Single-Equation GMM

José Jaén Delgado

Introduction - In this paper solutions to most exercises from Chapter 3 of Fumio Hayashi's *Econometrics* are provided

1 Questions for Review

1.1 Endogeneity Bias

1. Estimating causal effects in demand and supply functions is imperative for econometricians. Note that precise curve fitting is not the main interest but rather gaining insights from an interpretable model (thus no Machine Learning algorithm is considered as of now). Nevertheless, a naive approach of blindly employing OLS for such tasks procures undesired results.

Suppose we seek to estimate the following Linear Regression Models:

$$q_i^D(p_i) = \alpha_0 + \alpha_1 p_i + u_i$$
$$q_i^S(p_i) = \beta_0 + \beta_1 p_i + v_i$$

Note that the same regressors are being used in both equations. Thus, it is not clear whether the constant and price features estimate demand or supply. In fact, it can be shown that no function is being correctly estimated as the input variable price is not predetermined. Causal Inference cannot be performed directly as we are no longer dealing with exogeneous regressors. The OLS estimate of the price coefficient is biased.

Let us illustrate our claim by combining economic theory and mathematical logic. In market equilibrium: $q_i^D = q_i^S$

$$\alpha_0 + \alpha_1 p_i + u_i = \beta_0 + \beta_1 p_i + v_i$$
$$p_i = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_i - u_i}{\alpha_1 - \beta_1}$$

Then the covariance between price and error term v_i is expressed as:

$$\mathrm{cov}(p_i, v_i) = \mathrm{cov}\left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_i - u_i}{\alpha_1 - \beta_1}, v_i\right)$$

$$= \mathrm{cov}\left(\frac{v_i - u_i}{\alpha_1 - \beta_1}, v_i\right) \quad \text{(since the first fraction is composed of parameters)}$$

$$= \frac{\mathrm{V}[v_i] - \mathrm{cov}(u_i, v_i)}{\alpha_1 - \beta_1} \quad (\neq 0)$$

Where $\alpha_1 < 0$ and $\beta_1 > 0$ from competitive market theory.

- If $\mathrm{V}[v_i] > \mathrm{cov}(u_i, v_i) \implies \mathrm{cov}(p_i, v_i) < 0$
- If $\mathrm{V}[v_i] < \mathrm{cov}(u_i, v_i) \implies \mathrm{cov}(p_i, v_i) > 0$
- If $\mathrm{V}[v_i] = \mathrm{cov}(u_i, v_i) \implies \mathrm{cov}(p_i, v_i) = 0$

Deriving the asymptotic expression of α_1:

$$\begin{aligned}\mathrm{cov}(p_i, q_i) &= \mathrm{cov}(p_i, \alpha_0 + \alpha_1 p_i + u_i) \\ &= \mathrm{cov}(p_i, \alpha_1 p_i + u_i) \quad \text{(since } \alpha_0 \text{ is a constant)} \\ &= \alpha_1 \mathrm{V}[p_i] - \mathrm{V}[u_i] \quad \text{(by the definition of } p_i)\end{aligned}$$

The probability limit of the OLS estimator of price takes the form:

$$\begin{aligned}\hat{\alpha}_1 &= \frac{\alpha_1 \mathrm{V}[p_i] - \mathrm{V}[u_i]}{\mathrm{V}[p_i]} \\ &= \alpha_1 - \mathrm{V}[u_i] \quad (\hat{\alpha_1} \neq \alpha_1)\end{aligned}$$

2. It has been shown that the OLS estimate for the price coefficient is biased in the demand and supply functions. It is also the case for the intercept vector $(\alpha_0,\, \beta_0)'$. Note that the asymptotic expression for the intercept is:

$$\hat{\alpha}_0 = \mathrm{E}[q_i] - \left(\frac{\mathrm{cov}(p_i, q_i)}{\mathrm{V}[p_i]}\right) \mathrm{E}[p_i]$$

Following the expression for the demand function:

$$\mathrm{E}[q_i] = \alpha_0 + \alpha_1 \mathrm{E}[p_i] \quad (\text{since } \ \mathrm{E}[u_i] = 0)$$

