第8章 工具变量法

- OLS 能够成立的最重要条件是解释变量与扰动项不相关(前定变量或同期外生)。否则，OLS 不一致。
- 但解释变量与扰动项相关(内生性)的例子比比皆是。
- 解决内生性的主要方法之一为工具变量法。
- 内生性的来源包括**遗漏变量偏差**、**联立方程偏差(双向因果关系)**，及**测量误差偏差**(measurement error bias)。

联立方程偏差

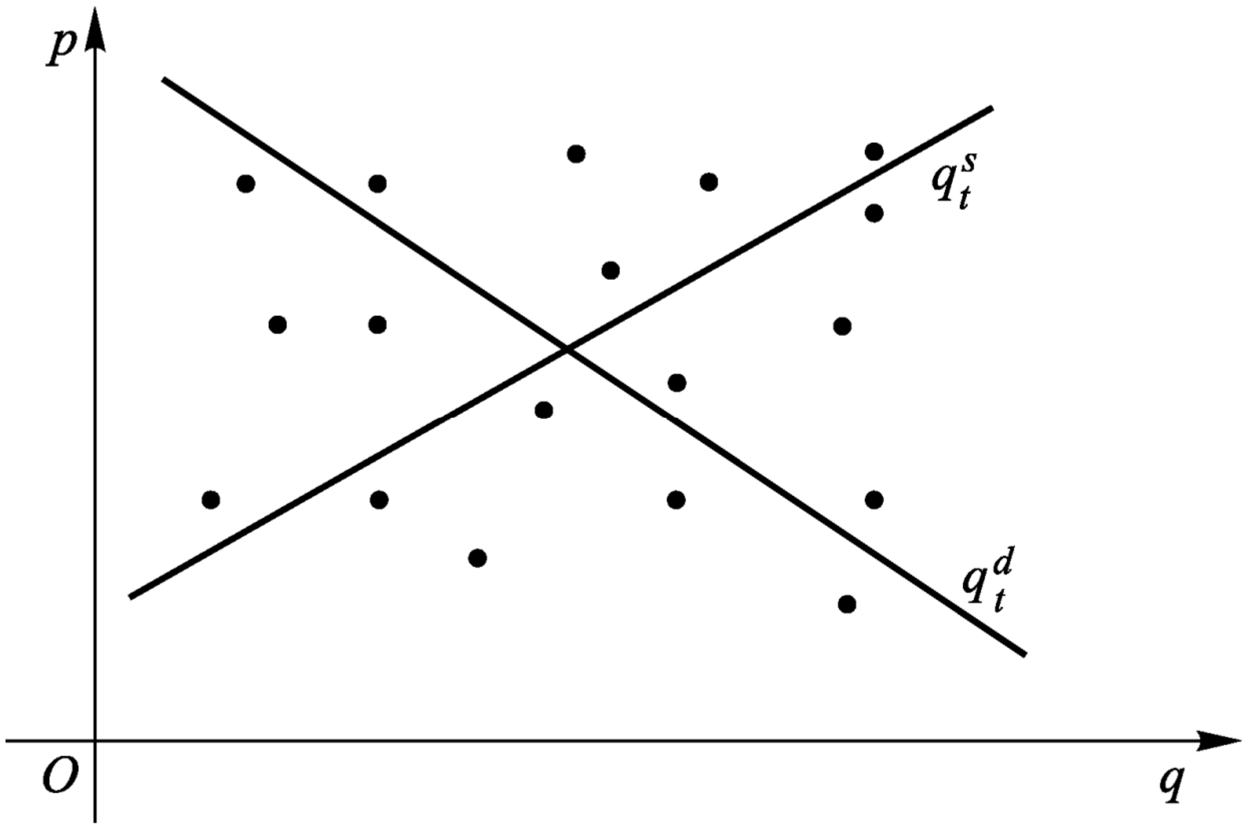
- **例** 考察农产品市场均衡模型：

$$\begin{cases} q_t^d = \alpha + \beta p_t + u_t & (\text{需求}) \\ q_t^s = \gamma + \delta p_t + \nu_t & (\text{供给}) \\ q_t^d = q_t^s & (\text{均衡}) \end{cases}$$

- q_t^d为农产品需求，q_t^s为农产品供给，p_t为农产品价格。
- 市场出清(market clearing)的均衡条件要求$q_t^d = q_t^s$。
- 令$q_t = q_t^d = q_t^s$，可得

$$\begin{cases} q_t = \alpha + \beta p_t + u_t \\ q_t = \gamma + \delta p_t + \nu_t \end{cases}$$

- 两个方程的被解释变量与解释变量完全一样。
- 如直接作回归$q_t \xrightarrow{OLS} p_t$,估计的是需求还是供给函数?



