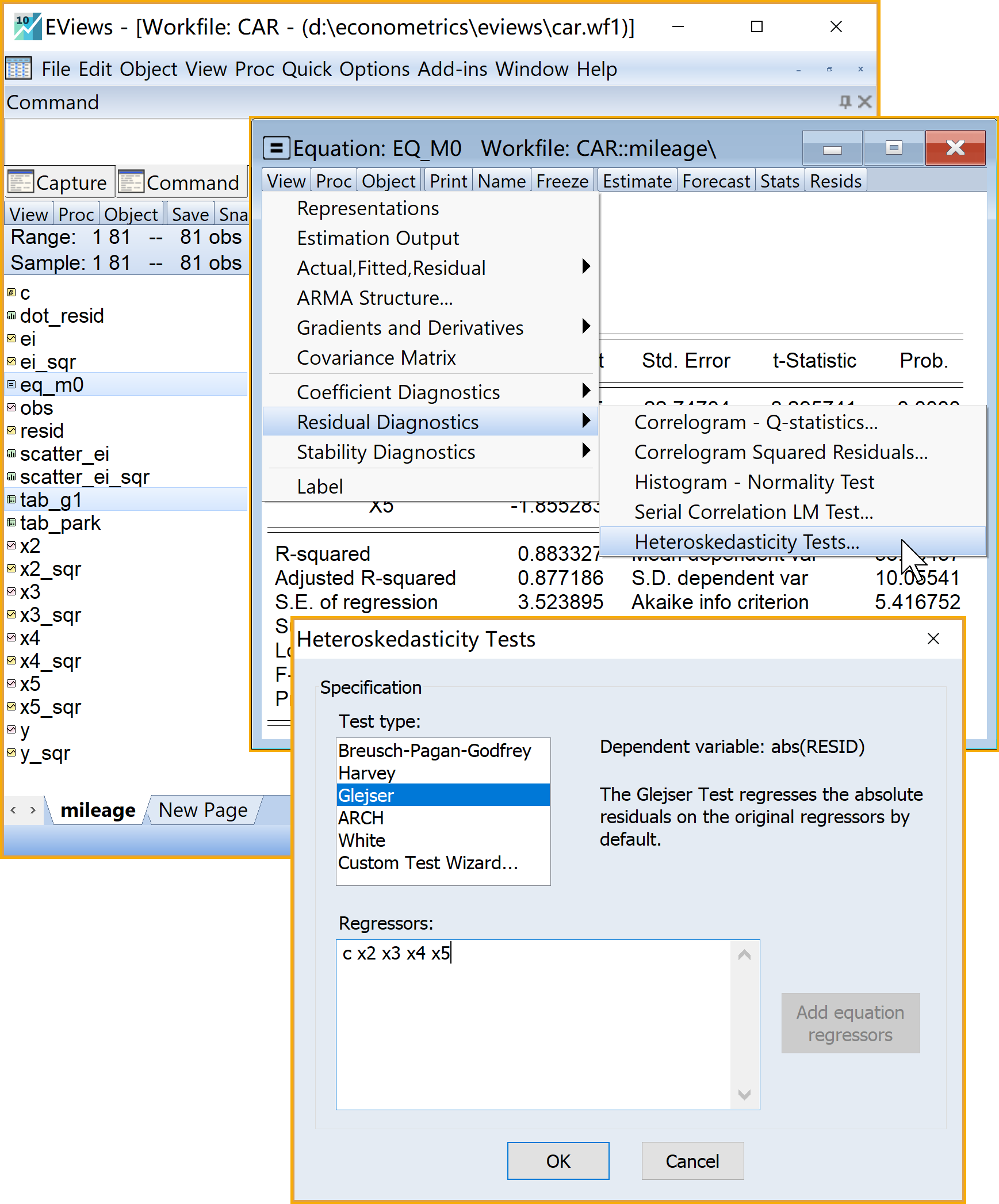


图11 Glejser异方差检验操作1

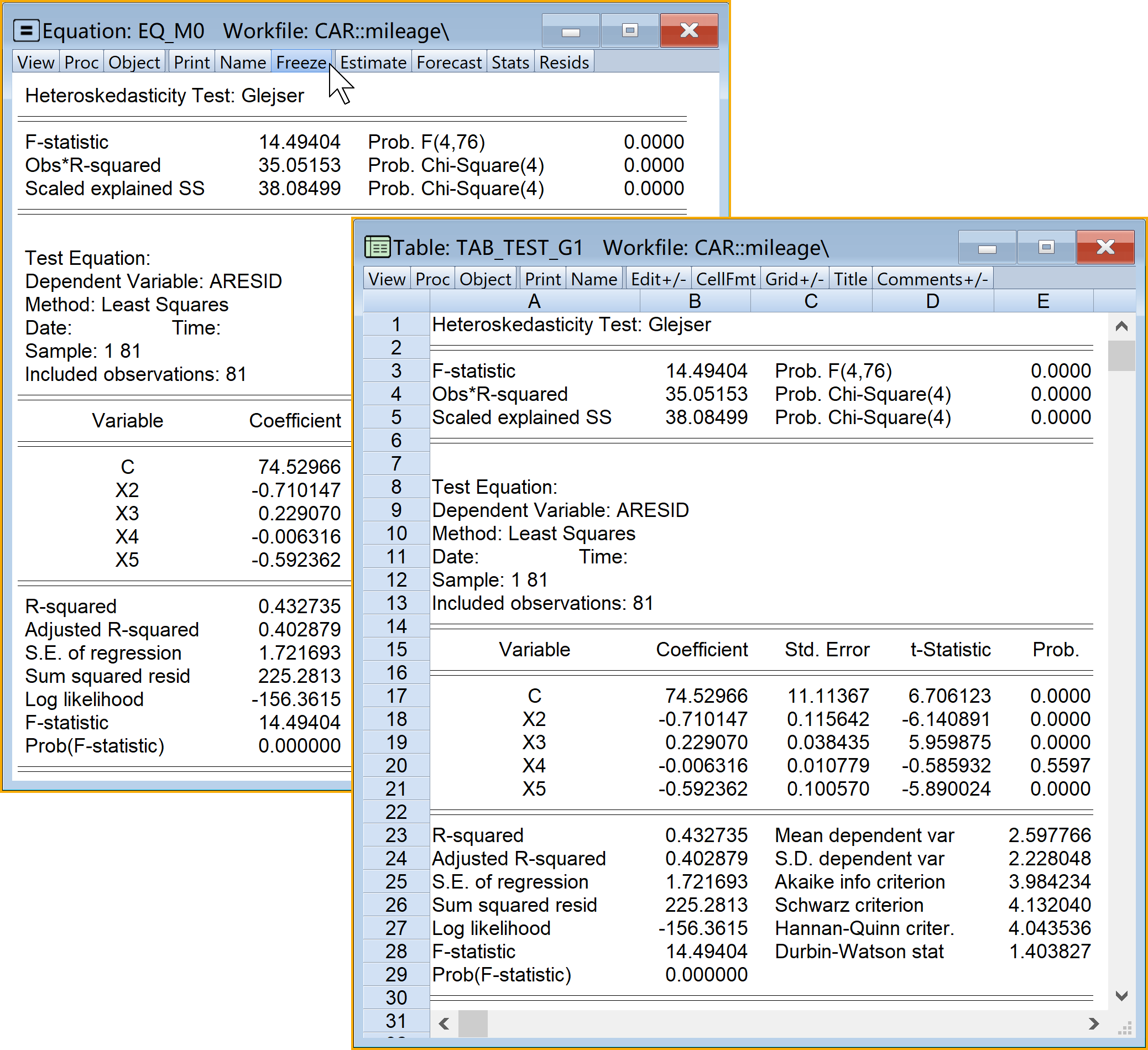


图12 Glejser异方差检验报告1

* Eviews操作2（Glejser辅助方程(4)，具体见图13）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择Glejser
     + 设置诊断方程（Regressors）：输入c X2^0.5 X3^0.5 X4^0.5 X5^0.5
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（table）对象：点击name(建议为tab\_G2)
     + 查看结果：双击tab\_G2