Then, the probability limit of the OLS estimate for the intercept can be expressed as:

$$\hat{\alpha}_0 = \alpha_0 - \left(\alpha_1 + \frac{\text{cov}(p_i, q_i)}{\text{V}[p_i]}\right) \text{E}[p_i]$$

$$\begin{aligned}
\hat{\alpha}_0 - \alpha_0 &= -\left(\alpha_1 + \frac{\text{cov}(p_i, q_i)}{\text{V}[p_i]}\right) \text{E}[p_i] \\
\hat{\alpha}_0 - \alpha_0 &\neq 0
\end{aligned}$$

3. Let us present yet another expression for the price coefficient, assuming that $\text{cov}(u_i, v_i) = 0$. We substitute p_i derived in **1.** in the supply function:

$$\begin{aligned}
q_i &= \beta_0 + \beta_1 \left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_i - u_i}{\alpha_1 - \beta_1}\right) + v_i \\
&= \beta_0 + \frac{\beta_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_1 v_i - \beta_1 u_i}{\alpha_1 - \beta_1} + v_i \\
&= \frac{\alpha_1\beta_0 - \beta_1\beta_0 + \beta_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_1 v_i - \beta_1 u_i + \alpha_1 v_i - \beta_1 v_i}{\alpha_1 - \beta_1} \\
&= \frac{\alpha_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1} + \frac{\alpha_1 v_i - \beta_1 u_i}{\alpha_1 - \beta_1}
\end{aligned}$$

Then the covariance between the target variable and feature p_i is:

$$\begin{aligned}
\text{cov}(p_i, q_i) &= \text{cov}\left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_i - u_i}{\alpha_1 - \beta_1}, \frac{\alpha_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1} + \frac{\alpha_1 v_i - \beta_1 u_i}{\alpha_1 - \beta_1}\right) \\
&= \frac{1}{(\alpha_1 - \beta_1)^2} \text{cov}\left(v_i - u_i, \alpha_1 v_i - \beta_1 u_i\right) \\
&= \frac{1}{(\alpha_1 - \beta_1)^2} \left(\alpha_1 \text{V}[v_i] + \beta_1 \text{V}[u_i]\right)
\end{aligned}$$

Let us calculate the variance of p_i to obtain the probability limit of the OLS estimator since: $\alpha_1 = \text{cov}(x, y)/\text{V}[x]$

$$\mathrm{V}[p_i] = \mathrm{V}\left[\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_i - u_i}{\alpha_1 - \beta_1}\right] = \frac{1}{(\alpha_1 - \beta_1)^2}\mathrm{V}\left[v_i - u_i\right]$$

$$= \frac{\mathrm{V}[v_i] + \mathrm{V}[u_i]}{(\alpha_1 - \beta_1)^2}$$

Finally (in probability limit terms):

$$\hat{\alpha_1} = \frac{\alpha_1 \mathrm{V}[v_i] + \beta_1 \mathrm{V}[u_i]}{\mathrm{V}[v_i] + \mathrm{V}[u_i]}$$

4. Let us propose a solution to consistently estimate the effect of price on quantity supplied or demanded. For such purpose, we need some observable factors that shift the supply curve. Suppose v_i can be decomposed into x_i and ξ_i. Imagine there is data available for the former but not for the latter. Further assume $\mathrm{cov}(x_i, u_i) = 0$ and also that $\mathrm{cov}(x_i, \xi_i) = 0$, so the observable supply shifter x_i is predetermined. Then the equations become:

$$\begin{aligned} q_i^D &= \alpha_0 + \alpha_1 p_i + u_i \\ q_i^S &= \beta_0 + \beta_1 p_i + \beta_2 x_i + \xi_i \end{aligned}$$

In market equilibrium $q_i^D = q_i^S$:

$$\begin{aligned} \alpha_0 + \alpha_1 p_i + u_i &= \beta_0 + \beta_1 p_i + \beta_2 x_i + \xi_i \\ p_i &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \beta_2 \frac{x_i}{\alpha_1 - \beta_1} + \frac{\xi_i - u_i}{\alpha_1 - \beta_1} \end{aligned}$$

Proving that p_i is still endogenous:

$$\mathrm{cov}(p_i, u_i) = \mathrm{cov}\left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \beta_2 \frac{x_i}{\alpha_1 - \beta_1} + \frac{\xi_i - u_i}{\alpha_1 - \beta_1}, u_i\right) = -\frac{\mathrm{V}[u_i]}{\alpha_1 - \beta_1}$$

For which $\mathrm{cov}(\xi_i, u_i) = 0$ was also assumed. Also, note that:

$$\mathrm{cov}(p_i, \xi_i) = \mathrm{cov}\left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \beta_2 \frac{x_i}{\alpha_1 - \beta_1} + \frac{\xi_i - u_i}{\alpha_1 - \beta_1}, \xi_i\right) = \frac{\mathrm{V}[\xi_i]}{\alpha_1 - \beta_1}$$

Since x_i is uncorrelated with $(u_i, \xi_i)'$, we could clean out p_i with the variation exaplined by x_i, obtaining a Local Average Treatment Effect (LATTE).