- 把线性方程组的(p_t, q_t)看成是未知数(内生变量)，把(u_t, v_t)看作已知，可求解(p_t, q_t)为(u_t, v_t)的函数。
- 故解释变量p_t与两个方程的扰动项(u_t, v_t)都相关，即$Cov(p_t, u_t) \neq 0$，$Cov(p_t, v_t) \neq 0$。

- 对于需求函数的正冲击$(u_t > 0)$，使均衡价格p_t上升，故二者正相关。
- 对于供给函数的正冲击$(v_t > 0)$，使均衡价格p_t下降，故二者负相关。
- 故OLS不一致，称为“联立方程偏差”(simultaneity bias)或“内生性偏差”(endogeneity bias)。

- **例** 考察宏观经济模型中的消费函数：

$$\begin{cases} C_t = \alpha + \beta Y_t + \varepsilon_t \\ Y_t = C_t + I_t + G_t + X_t \end{cases}$$

- Y_t，C_t，I_t，G_t，X_t分别为国民收入、总消费、总投资、政府净支出与净出口。
- 第一个方程为消费方程，第二个方程为国民收入恒等式。
- 如单独对消费方程进行 OLS 回归，存在联立方程偏差，得不到一致估计。

测量误差偏差

- 内生性的另一来源是解释变量的测量误差(measurement error 或 errors-in-variables)。
- 例 假设真实模型为

$$y = \alpha + \beta x^* + \varepsilon$$

- 其中，$\beta \neq 0$，$Cov(x^*, \varepsilon) = 0$。
- x^*无法观测，只能观测到x，二者满足如下关系：

$$x = x^* + u$$

- 其中，$Cov(x^*, u) = 0$，$Cov(u, \varepsilon) = 0$。
- 代入上式可得，

$$y = \alpha + \beta x + (\varepsilon - \beta u)$$

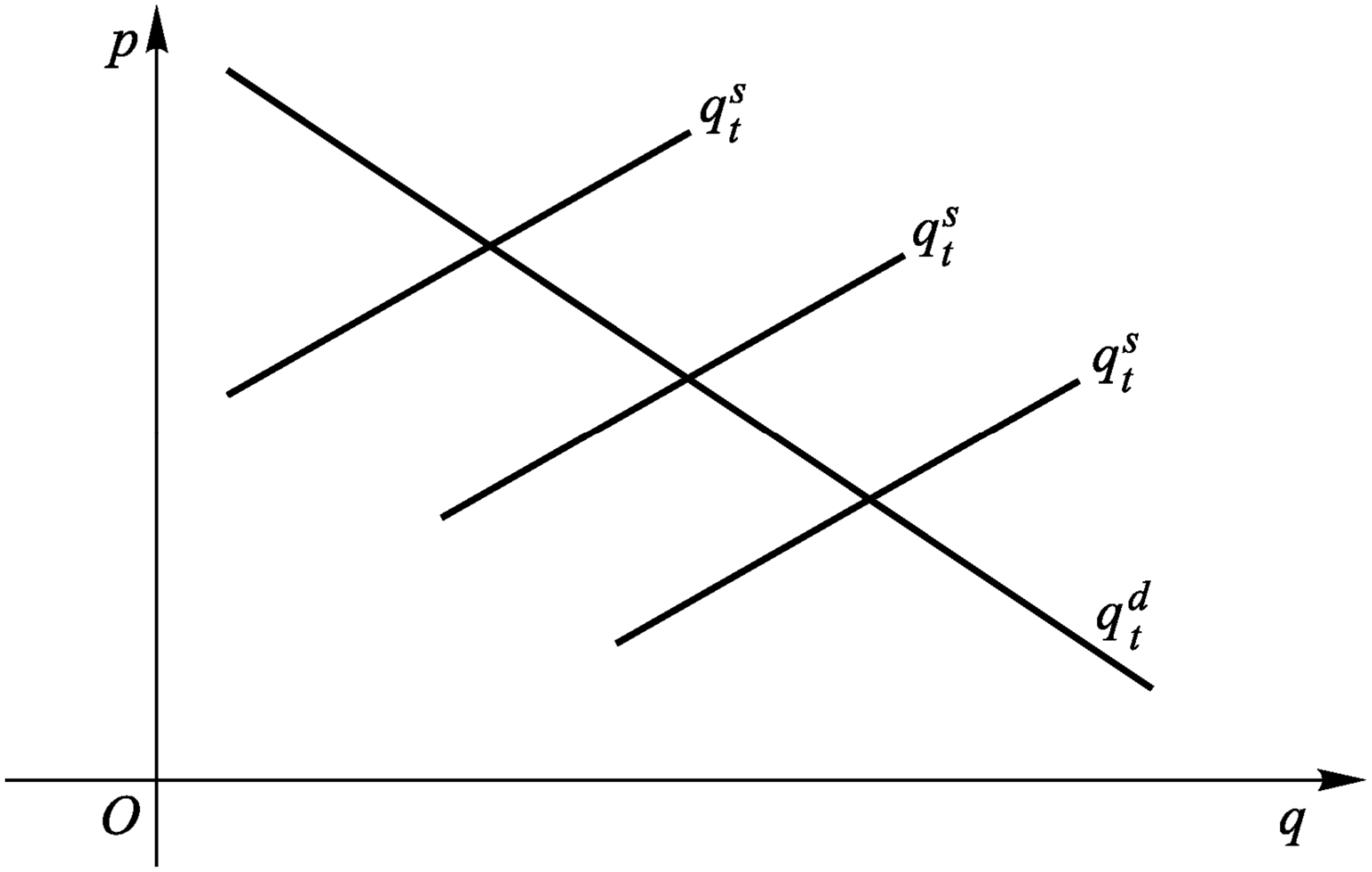
- 新扰动项$(\varepsilon - \beta\mu)$与解释变量x存在相关性：