具体Eviews报告见14：

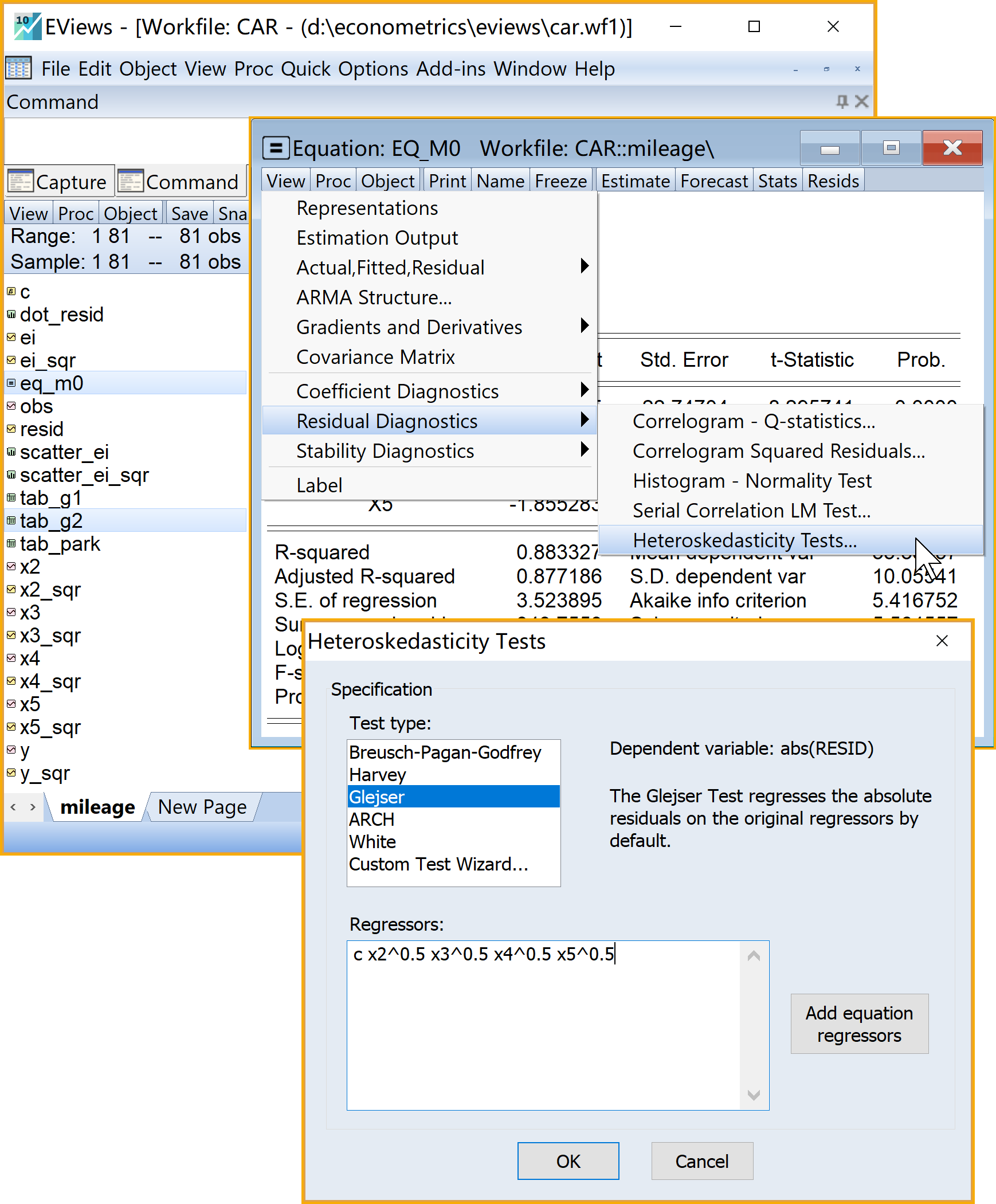


图13 Glejser异方差检验操作2

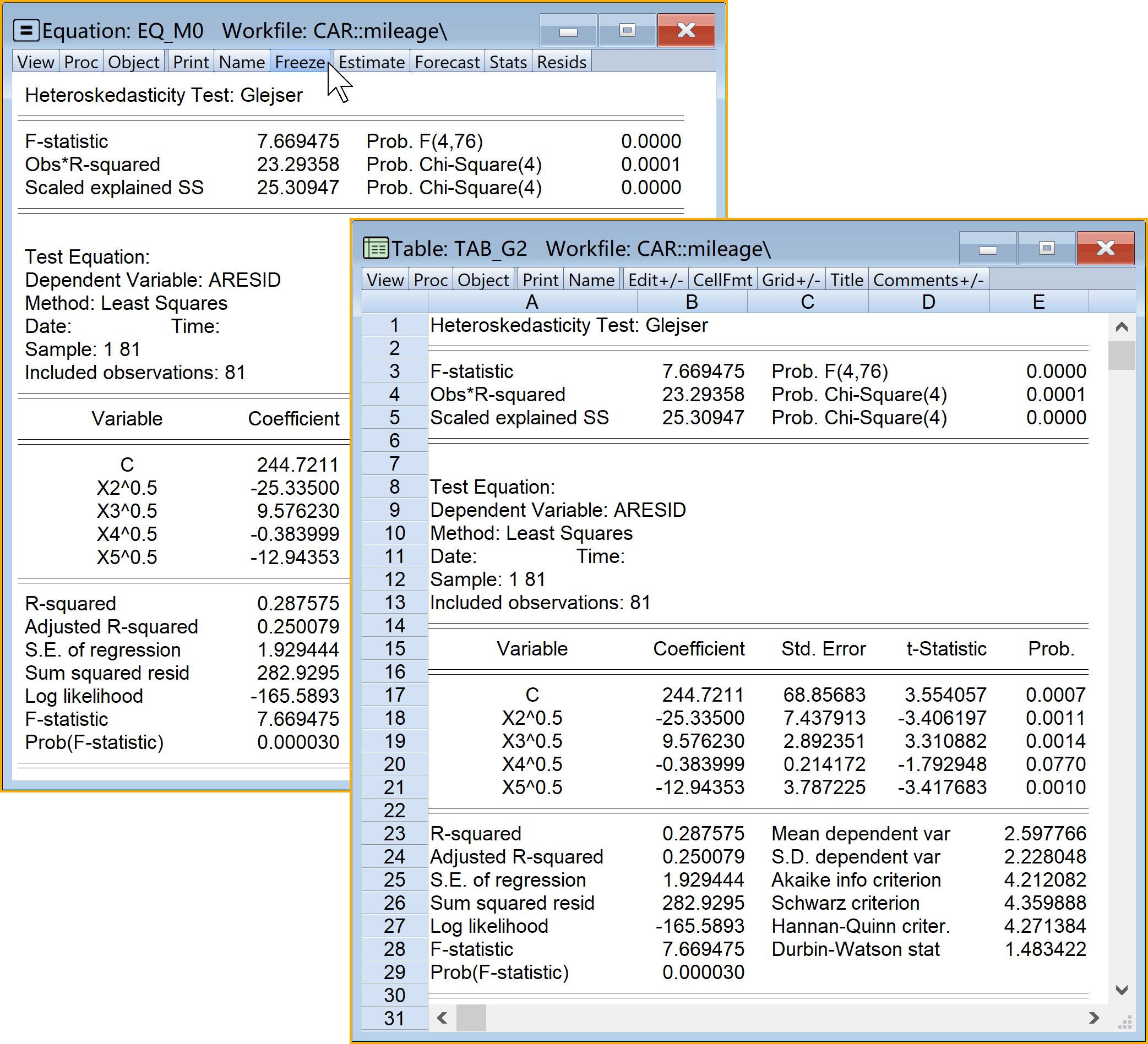


图14 Glejser异方差检验报告2

* Eviews操作3（Glejser辅助方程(5)，具体见图15）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择Glejser
     + 设置诊断方程（Regressors）：输入c 1/ X2 1/ X3 1/ X4 1/ X5
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（table）对象：点击name(建议为tab\_G3)
     + 查看结果：双击tab\_G3