In econometrics, x_i is said to be an Instrumental Variable (IV), whose desirable properties (orthogonality with respect to the error terms and correlation with the variable of interest), will eventually allow to estimate the causal effect of p_i on q_i. Let us illustrate this with a Directed Acyclic Graph (DAG):

$$\begin{tikzcd}
x_i \arrow[r] & p_i \arrow[r] \arrow[d] & q_i \\
 & u_i \arrow[ur] &
\end{tikzcd}$$

p_i suffers from endogeneity but x_i does not. Also, x_i affects q_i only through p_i (imperative assumption!). x_i is a valid instrument only if this last characteristic holds.

In our example, think of x_i as temperature and q_i as coffee supplied/demanded. It is reasonable to assume that coffee production is directly affected by temperature, whereas consumers are not necessarily influenced by it. Changes in x_i will lead to the shifts in the supply function, while demand will remain unaltered. We can exploit this fact to estimate the causal effect of p_i on q_i due to temperature.

First, let us prove that x_i is correlated with p_i (instrument relevance):

$$\mathrm{cov}(x_i, p_i) = \frac{\beta_2}{\alpha_1 - \beta_1}\mathrm{V}[x_i] \quad (\neq 0)$$

Then, since we imposed restrictions on $\{x_i, u_i, \xi_i\}$ so that x_i is predetermined, temperature is uncorrelated with the demand error term (instrument exogeneity).

The Instrumental Variables (IV) Estimator can be expressed as:

$$\hat{\alpha_1}^{\mathrm{IV}} = \frac{\hat{\mathrm{cov}}(x_i, q_i)}{\hat{\mathrm{cov}}(x_i, p_i)}$$

This is unbiased since the probability limit of the numerator and denominator equals α_1:

$$\begin{aligned}
\mathrm{cov}(x_i, q_i) &= \alpha_1 \mathrm{cov}(x_i, p_i) + \mathrm{cov}(x_i, u_i) \\
&= \alpha_1 \mathrm{cov}(x_i, p_i)
\end{aligned}$$

So (asymptotically):

$$\hat{\alpha}_1^{\mathrm{IV}} = \frac{\hat{\mathrm{cov}}(x_i, q_i)}{\hat{\mathrm{cov}}(x_i, p_i)} = \alpha_1 \frac{\mathrm{cov}(x_i, p_i)}{\mathrm{cov}(x_i, p_i)} = \alpha_1$$

1.2 More Endogeneity Examples

1 . Endogeneity bias is ubiquitous and so econometricians must be well-equipped with statistical theory to effectively tackle it as to estimate causal effects. Let Friedman's Permanent Income Hypothesis serve as our first example:

$$C_i^* = kY_i^* \quad k \in (0,1)$$

Where C_i^* is permanent consumption, Y_i^* is permanent income, k is a scalar assumed to be uniformly distributed across each household i. It is also assumed that observed consumption C_i and observed income Y_i are error-ridden measured variables (microdata normally suffers from non-exact information provided by repondents). Formally:

$$\begin{aligned} C_i &= C_i^* + c_i \\ Y_i &= Y_i^* + y_i \end{aligned}$$

Further assumptions of Friedman's Hypothesis are listed below:

$$\begin{aligned} &\mathrm{E}[c_i] = 0,\ \mathrm{E}[C_i^* c_i] = 0,\ \mathrm{E}[Y_i^* c_i] = 0,\ \mathrm{E}[c_i y_i] = 0 \\ &E[y_i] = 0,\ \mathrm{E}[Y_i^* y_i] = 0,\ \mathrm{E}[C_i^* y_i] = 0 \end{aligned}$$