$$
\begin{aligned}
\mathrm{Cov}(x, \varepsilon - \beta u) &= \mathrm{Cov}(x^* + u, \varepsilon - \beta u) \\
&= \underbrace{\mathrm{Cov}(x^*, \varepsilon)}_{=0} - \beta \underbrace{\mathrm{Cov}(x^*, u)}_{=0} + \underbrace{\mathrm{Cov}(u, \varepsilon)}_{=0} - \beta \mathrm{Cov}(u, u) \\
&= -\beta \mathrm{Var}(u) \neq 0
\end{aligned}
$$

- 故 OLS 不一致，称为“测量误差偏差”(measurement error bias)。
- 如果被解释变量存在测量误差，后果却不严重。
- 比如，只要被解释变量的测量误差与解释变量不相关，则 OLS依然一致。

工具变量法

- OLS 不一致因内生变量与扰动项相关而引起。
- 如能将内生变量分成两部分，一部分与扰动项相关，另一部分与扰动项不相关，可用与扰动项不相关的那部分得到一致估计。
- 通常借助另外一个“工具变量”实现这种分离。
- 假设在农产品市场均衡的图中，存在某因素使供给曲线经常移动，而需求曲线基本不动。
- 可估计需求曲线。使得供给曲线移动的变量就是工具变量。