具体Eviews报告见16：

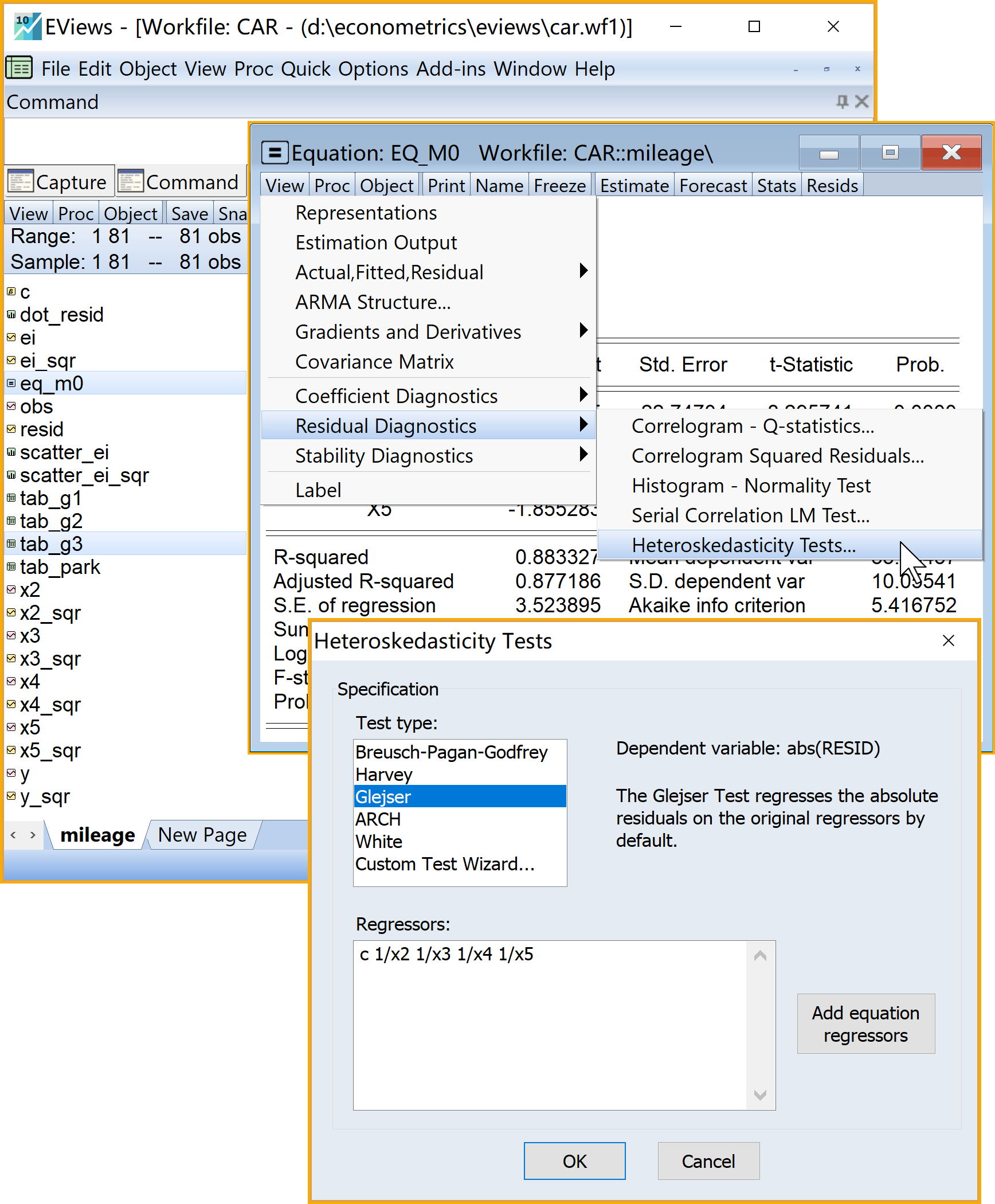


图15 Glejser异方差检验操作3

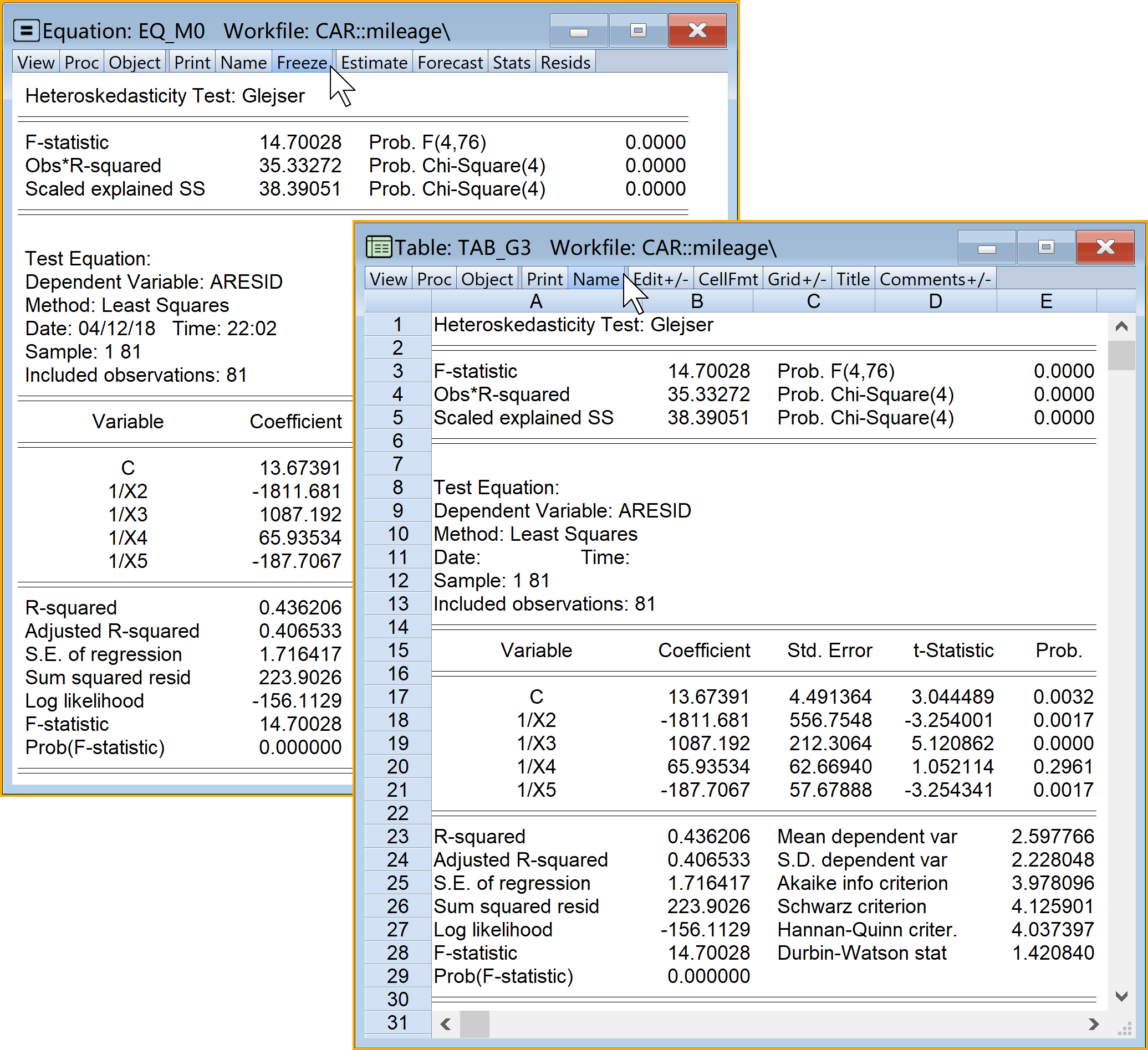


图16 Glejser异方差检验报告3

##### BPG检验法

* 诊断辅助方程：
* 诊断标准：
  + 如果诊断辅助方程(6)的检验（scaled explained SS）**不显著**（对应的概率值P>0.1），则表明主模型(1)是同方差
  + 如果诊断辅助方程(6)的检验（scaled explained SS）**显著**（对应的概率值P<0.1），则表明主模型(1)是异方差。
* Eviews操作（菜单操作实现，具体见图17）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择Breusch-Pagan-Godfrey
     + 设置诊断方程（Regressors）：输入c X2 X3 X4 X5
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（table）对象：点击name(建议为tab\_bpg)
     + 查看结果：双击tab\_bpg