Regressing C_i^* on a constant and Y_i^* and following the error-riden expressions:

$$\begin{aligned} C_i^* &= \gamma_0 + \gamma_1 Y_i^* + u_i \quad \text{(original regression)} \\ C_i - c_i &= \gamma_0 + \gamma_1 (Y_i - y_i) + u_i \quad \text{(by definition of } \{C_i^*, Y_i^*\}) \\ C_i &= \gamma_0 + \gamma_1 Y_i + \xi_i \quad (\xi_i \coloneqq c_i - \gamma_1 y_i + u_i) \end{aligned}$$

Deriving the probability limit of the OLS estimator γ_1:

$$\begin{aligned} \mathrm{cov}(Y_i, C_i) &= \mathrm{cov}\left(Y_i, \gamma_0 + \gamma_1 Y_i + \xi_i\right) \\ &= \gamma_1 \mathrm{V}[Y_i] + \mathrm{cov}\left(\xi_i, Y_i\right) \\ &= \gamma_1 (\mathrm{V}[Y_i] - \mathrm{V}[y_i]) \end{aligned}$$

The variance of the regressor Y_i is:

$$\mathrm{V}[Y_i] = \mathrm{V}[Y_i^*] + \mathrm{V}[y_i]$$

Combining both expressions we obtain α_1:

$$\alpha_1 = \frac{\text{cov}(Y_i, C_i)}{\text{V}[Y_i]} = \gamma_1 \frac{\text{V}[Y_i] - \text{V}[y_i]}{\text{V}[Y_i^*] + \text{V}[y_i]}$$

Clearly, $\hat{\alpha_1} \overset{p}{\nrightarrow} \alpha_1$ and so the OLS estimator is biased. Suppose there existed a valid instrument x_i for Y_i so that $\text{E}[x_i \xi_i] = 0$ and $\text{E}[x_i Y_i] \neq 0$.

$$\alpha_1^{IV} = \frac{\text{E}[x_i C_i]}{\text{E}[x_i Y_i]}$$

The x_i that could make this possible is none other than a scalar $k = 1$. The IV estimator can be interpreted as the ratio of the sample mean of C_i to the sample mean of Y_i. Indeed, Friedman proposed such instrument for consistently estimating households consumption.

2. Another instance of endogeneity arises when economic agents make decisions based on factors for which no data has been collected. Let us consider a cross-sectional sample of firms maximizing profits. Specifically, A_i denotes the i-th firm's efficiency level, L_i represents the units of labor employed and v_i some technology shock. Formally:

$$Q_i = A_i L_i^{\phi_1} e^{v_i} \quad \phi_1 \in (0, 1)$$

Let $B \coloneqq \text{E}[e_i^v]$ and assume B is uniformly distributed across firms and time. The production function becomes:

$$Q_i = A_i L_i^{\phi_1} B$$

Introducing labor cost w_i (wages) and prices p_i in a competitive market with no technological advantages, firms maximize the following profit function:

$$\arg\max_{L_i} \Pi(L_i) = pA_i L_i^{\phi_1} B - wL_i$$

First Order Condition (FOC) takes the form:

$$\Pi'(L_i) = \phi_1 p A_i L_i^{\phi_1 - 1} B - w$$

Solving for L_i:

$$L_i = \left(\frac{w}{\phi_1 p A_i B}\right)^{\frac{1}{\phi - 1}}$$

Let us present new notation to be used:

$$u_i := \ln A_i - \mathrm{E}[\ln A_i], \quad \phi_0 := \mathrm{E}[\ln A_i], \quad A_i := e^{\phi_0 + u_i}$$

$$\begin{aligned}
\ln L_i &= \frac{1}{\phi_1 - 1}\left[\ln \frac{w}{p} - \ln \phi_1 A_i B\right] \\
&= \frac{1}{\phi_1 - 1}\left[\ln \frac{w}{p} - \ln \phi_1 B - \ln A_i\right] \\
&= \frac{1}{\phi_1 - 1}\left[\ln \frac{w}{p} - \ln \phi_1 B\right] - \frac{1}{\phi_1 - 1}(\phi_0 + u_i) \\
&= \underbrace{\frac{1}{\phi_1 - 1}\left[\ln \frac{w}{p} - \phi_0 - \ln \phi_1 B\right]}_{\beta_0} + \frac{u_i}{1 - \phi_1}
\end{aligned}$$

Taking logarithms in the production function:

$$\begin{aligned}
\ln Q_i &= \ln A_i L_i^{\phi_1} e^{v_i} \\
&= \phi_0 + u_i + \phi_1 \ln L_i + v_i
\end{aligned}$$