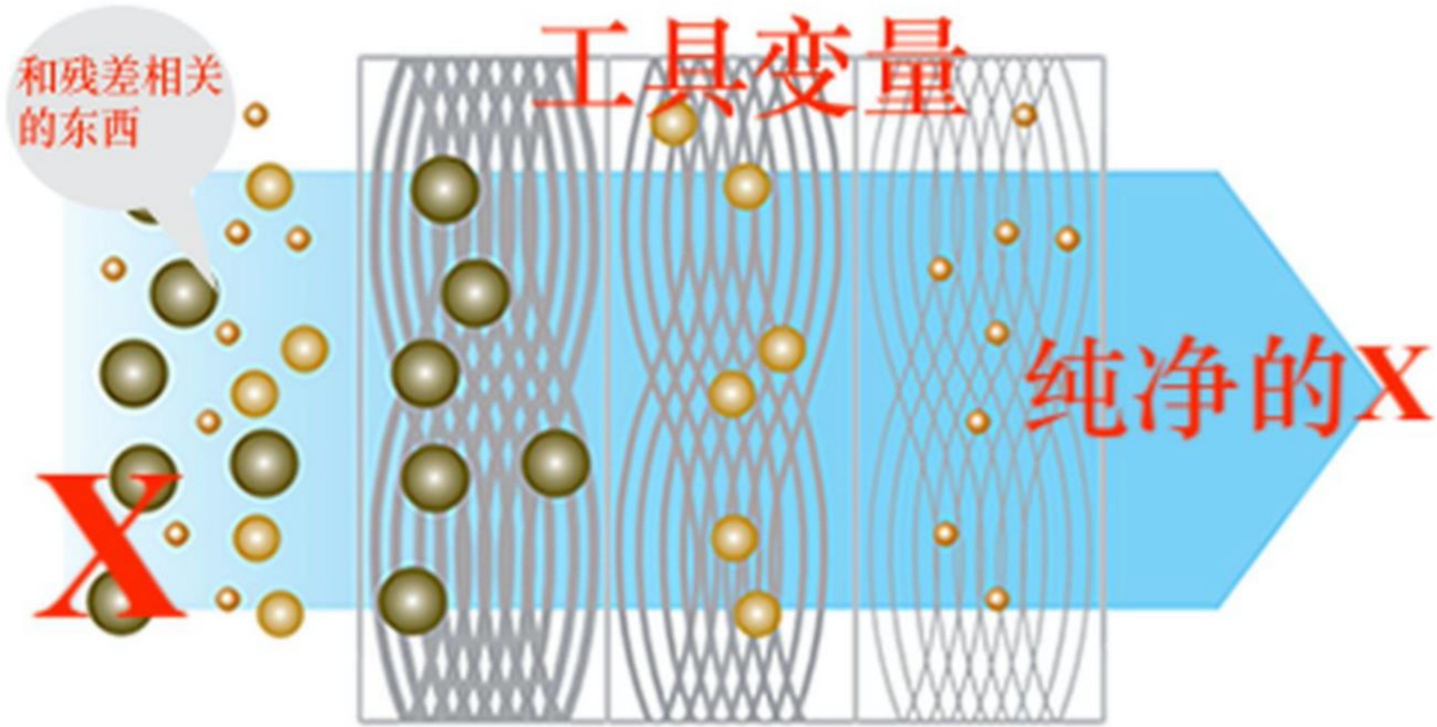


p
q_t^s
q_t^s
q_t^s
q_t^d
O
q

- 假设影响供给方程扰动项的因素可分解为两部分，即可观测的气温z_t与不可观测的其他因素：

$$q_t^s = \gamma + \delta p_t + \eta z_t + \nu_t$$

- 假定气温z_t是前定变量，与需求方程的扰动项不相关，即$Cov(z_t, u_t) = 0$。
- 由于气温z_t的变化使得供给函数q_t^s沿着需求函数q_t^d移动，故可估计需求函数q_t^d。
- 称z_t为“工具变量”(Instrumental Variable，IV)。


和残差相关
的东西
工具变量
纯净的X
X

- 需求方程为

$$q_t = \alpha + \beta p_t + u_t$$

- 在回归方程中，一个**有效(valid)的工具变量**应满足以下两个条件。
- (i) **相关性(relevance)**：工具变量与内生解释变量相关，即 $Cov(z_t, p_t) \neq 0$。
- (ii) **外生性 (exogeneity)**： 工具变量与扰动项不相关，即 $Cov(z_t, u_t) = 0$。

在本例中，气温z_t满足这两个条件。

- (i) 相关性：从联立方程组可解出$p_t = p_t(z_t, \mu_t, \nu_t)$，故$Cov(z_t, p_t) \neq 0$。
- (ii) 外生性：假设气温z_t是前定变量，$Cov(z_t, \mu_t) = 0$。

- 利用工具变量的这两个性质，可得到对需求方程回归系数β的一致估计。
- 需求方程为

$$q_t = \alpha + \beta p_t + \mu_t$$

- 需求方程两边同时求与z_t的协方差：

$$\begin{aligned}\mathrm{Cov}(q_t, z_t) &= \mathrm{Cov}(\alpha + \beta p_t + u_t, z_t) \\ &= \beta \mathrm{Cov}(p_t, z_t) + \underbrace{\mathrm{Cov}(u_t, z_t)}_{=0} = \beta \mathrm{Cov}(p_t, z_t)\end{aligned}$$

- 由于工具变量的外生性，故$Cov(u_t, z_t) = 0$。
- 根据工具变量的相关性，$Cov(p_t, z_t) \neq 0$。

- 两边同除$Cov(p_t,z_t)$：

$$\beta=\frac{Cov(q_t,z_t)}{Cov(p_t,z_t)}$$

- 以样本矩取代总体矩(以样本协方差替代总体协方差)，可得一致的“工具变量估计量”(Instrumental Variable Estimator)：

$$\hat{\beta}_{\mathrm{IV}}=\frac{\widehat{\mathrm{Cov}(q_t,z_t)}}{\widehat{\mathrm{Cov}(p_t,z_t)}}=\frac{\sum_{t=1}^{n}(q_t-\bar{q})(z_t-\bar{z})}{\sum_{t=1}^{n}(p_t-\bar{p})(z_t-\bar{z})}\xrightarrow{p}\frac{\mathrm{Cov}(q_t,z_t)}{\mathrm{Cov}(p_t,z_t)}=\beta$$

- $\bar{q},\bar{p},\bar{z}$分别为q,p,z的样本均值。
- 如工具变量与内生变量无关，$Cov(z_t,p_t)=0$，则无法定义工具变量法。
- 如果工具变量与内生变量的相关性很弱，$Cov(z_t,p_t)\approx 0$，会导致估计量$\hat{\boldsymbol{\beta}}_{IV}$的方差变得很大，称为“弱工具变量问题”。

二阶段最小二乘法

- 工具变量法一般通过“二阶段最小二乘法”(Two Stage Least Square，2SLS 或 TSLS)来实现。
- 第一阶段回归：用内生解释变量对工具变量回归，即$p_t \xrightarrow{OLS} z_t$，得到拟合值$\hat{p}_t$。
- 第二阶段回归：用被解释变量对第一阶段回归的拟合值进行回归，即$q_t \xrightarrow{OLS} \hat{p}_t$。

- 为什么这样做能得到一致估计?
- 首先，把需求方程$q_t = \alpha + \beta p_t + u_t$分解为

$$q_t = \alpha_0 + \beta \hat{p}_t + \underbrace{\left[u_t + \beta(p_t - \hat{p}_t)\right]}_{\equiv \varepsilon_t}$$