具体Eviews报告见18：

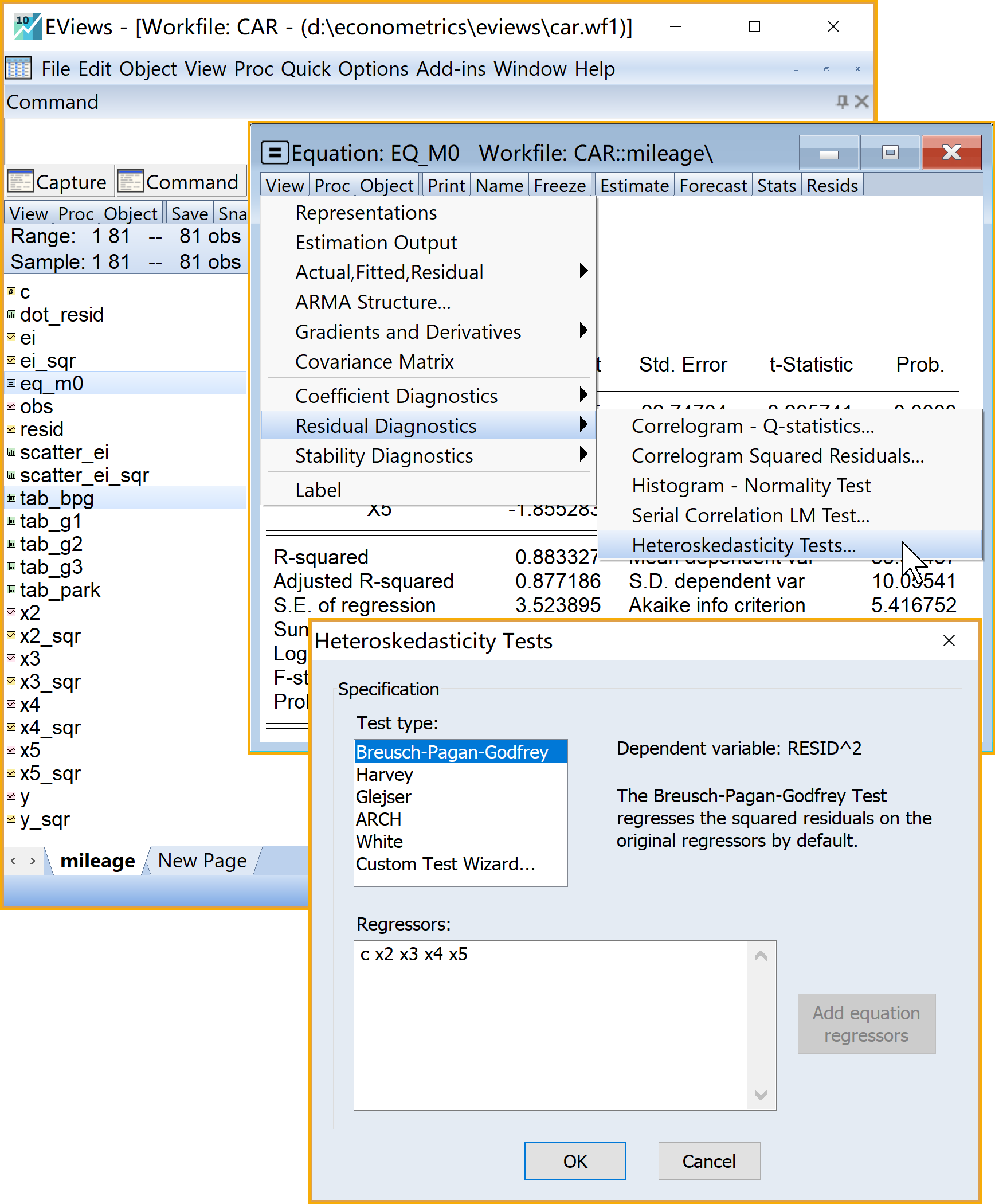


图17 BPG异方差检验操作

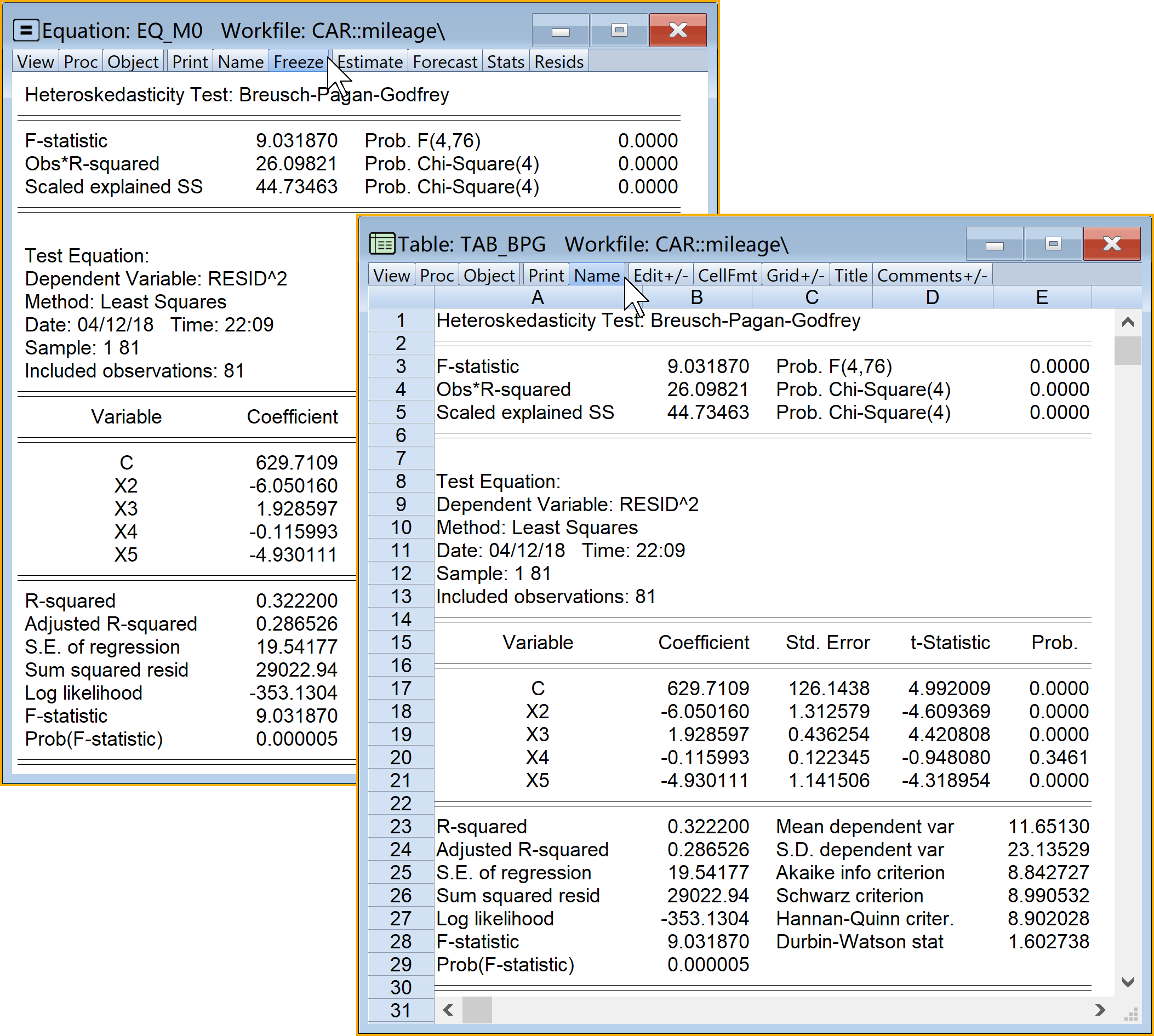


图18 BPG异方差检验报告

##### White检验法

* 诊断辅助方程：
* 诊断标准：
  + 如果诊断辅助方程(8)的检验（scaled explained SS）**不显著**（对应的概率值P>0.1），则表明主模型(7)是同方差
  + 如果诊断辅助方程(8)的检验（scaled explained SS）**显著**（对应的概率值P<0.1），则表明主模型(7)是异方差。
* Eviews操作（菜单操作实现，具体见图19）：
  1. 打开主方程：双击方程（equation）对象eq\_m0
  2. 进入引导菜单： View Residual Diagnostics Heteroskedasticity Test Specification
     + 设置诊断方法（Test type）: 点击选择White
     + 设置诊断方程（下面两类方程自行选择其一）：
       - 交叉项方程：**勾选** Include White cross term
       - 非交叉项方程：**不勾选** Include White cross term
  3. 完成设置：点击Ok
  4. 命名并保存表格（table）对象
     + 另存为表格（table）对象：点击Freeze
     + 命名并保存表格（table）对象：点击name(建议为tab\_white)
     + 查看结果：双击tab\_white