Notice that $(\ln L_i - \beta_0)(1 - \phi_1) = u_i$, thus:

$$\begin{aligned}
\ln Q_i &= \phi_0 + \phi_1 \ln L_i + v_i + (\ln L_i - \beta_0)(1 - \phi_1) \\
&= \phi_0 + \phi_1 \ln L_i + v_i + \ln Li - \beta_0 - \phi_1 \ln L_i + \phi_1 \beta_0 \\
&= \phi_0 + \phi_1 \beta_0 - \beta_0 + \ln L_i + v_i \\
&= \phi_0 + \phi_1 (1 - \beta_0) + \ln L_i + v_i
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(\ln L_i, \ln Q_i) &= \operatorname{cov}\left(\phi_0 + \phi_1(1 - \beta_0) + \ln L_i + v_i, \ln L_i\right) \\
&= \mathrm{V}[\ln L_i] \quad \text{(since technology is uniform)}
\end{aligned}$$

Consequently $\phi_1 = 1$.

3. Supposing firms could observe $(v_i, u_i)'$ before selecting labor input, we claim that $\ln Q_i$ would be an exact linear function of $\ln L_i$:

$$
\begin{aligned}
Q_i &= A_i \left[\left(\frac{w}{p} \right) \phi_1 A_i e^{v_i} \right]^{\frac{\phi_1}{1-\phi_1}} e^{v_i} \\
&= \left(\frac{w}{p} \right)^{\frac{\phi_1}{\phi_1 - 1}} \phi_1^{\frac{\phi_1}{\phi_1-1}} A_i^{\frac{1}{\phi_1-1}} e^{v_i \frac{1}{\phi_1 - 1}} \\
\ln Q_i &= \frac{\phi_1}{\phi_1 - 1} \ln \frac{w}{p} + \frac{\phi_1}{\phi_1 - 1} \ln \phi_1 A_i + \frac{1}{\phi_1 - 1} \ln e^{v_i}
\end{aligned}
$$

Note the absence of stochastic elements in $\ln Q_i$. Since we plugged $\ln L_i$ in the former variable, $\ln Q_i$ is an exact linear function of the latter.

1.3 The General Formulation

1. Let us consider the market equilibrium model of exercise 4 from the first subsection. It was assumed that an observable supply shifter x_i existed. Consider:

$$
\begin{aligned}
y_i &= \mathbf{z}_i' \boldsymbol{\delta} + \epsilon_i \quad \text{(general format)} \\
q_i &= \beta_0 + \beta_1 p_i + u_i \quad \text{(market equilibrium model)}
\end{aligned}
$$

Where $\mathbf{z}_i$ is an L-dimensional vector of regressors, $\boldsymbol{\delta}$ is an L-dimensional coefficient vector and ϵ_i is the error term for the i-th observation. Furthermore, let $\mathbf{x}_i$ represent a K-dimensional vector of instrumental variables, where $\mathrm{E}[\mathbf{x}_i \epsilon_i] = 0$ and $\mathrm{E}[\mathbf{x}_i \mathbf{z}_i'] < \infty$

Then, for the market equilibrium model:

$$
y_i = q_i, \;\; \boldsymbol{\delta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \;\; \mathbf{z}_i = \begin{bmatrix} 1 \\ p_i \end{bmatrix}, \epsilon_i = \xi_i, \;\; \mathbf{x}_i = \begin{bmatrix} 1 \\ x_i \end{bmatrix}
$$

Where $L = 2$ and $K = 2$ (just use `len`($\mathbf{z}_i$) and `len`($\mathbf{x}_i$) in `Python` to quickly see it). Since $L = K$, the demand equation is exactly identified. The resulting estimator when introducing x_i is the IV estimator.

2. Let us now consider a simplified version of the wage equation:

$$\ln w_i = \delta_0 + \delta_1 S_i + \delta_2 EXP_i + \delta_3 IQ_i + \epsilon_i$$

Where w_i is the wage for the i-th worker, S_i is completed years of schooling, EXP_i denotes experience in years and IQ_i is Intellectual Quotient. Assume all regressors but IQ_i are exogeneous, since the latter is an error-ridden measure of ability. Consider AGE_i (age in years) and MED_i (mother's education in years) for the i-th observation as instrumental variables. The general model takes the form:

$$y_i = \ln w_i, \ \boldsymbol{\delta} = \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}, \ \mathbf{z}_i = \begin{bmatrix} 1 \\ S_i \\ EXP_i \\ IQ_i \end{bmatrix}, \ \mathbf{x}_i = \begin{bmatrix} 1 \\ S_i \\ EXP_i \\ AGE_i \\ MED_i \end{bmatrix}$$