- 这是第二阶段的回归，扰动项为$\varepsilon_t \equiv \mu_t + \beta(p_t - \hat{p}_t)$。
- **命题** 在第二阶段回归中，$\hat{p}_t$与扰动项ε_t不相关。
- **证明：** 由于$\varepsilon_t \equiv \mu_t + \beta(p_t - \hat{p}_t)$，故

$$\mathrm{Cov}(\hat{p}_t, \varepsilon_t) = \mathrm{Cov}(\hat{p}_t, u_t) + \beta\,\mathrm{Cov}(\hat{p}_t, p_t - \hat{p}_t)$$

- 首先，由于$\hat{p}_t$是z_t的线性函数（$\hat{p}_t$为第一阶段回归的拟合值），而$Cov(z_t, u_t) = 0$（工具变量的外生性），故$Cov(\hat{p}_t, u_t) = 0$。
- 其次，在第一阶段回归中，拟合值$\hat{p}_t$与残差$(p_t - \hat{p}_t)$正交（OLS的正交性），故$Cov(\hat{p}_t, p_t - \hat{p}_t) = 0$。
- 第二阶段回归的解释变量$\hat{p}_t$与扰动项ε_t不相关，故2SLS一致。

- **2SLS的实质**：把内生解释变量p_t分成两部分，由工具变量z_t所造成的外生部分$(\hat{p}_t)$，及与扰动项相关的其余部分$(p_t - \hat{p}_t)$；
- 把被解释变量q_t对p_t中的外生部分$(\hat{p}_t)$进行回归，从而满足OLS对前定变量的要求而得到一致估计。

- 如存在多个工具变量，仍可用 2SLS 法。
- 假设z_1与z_2为两个有效工具变量（满足相关性与外生性），则第一阶段回归变为

$$p = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + u$$

可得拟合值$\hat{p} = \hat{\alpha}_0 + \hat{\alpha}_1 z_1 + \hat{\alpha}_2 z_2$，则第二阶段回归不变。

- 考虑多个内生变量的情形：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

- 其中，x_1与x_2均内生，都与ε相关。
- 由于有两个内生变量，至少需要两个工具变量，才能进行 2SLS估计。

- 如只有一个工具变量z，由第一阶段回归可得，$\hat{x}_1 = \hat{\alpha}_0 + \hat{\alpha}_1 z$，$\hat{x}_2 = \hat{\gamma}_0 + \hat{\gamma}_1 z$。
- 将$\hat{x}_1$与$\hat{x}_2$代入原方程：

$$y = \beta_0 + \beta_1 \hat{x}_1 + \beta_2 \hat{x}_2 + \nu_t$$

$\hat{x}_1$与$\hat{x}_2$都是z的线性函数，故存在严重多重共线性。

- **阶条件**：进行 2SLS 估计的必要条件是工具变量个数不少于内生解释变量的个数，称为“阶条件”(order condition)。
- (1) 不可识别(unidentified)：工具变量个数小于内生解释变量个数，无法使用2SLS；
- (2) 恰好识别(just or exactly identified)：工具变量个数等于内生解释变量个数;
- (3) 过度识别(overidentified)：工具变量个数大于内生解释变量个数。
- 在恰好识别与过度识别的情况下，都可使用 2SLS;
- 在不可识别的情况下，无法使用 2SLS。

- 考虑多个内生变量，且包含外生解释变量的情形：
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 w + \varepsilon$$
- x_1与x_2为内生变量，w为外生变量（与ε不相关）。假设有三个有效工具变量z_1,z_2,z_3。
- 在第一阶段回归中，分别将两个内生变量(x_1, x_2)对所有外生变量(包括工具变量z_1,z_2,z_3及外生变量w）回归：

$$x_1 = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3 + \alpha_4 w + u$$

$$x_2 = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \gamma_3 z_3 + \gamma_4 w + v$$

- 外生变量w可视为自己的工具变量，因为满足工具变量的定义。
- 首先，w与w高度相关，满足相关性。
- 其次，w与扰动项ε不相关，因为w为外生变量。
- 有时也称z_1, z_2, z_3为方程外的工具变量。
- 将上述两个方程的拟合值分别记为$\hat{x}_1$，$\hat{x}_2$，并代入原方程，进行第二阶段回归：

$$y = \beta_0 + \beta_1 \hat{x}_1 + \beta_2 \hat{x}_2 + \beta_3 w + \xi$$

ξ为第二阶段回归的扰动项。记此估计量为$\widehat{\boldsymbol{\beta}}_{IV}$。

- **注意1** 如果不将外生解释变量w放入第一阶段回归，则不能保证第一阶段残差与w正交（不相关），而w仍出现在第二阶段回归中，故导致2SLS不一致。
- **注意2**，第二阶段回归所得残差为

$$\hat{\xi} = y - (\hat{\beta}_0 + \hat{\beta}_1\hat{x}_1 + \hat{\beta}_2\hat{x}_2 + \hat{\beta}_3 w)$$