具体Eviews报告见20：

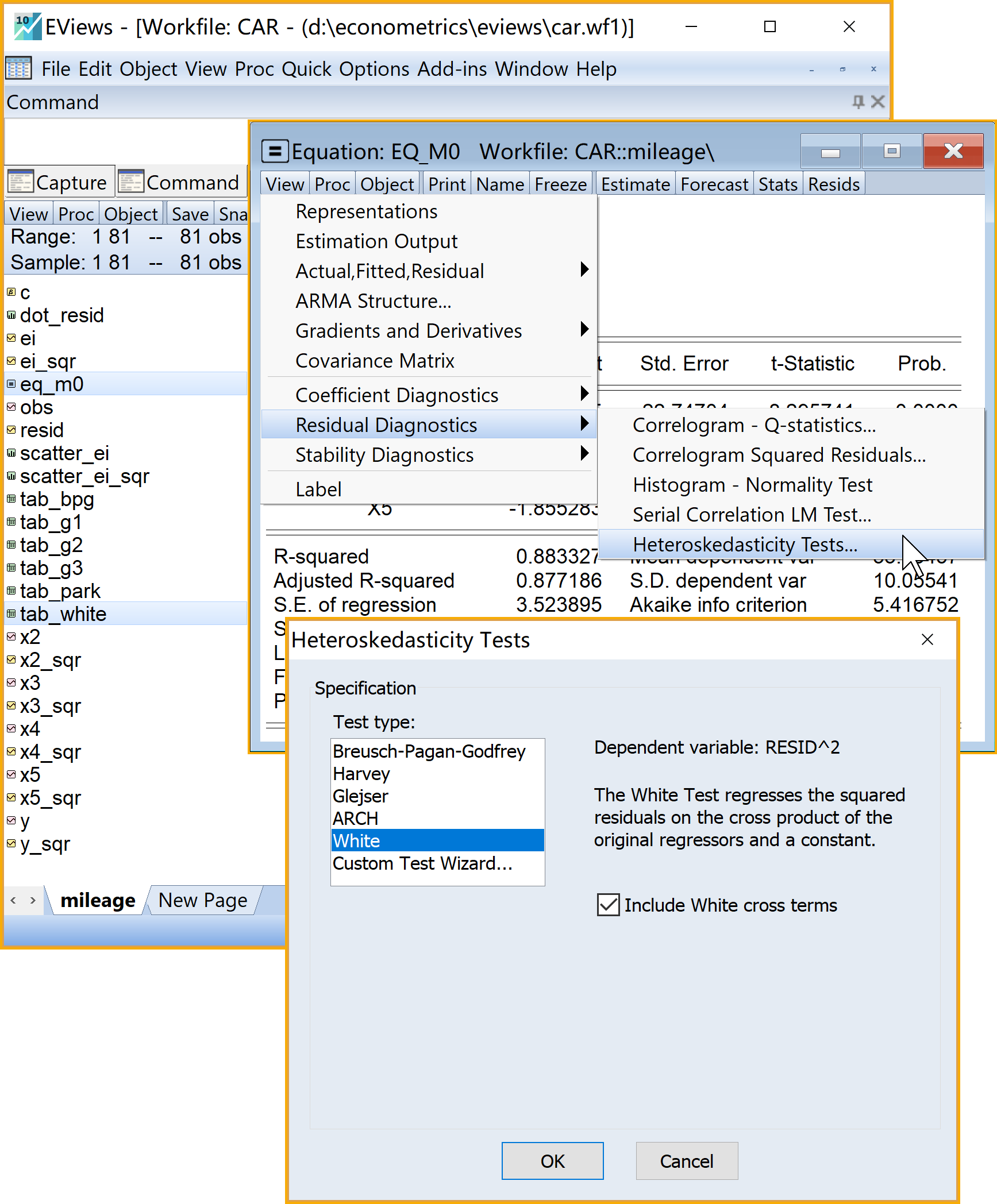


图19 White异方差检验报告操作

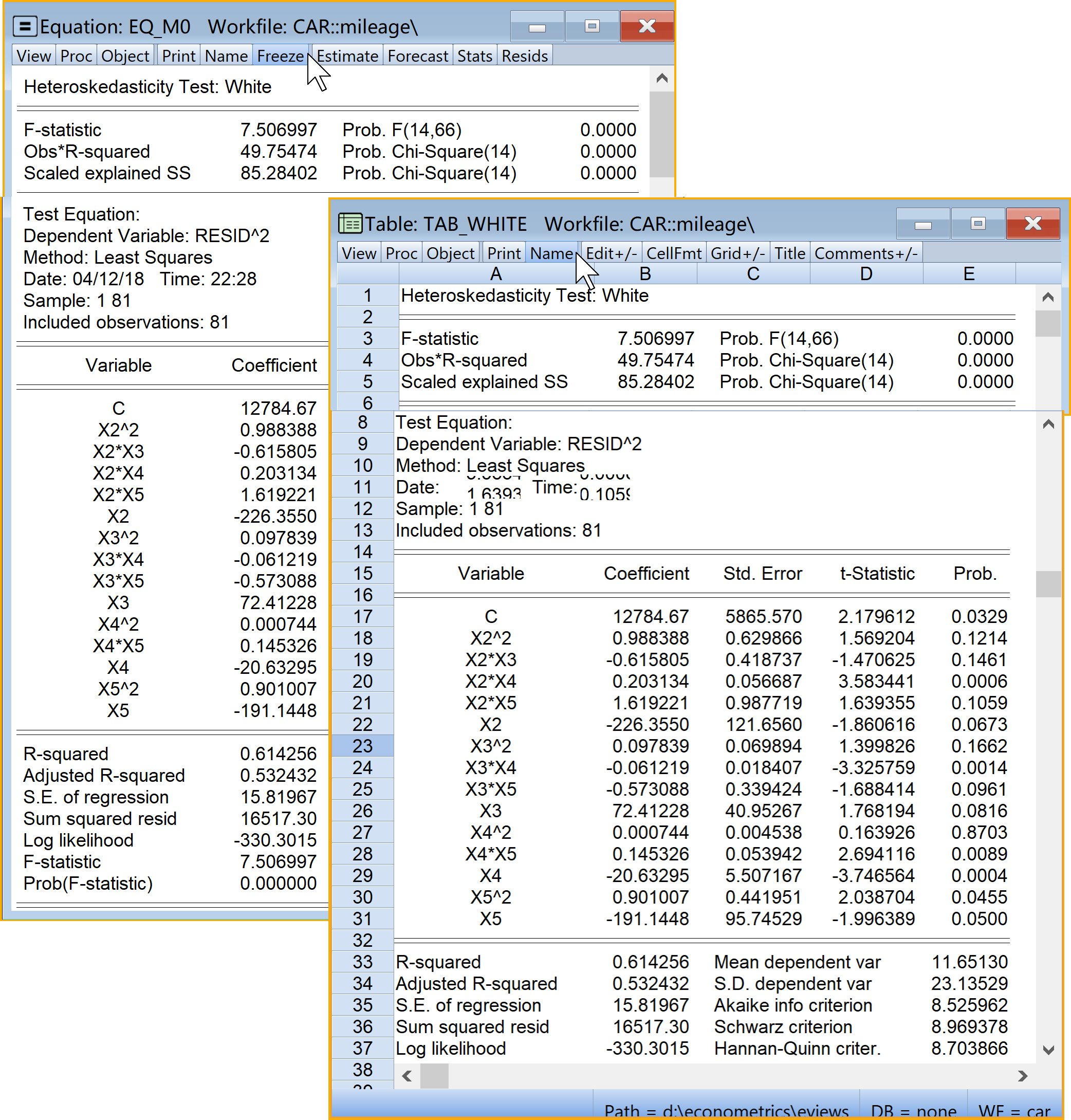


图20 White异方差检验报告

### 存在异方差问题的模型矫正

#### 使用加权最小二乘法（WLS）矫正异方差问题

##### WLS矫正情形1： 方差已知且等于样本方差

* 理论提示：

如果主模型(1)存在异方差问题，且假设方差正比于样本方差，则有：

对主模型(1)两边同时除以，得到：

##### WLS矫正情形2： 方差正比于

* 理论提示：

如果主模型(1)存在异方差问题，且假设方差正比于，则有：

对主模型(1)两边同时除以，得到：

* Eviews操作：

##### WLS矫正情形3： 方差正比于

* 理论提示：

如果主模型(1)存在异方差问题，且假设方差正比于，则有：

对主模型(1)两边同时除以，得到：

##### WLS矫正情形4： 方差正比于

* 理论提示： 如果主模型(1)存在异方差问题，且假设方差正比于，则有：

对主模型(1)两边同时除以，得到：

* Eviews操作：

##### WLS矫正情形5： 方差未知

* 理论提示： 如果主模型(1)存在异方差问题，且假设方差正比于，则有：

对主模型(1)两边同时取对数，得到：

#### 使用White校正法矫正异方差问题

* 理论提示
* Eviews操作（见图21）
  1. 依次选择 Quick Estimation Equation
  2. 引导设置Equation Estimation Specification
     1. 方程设置（Equation Specification）： Y （此处如果仅填Y变量，则任何X变量都没有强制一定要留在模型中）
     2. 输入自变量（List of search regressors）：Y c X2 X3 X4 X5
     3. 估计方法（Estimation settings）：
        + Method：选择LS - Least Squares(NLS and ARMA)
        + Sample： **默认设置**
  3. 引导设置Equation Estimation Options
     1. 系数协方差设置（Coefficient covariance）
        + 协方差方法（Coefficient method）：下拉选择Huber-White
     2. 权重设置（Weights）：默认设置
     3. 最优化设置（Optimization）：默认设置
     4. 完成设置：点击OK
  4. 模型命名：建议为eq\_adj\_white
  5. 查看分析报告（见图22）