Then $L = 4$ and $K = 5$ (just like `length`$(\cdot)$ in `R`). Note that since $L < K$ the wage equation is overdetermined, assuming that the rank condition is satisfied, i.e: $\Pr\left(\text{rank}\left(\mathrm{E}[\mathbf{x}_i \mathbf{z}_i']\right) = L\right) = 1$. Thus, the resulting estimator when introducing $\mathbf{x}_i$ is the General Method of Moments (GMM) estimator.

3. The production function example presented in exercise 2 from the previous subsection considered no more than a single instrument (a constant) but there were two regressors (a constant and L_i). We write the orthogonality condition and verify there exist infinite combinations of (ϕ_0, ϕ_1) that satisfy it.

$$\ln Q_i = \phi_0 + \phi_1 \ln L_i + (v_i + u_i)$$

Then:

$$\begin{aligned} \mathrm{E}[x_i(y_i - \mathbf{z}_i)'\boldsymbol{\delta}] &= 0 \\ \mathrm{E}[1(\ln Q_i - [1 \ln L_i])\boldsymbol{\delta}] &= 0 \\ \mathrm{E}[\ln Q_i - \phi_0 - \phi_1 \ln L_i] &= 0 \end{aligned}$$

Consequently:

$$\phi_0 = \ln Q_i - \phi_1 \ln L_1$$
$$\phi_1 = \frac{\phi_0 - \ln Q_i}{\ln L_i}$$

It is easy to see that for one single equation there exist two unkowns. The production equation is underidentified since $K < L$ (one instrumental variable for two regressors). Infinitely many (ϕ_0, ϕ_1) can be proposed to solve said equation.

4. We know proceed to specify $(y_i, \mathbf{z}_i, \mathbf{x}_i)$ and rank condition for some examples that were analyzed before.

· Haavelmo's macroeconomic model:

$$\begin{aligned} C_t &= \alpha_0 + \alpha_1 Y_t + u_t, \quad \alpha_1 \in (0,1) \\ Y_t &= C_t + I_t \end{aligned}$$

Where C_i is aggregate consumption at year t, Y_t denotes Gross National Product (GNP), I_t represents investment and α_1 is the Marginal Propensity to Consume (MPC).

Substituting C_t in the second equation and solving for Y_t yields:

$$Y_t = \frac{\alpha_0}{1-\alpha_1} + \frac{I_t}{1-\alpha_1} + \frac{u_t}{1-\alpha_1}$$

Thus, assuming I_t is exogeneous, it can be shown that Y_t is endogeneous:

$$\operatorname{cov}(Y_t, u_t) = \operatorname{cov}\left(\frac{\alpha_0}{1-\alpha_1} + \frac{I_t}{1-\alpha_1} + \frac{u_t}{1-\alpha_1}, u_t\right) = \frac{\mathrm{V}[u_i]}{1-\alpha_1} \quad (\neq 0)$$

I_t is an instrumental variable to estimate the causal effect of Y_t on C_t:

$$\begin{tikzcd} I_t \arrow[r] & Y_t \arrow[r] \arrow[d] & C_t \\ & u_t \arrow[ur] & \end{tikzcd}$$

$$y_i = C_t,\ \boldsymbol{\delta} = \begin{bmatrix} \delta_0 \\ \delta_1 \end{bmatrix},\ \mathbf{z}_i = \begin{bmatrix} 1 \\ Y_t \end{bmatrix},\ \mathbf{x}_i = \begin{bmatrix} 1 \\ I_t \end{bmatrix} \quad L = 2,\ K = 2$$

· Friedman's Permanent Income Hypothesis (already presented model):

$$\begin{aligned} C_i^* &= kY_i^* \quad k \in (0,1) \\ C_i &= C_i^* + c_i \\ Y_i &= Y_i^* + y_i \end{aligned}$$

The DAG can be drawn as:

$$\begin{tikzcd} 1 \arrow[r] & Y_i^* \arrow[r] \arrow[d] & C_i^* \\ & \xi_i \arrow[ur] & \end{tikzcd}$$