- 原方程真正的残差却是

$$e_{IV} \equiv y - (\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 w)$$

- 二者并不相等，即$e_{IV} \neq \hat{\xi}$
- 进行 2SLS 估计，最好不要自己手工进行两次回归，而直接使用 Stata 命令。

- $\widehat{\boldsymbol{\beta}}_{IV}$为$\boldsymbol{\beta}$的一致估计，且渐近正态，可照常进行大样本统计推断。
- 2SLS 的第二阶段回归就是 OLS，故$\widehat{\boldsymbol{\beta}}_{IV}$的协方差矩阵$Var(\widehat{\boldsymbol{\beta}}_{IV})$在形式上与OLS估计量相似。
- 考虑到可能存在异方差，建议使用异方差稳健的标准误。

- Stata命令格式
- `ivregress 2sls y x1 x2 (x3 = z1 z2),robust first`

- 在球形扰动项的情况下，2SLS 是最有效率的工具变量法。
- 在异方差的情况下，存在更有效率的工具变量法，即“广义矩估计”(Generalized Method of Moments，GMM)。
- GMM 是数理统计“矩估计”(Method of Moments，MM)的推广。
- GMM 之于 2SLS，正如 GLS 与 OLS 的关系。
- 在恰好识别或同方差的情况下，GMM 等价于 2SLS。

弱工具变量

- 如工具变量与内生变量仅微弱相关，$\hat{\boldsymbol{\beta}}_{IV}$的方差将变得很大。
- 由于工具变量仅包含极少与内生变量有关的信息，利用这部分信息进行的工具变量法估计就不准确。
- 这种工具变量称为“弱工具变量”(weak instruments)。
- 弱工具变量的后果类似于样本容量过小，会导致$\hat{\boldsymbol{\beta}}_{IV}$的小样本性质变得很差，即$\hat{\boldsymbol{\beta}}_{IV}$的小样本真实分布离大样本的渐近正态分布相去甚远，致使基于大样本的统计推断失效。

- 为检验是否存在弱工具变量，可在第一阶段回归中，检验所有方程外的工具变量的系数是否联合为零。
- 假设原模型为

$$y = \beta_0 + \beta_1 x + \beta_2 w + \varepsilon$$

x为内生变量，w为外生变量。

- 假设有两个有效工具变量z_1,z_2，第一阶段回归为

$$x = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 w + \mu$$

- 检验H_0: $\alpha_1 = \alpha_2 = 0$。
- 由于工具变量的强弱连续变化，很难确定明确的检验标准。
- 经验规则：此检验的**F 统计量大于 10** (由于技术性原因，此处使用普通标准误)，则拒绝“存在弱工具变量”的原假设。

- 如发现存在弱工具变量，可能的解决方法包括：
- (i) 寻找更强的工具变量；
- (ii) 使用对弱工具变量更不敏感的“有限信息最大似然估计法”(Limited Information Maximum Likelihood Estimation，LIML)。
- 在大样本下，LIML 与 2SLS 渐近等价。
- 在弱工具变量的情况下，LIML 的小样本性质可能优于 2SLS。

对工具变量外生性的过度识别检验

- 工具变量的外生性是保证 2SLS 一致性的重要条件。
- 如果“工具变量”与扰动项相关，可导致严重的偏差。
- 在恰好识别的情况下，无法检验工具变量的外生性。
- 只能进行定性讨论或依赖于专家的意见。
- 定性讨论：如果工具变量外生，则它影响被解释变量的唯一渠道就是通过内生变量，除此以外别无其他渠道。
- 由于此唯一渠道(内生变量)已包括在回归方程中，故工具变量不会再出现在被解释变量的扰动项中，或对扰动项有影响。
- 此条件称为“排他性约束”(exclusion restriction)，它排除了工具变量除了通过内生变量而影响被解释变量的其他渠道。

- 实践中，需找出工具变量影响被解释变量的所有其他可能渠道，然后一一排除，才能说明工具变量的外生性。
- 在过度识别情况下，可进行“过度识别检验”(overidentification test)。
- 过度识别检验的大前提(maintained hypothesis)是该模型至少恰好识别，即有效工具变量至少与内生解释变量一样多。
- 在此大前提下，过度识别检验的原假设为

$$H_0\text{: 所有工具变量都外生}$$

- 如拒绝原假设，则认为至少某个变量与扰动项相关。

- 假设共有 K个解释变量$\{x_1, \ldots, x_K\}$，其中前$(K-r)$个解释变量$\{x_1, \ldots, x_{K-}\}$为外生变量，而后r个解释变量$\{x_{K-r+1}, \ldots, x_K\}$为内生变量：

$$y = \underbrace{\beta_1 x_1 + \cdots + \beta_{K-r} x_{K-r}}_{\text{外生}} + \underbrace{\beta_{K-r+1} x_{K-r+1} + \cdots \beta_K x_K}_{\text{内生}} + \varepsilon$$

- 假设共有m个方程外的工具变量$\{z_1, \ldots, z_m\}$，其中$m > r$；则过度识别的原假设为

$$H_0: \ \mathrm{Cov}(z_1, \varepsilon) = 0, \ \cdots, \ \mathrm{Cov}(z_m, \varepsilon) = 0$$