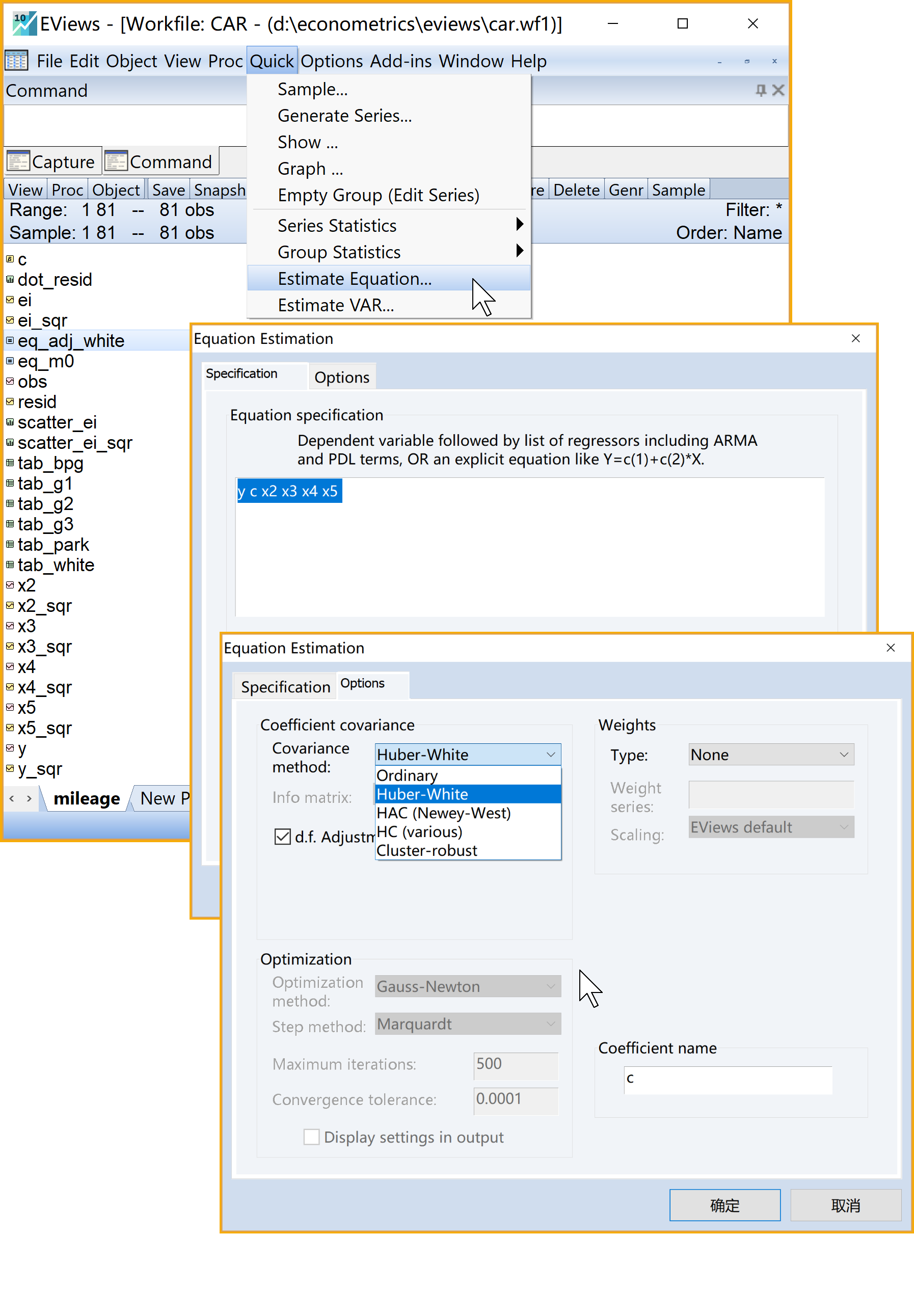


图21 White异方差矫正法的操作

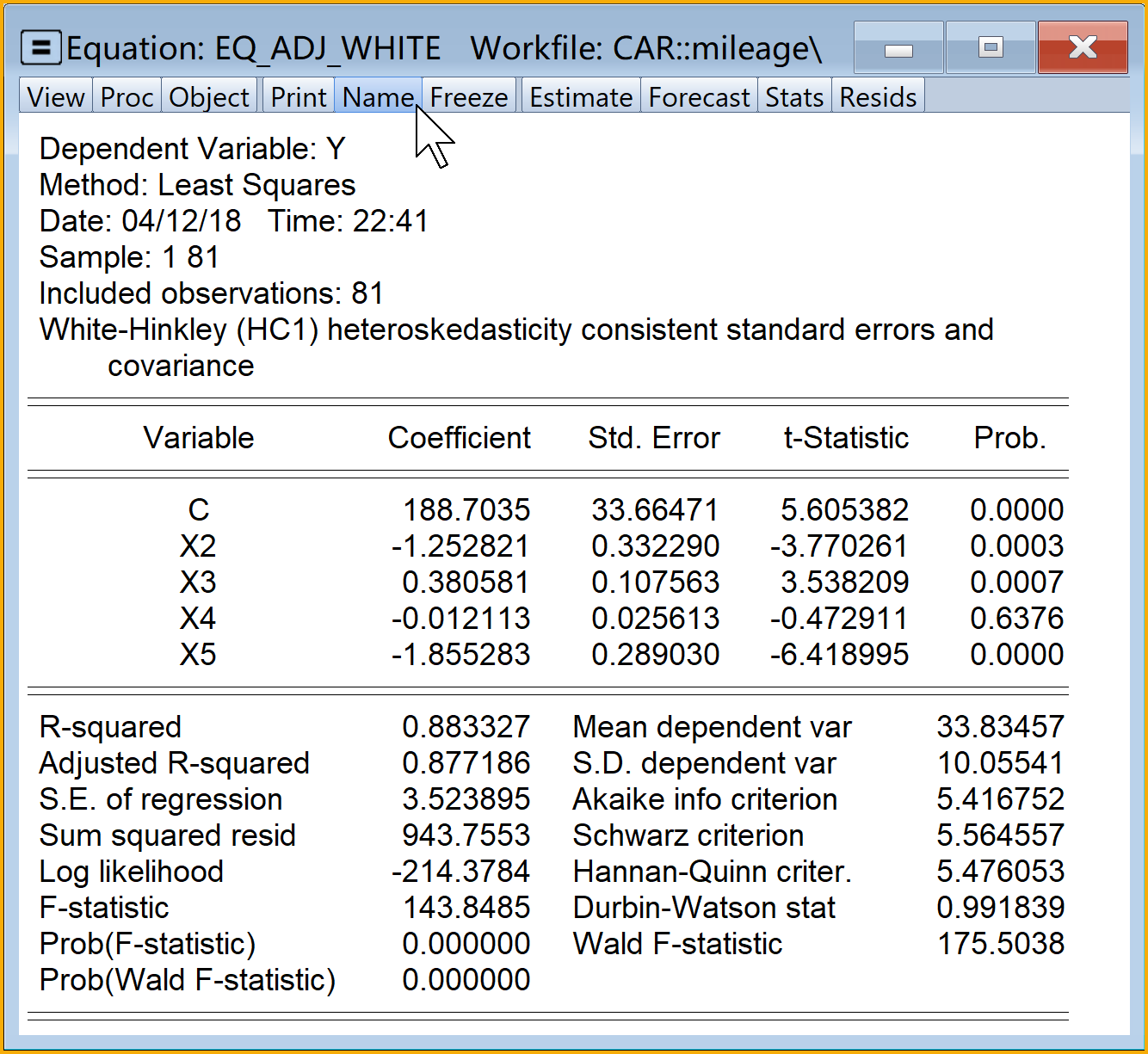


图22 White异方差矫正法的Eviews报告

## 作业题

**财富500强企业高管薪水数据**：表3给出给出了447辆汽车在Y高管薪水及分红，X2担任 CEO年数（不足6个月视为的0），X3总裁CEO的年龄，X4企业的总销售收入，X5企业的利润，X6企业的总资产等方面的数据。

表3 财富500强企业高管薪水数据(n=447)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| obs | Y | X2 | X3 | X4 | X5 | X6 |
| 1 | 3030 | 7 | 61 | 161315 | 2956.0 | 257389 |
| 2 | 6050 | 0 | 51 | 144416 | 22071.0 | 237545 |
| 3 | 3571 | 11 | 63 | 139208 | 4430.0 | 49271 |
| 4 | 3300 | 6 | 60 | 100697 | 6370.0 | 92630 |
| 5 | 10000 | 18 | 63 | 100469 | 9296.0 | 355935 |
| 443 | 1866 | 10 | 59 | 2934 | 375.0 | 35800 |
| 444 | 906 | 6 | 59 | 2933 | 193.9 | 4986 |
| 445 | 2300 | 9 | 57 | 2910 | 182.9 | 2739 |
| 446 | 875 | 7 | 50 | 2905 | 132.0 | 5009 |
| 447 | 1758 | 5 | 62 | 2896 | 16.6 | 2855 |

变量说明见表4：

表4 变量定义及说明

|  |  |
| --- | --- |
| variable | label |
| obs | TOP500公司序号 |
| Y | 高管薪水及分红 |
| X2 | 担任 CEO年数（不足6个月视为的0） |
| X3 | 总裁CEO的年龄 |
| X4 | 企业的总销售收入 |
| X5 | 企业的利润 |
| X6 | 企业的总资产 |

请考虑如下样本回归模型：

请回答如下问题:

1. 根据回归模型(22)，写出总体回归模型（PRM），并对参数的理论预期（符号、大小、关系）进行说明。
2. 利用Eviews对样本回归模型(22)进行回归分析（将报告截图过来，并写出相应的简要报告形式——三行式或四行式）。参数估计结果符合你的理论预期么？
3. 回归模型(22)存在异方差问题的证据吗？请以此按照下列方法进行诊断，并分别得到分析结论（要求截图过来并进行简要说明）：
   1. 非正式检验法（图解法）：
      * 绘制序列和序列的描点图（dot plot），得到你的初步结论。
      * 分别绘制序列分别与、序列的散点图（scatter plot），得到你的初步结论
      * 分别绘制与、序列的散点图（scatter plot），得到你的初步结论
   2. 正式检验法
      * 利用Park检验法，并得出你的初步结论（要求写出park诊断方程，并将park检验结果截图过来）
      * 利用Glejser检验法，并得出你的初步结论(请确定X，并做三个类型的辅助回归。分别将Glejser检验结果截图过来）
      * 利用BPG检验法，并得出你的初步结论（要求写出BPG诊断方程，并将BPG检验结果截图过来）
      * 利用White检验法（注意交叉项的使用），并得出你的初步结论（要求写出White诊断方程，并将White检验结果截图过来）