- 通过2SLS 的残差e_{IV}考察工具变量与扰动项的相关性。

- 把e_{IV}对所有外生变量(所有外生变量与工具变量)进行辅助回归:

$$e_{IV} = \gamma_1 x_1 + \cdots + \gamma_{K-r} x_{K-r} + \delta_1 z_1 + \cdots + \delta_m z_m + error$$

- 原假设可写为

$$H_0: \delta_1 = \cdots = \delta_m = 0$$

- 记辅助回归的可决系数为R^2，Sargan 统计量为

$$nR^2 \xrightarrow{d} \chi^2(m-r)$$

- Sargan 统计量的渐近分布为$\chi^2(m-r)$，其自由度$(m-r)$是过度识别约束的个数，即方程外工具变量个数(m)，减去内生变量个数(r)，也就是“多余”的工具变量个数。
- 如恰好识别，则$m-r=0$（自由度为0），$\chi^2(0)$无定义，无法使用“过度识别检验”。

- 此检验的直观思想：在过度识别的情况下，可用不同的工具变量组合来进行工具变量法估计。
- 如果所有工具变量都有效，则这些不同的工具变量估计量$\hat{\boldsymbol{\beta}}_{IV}$都将收敛到相同的真实参数$\boldsymbol{\beta}$。
- 可检验不同的工具变量估计量之间的差是否收敛于 $\mathbf{0}$ ；如果不是，则说明这些工具变量不全有效。
- 在恰好识别的情况下，只有唯一的工具变量估计量，无法进行比较，故过度识别检验失效。
- 即使接受了过度识别的原假设，也并**不能**证明这些工具变量的外生性。
- 因为，即使不同的工具变量估计量$\hat{\boldsymbol{\beta}}_{IV}$的概率极限相同，并不能保证它们都收敛到真实的参数$\boldsymbol{\beta}$；也可能都收敛到其他值，比如$\boldsymbol{\beta}^* \neq \boldsymbol{\beta}$

- 事实上，过度识别检验还有一个大前提（maintained hypothesis），即该模型至少恰好识别，即有效工具变量至少与内生解释变量一样多。
- 此大前提无法检验，只能假定成立（可做定性讨论）。
- 比如，只有一个内生变量，在进行过度识别检验时，隐含地假定至少有一个工具变量外生，然后检验所有其他工具变量的外生性。
- 恰好识别的大前提保证了，在这些工具变量估计量至少有一个收敛到真实参数。此时，如果所有的工具变量估计量$\hat{\boldsymbol{\beta}}_{IV}$均收敛到同一地方，则都是一致估计。

对解释变量内生性的豪斯曼检验：究竟该用 OLS 还是 IV

- 使用工具变量法的前提是存在内生解释变量。
- 如何检验解释变量是否内生？
- 扰动项不可观测，无法直接检验解释变量与扰动项的相关性。
- 如找到有效工具变量，可借助工具变量来检验。

- 假设存在方程外的有效工具变量。
- 如果所有解释变量都外生，则 OLS 与 IV 都一致，但 OLS 比 IV更有效。
- 在这种情况下，虽然 IV 一致，但相当于“无病用药”，反而增大估计量的方差。
- 反之，如果存在内生变量，则 OLS 不一致，而 IV 一致。

- “豪斯曼检验”(Hausman specification test)(Hausman, 1978)的原假设为“H_0：所有解释变量均为外生变量”。
- 如果H_0成立，则 OLS 与 IV 都一致，在大样本下$\hat{\boldsymbol{\beta}}_{IV}$与$\hat{\boldsymbol{\beta}}_{OLS}$都收敛于真实的参数值$\boldsymbol{\beta}$；故$(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})$依概率收敛于 **0**。
- 反之，如果H_0不成立，则 IV 一致而 OLS 不一致，故$(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})$不会收敛于 **0**。

- 如果$(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})$的距离很大，则倾向于拒绝原假设。
- 根据沃尔德检验原理，以二次型度量此距离：

$$(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})' \left[\widehat{Var(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})}\right]^{-1} (\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS}) \xrightarrow{d} \chi^2(r)$$

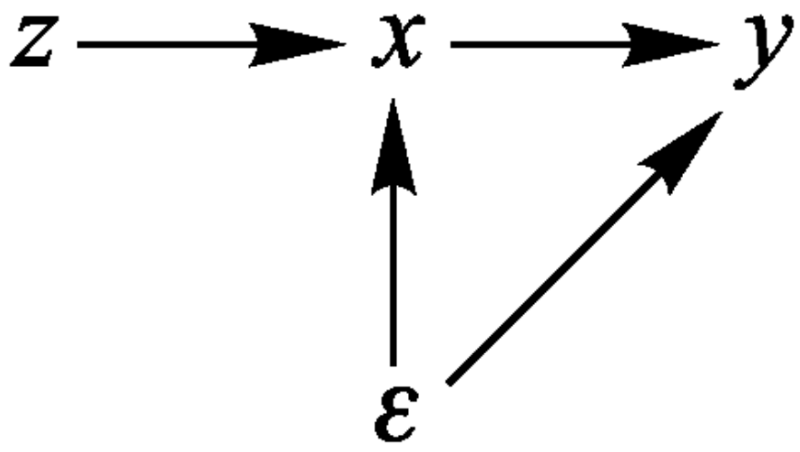
- r 为内生解释变量的个数，
- $\left[\widehat{Var(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})}\right]$为$(\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS})$的协方差矩阵之样本估计值。
- 如果此豪斯曼统计量很大，超过了其渐近分布$\chi^2(r)$的临界值，则拒绝“所有解释变量均外生”的原假设，认为存在内生变量，应使用 IV。

- 传统豪斯曼检验的缺点是，为简化矩阵 $\left[Var(\widehat{\hat{\boldsymbol{\beta}}_{IV} - \hat{\boldsymbol{\beta}}_{OLS}})\right]$ 的计算，假设在H_0成立的情况下，OLS最有效率，故不适用异方差的情形(OLS 只在球形扰动项的情况下才最有效率)。
- 改进的“杜宾-吴-豪斯曼检验”(Durbin-Wu-Hausman Test，DWH)在异方差的情况下也适用。

如何获得工具变量

- 工具变量的两个要求(相关性与外生性)常自相矛盾。
- 寻找合适的工具变量通常较困难，需要一定的创造性与想象力。
- 寻找工具变量的步骤大致可以分为两步:
- (i) 列出与内生解释变量(x)相关的尽可能多的变量的清单;
- (ii) 从这一清单中剔除与扰动项相关的变量。
- 第(ii)步的操作较难，因为扰动项不可观测。

- 如何判断候选变量(z)是否与扰动项(ε)相关?
- 由于扰动项是y的扰动项，可从z与y的相关性着手。
- z与y相关，因为z与内生变量x相关。
- 但z对y的影响仅通过x起作用，因为如果z与ε相关，则z对y的影响必然还有除x以外的渠道。
- 是否“z对y的影响仅通过x起作用”，可通过定性讨论来确定，即“排他性约束”(exclusion restriction)。



- **例 滞后变量**。对于时间序列或面板数据，常使用内生解释变量的滞后作为工具变量。
- 相关性：内生解释变量与其滞后变量相关。
- 外生性：由于滞后变量已经发生(从当期的角度看，其取值已经固定)，故可能“前定”，与当期扰动项不相关。
- 比如，Groves et al. (1994)考察国企改革(员工奖金激励制度)对企业生产率的作用。
- 奖金占员工报酬比重越高，则越能促进生产率的提高。
- 但生产率越高的企业越有能力给员工发奖金，存在双向因果。
- Groves et al. (1994)使用奖金比重的滞后值作为当期奖金比重的工具变量。二者的相关性显然。另一方面，当期生产率不可能影响过去的奖金比重，故奖金比重的滞后值(可能)具有外生性。

- **例** 警察人数与犯罪率。
- 警察人数越多，执法力度越大，犯罪率应越低。
- 但城市犯罪率高，政府可能增加警察人数。
- Levitt(1997)使用“市长选举的政治周期”作为犯罪率的工具变量。
- 市长竞选连任时，为拉选票，会增加警察人数以保证治安，故满足相关性。
- 选举周期以机械方式确定，除影响警察人数外，不单独对犯罪率起作用，故满足外生性。

- **例** 制度对经济增长的影响。
- 好制度促进经济增长，但制度变迁也依赖经济增长。
- 从历史角度，Acemoglu et al. (2001)使用“殖民者死亡率”(settler mortality)作为制度的工具变量。
- 欧洲殖民者在全球殖民时，由于各地气候及疾病环境(disease environment)不同，殖民者死亡率十分不同。
- 在死亡率高的地方(比如非洲)，殖民者难以定居，在当地建立掠夺性制度(extractive institutions)。
- 在死亡率低的地方(比如北美)，建立有利于经济增长的制度(比如较好的产权保护)。
- 初始制度上的差异一直延续到今天。故殖民者死亡率与今天的制度相关，满足相关性。
- 殖民者死亡率除影响制度外，不对当前的经济增长有任何直接影响，故满足外生性。

- **例** 看电视过多引发小儿自闭症?
- 在美国，电视普及与小儿自闭症发生率的攀升几乎同步。
- Waldman et al. (2006, 2008)研究看电视是否引发小儿自闭症。
- 有自闭倾向的儿童可能更经常看电视，存在双向因果关系。
- 使用降雨量作为电视观看时间的工具变量。
- 降雨越多的地区，人们呆在室内时间越长，看电视时间也越长，故相关。
- 降雨量很可能外生 (只通过看电视时间而影响被解释变量)。
- 研究结果支持过多观看电视为小儿自闭症的诱因。