4. 若发现存在异方差问题，你如何进行纠正主模型(22)？
   1. 使用White校正法纠正异方差问题，并比较与主模型(22)的差别。（要求分别截图两个回归方程的Eviews报告，进行对照分析并得到结论）
   2. 使用加权最小二乘法纠正异方差问题（提醒：根据前述分析，几种处理方法选择一种合适的处理方法，进行模型矫正分析。请说明你选择这种处理办法的理由，并得到分析结论！）
5. 现在做以 为因变量的第二个模型。异方差性有所改善吗?（要求截图相关Eviews报告，并简要陈述理由）。

## 作业题2

**存货案例**：表5给出给出了41年的Y存货（百万美元），X2销售额（百万美元）等方面数据。

表5 存货与销售额(n=41)

|  |  |  |
| --- | --- | --- |
| Year | Y | X2 |
| 1950 | 84646 | 46486 |
| 1951 | 90560 | 50229 |
| 1952 | 98145 | 53501 |
| 1953 | 101599 | 52805 |
| 1954 | 102567 | 55906 |
| 1986 | 423082 | 326227 |
| 1987 | 408226 | 334616 |
| 1988 | 439821 | 359081 |
| 1989 | 479106 | 394615 |
| 1990 | 509902 | 411663 |

变量说明见表6：

表6 变量定义及说明

|  |  |
| --- | --- |
| variable | label |
| Year | 年份 |
| Y | 存货（百万美元） |
| X2 | 销售额（百万美元） |

请考虑如下样本回归模型：

请回答如下问题:

1. 根据总体回归模型(23)，请对模型参数的理论预期（符号、大小、关系）进行说明。
2. 利用Eviews对模型(23)进行回归分析（将报告截图过来，并写出相应的简要报告形式——三行式或四行式）。参数估计结果符合你的理论预期么？
3. 回归模型(23)存在自相关问题的证据吗？请按照下列方法分别进行诊断，并分别得到分析结论（要求截图过来并进行简要说明）:
   1. 残差序列观察法（描点图法）：绘制序列的描点图（dot plot），得到你的初步结论。
   2. 残差序列观察法（描点图法）：根据前述观察，确定滞后阶数并分别绘制序列与序列的散点图（scatter plot），得到你的初步结论
   3. 辅助回归法：构建残差序列对序列的辅助回归方程(25)，得到你的初步结论
   4. 自相关和偏相关分析法：Eviews菜单操作对残差序列进行自相关和偏相关分析（**注意滞后阶数的选择**）
   5. Durbin-Watson检验法：分析Eviews报告中的D-W统计量
   6. 拉格朗日检验法（LM-test）：Eviews菜单操作进行布罗施-戈弗雷(Breusch-Goldfrey)的拉格朗日检验(B-G LM test)
4. 若发现回归模型(23)存在自相关问题，你如何对该模型进行矫正？请按照下列方法分别进行诊断，并分别得到分析结论（要求截图过来并进行简要说明）：
   1. 广义最小二乘法(GLS)：一阶差分法变换
   2. 广义最小二乘法(GLS)：基于残差辅助方程(25)中估计出来的近似得到。
   3. 广义最小二乘法(GLS)：基于D-W统计量近似计算得到。
   4. 广义最小二乘法(GLS)：基迭代法近似计算得到。
   5. 一致标准误校正法（HAC）：尼威-威斯特(Newey-West)校正法
   6. 对数变换法：对数模型(24)进行Eviews回归分析
5. 如果模型(23)与上面各种方法修正的结果相差较大，你认为哪一个方法要相对较好，为什么？