Rank conditions and specifications:

$$y_i = C_i^*,\ \boldsymbol{\delta} = \begin{bmatrix} k \end{bmatrix},\ \mathbf{z}_i = \begin{bmatrix} 1 \\ Y_i^* \end{bmatrix},\ \mathbf{x}_i = \begin{bmatrix} 1 \\ \ \end{bmatrix} \quad L = 1,\ K = 1$$

5. & 6. & 8. We prove that adding any L-dimensional row vector to the rows of a matrix does not alter its rank. This can be shown in two ways. Consider:

$$\mathrm{E}[\mathbf{x}_i \mathbf{z}_i'] = \boldsymbol{\Sigma}_{\mathbf{xz}}, \quad \mathrm{E}[\mathbf{x}_i \mathbf{y}_i] = \boldsymbol{\sigma}_{\mathbf{xy}}$$

We claim:

$$\operatorname{rank}(\boldsymbol{\Sigma}_{\mathbf{xz}}) = \operatorname{rank}\left(\boldsymbol{\Sigma}_{\mathbf{xz}} \ \vdots\ \boldsymbol{\sigma}_{\mathbf{xy}}\right)$$

Let $K = L$, then by Method of Moments (MM) estimation:

$$\begin{aligned} \mathbf{g}(\mathbf{w}_i; \boldsymbol{\delta}) &= \mathbf{x}_i(y_i - \mathbf{z}_i'\boldsymbol{\delta}) \quad (\mathbf{w}_i \coloneqq \{y_i, \mathbf{z}_i, \mathbf{x}_i\}) \\ \mathrm{E}[\mathbf{g}(\mathbf{w}_i; \boldsymbol{\delta})] &= \mathbf{0} \quad \text{(orthogonality condition)} \\ \mathrm{E}[\mathbf{x}_i y_i] &= \mathrm{E}[\mathbf{x}_i \mathbf{z}_i']\boldsymbol{\delta} \\ \boldsymbol{\delta} &= \boldsymbol{\Sigma}_{\mathbf{xz}}^{-1} \boldsymbol{\sigma}_{\mathbf{xy}} \end{aligned}$$

Since $\boldsymbol{\Sigma}_{\mathbf{xz}}$ is of full-column rank and matrix dimensions match. For the IV estimator, $\boldsymbol{\sigma}_{\mathbf{xy}}$ is a linear combination of $\boldsymbol{\Sigma}_{\mathbf{xz}}$, so clearly: $\operatorname{rank}(\boldsymbol{\Sigma}_{\mathbf{xz}}) = \operatorname{rank}\left(\boldsymbol{\Sigma}_{\mathbf{xz}} \ \vdots\ \boldsymbol{\sigma}_{\mathbf{xy}}\right)$.

If the proof was too abstract, let us dumb it down by providing a clearer example in `Python`:

```python
import numpy as np
from numpy.random import normal
from numpy.linalg import matrix_rank

# Create a 3x3 matrix
A = np.array([
    [1, 2, 3],
    [1, 6, 17],
    [1, 8, 23]
])
print(matrix_rank(A)) # rank(A) = 3

# Generate random variable
random = normal(0, 1, 3)

# Add the new variable to matrix A
A = np.vstack([A, random])
print(matrix_rank(A)) # rank(A) = 3
```

Summing up, a full-column matrix A was created as well as a random variable $x \sim \mathcal{N}(0,1)$. Clearly, $\text{cov}(\mathbf{z}_i, x) = 0 \ \forall i \in \{1,2,3\}$. Consequently, an already full-column rank matrix does not have its rank altered by adding uncorrelated vectors to it.

7. In the wage equation, suppose $AGE_i = S_i + EXPR_i \ \ \forall i$. In this case an instrument is formed by linearly combining some exogenous regressors, effectively rendering it redundant. The rank condition is still satisfied since $K = L$, but now $\hat{\boldsymbol{\delta}}$ becomes an IV estimator rather than a GMM estimator, as the equation is exactly identified. For asymptotic normally though, note that $\text{E}[\mathbf{g}_i\mathbf{g}_i']$ should be nonsingular. This is no longer the case as it will be shown.

Since $\text{E}[\mathbf{g}_i\mathbf{g}_i'] = \text{E}[\epsilon_i^2\mathbf{x}_i\mathbf{x}_i']$, it suffices to show that $\boldsymbol{\alpha}'\mathbf{g}_i\mathbf{g}_i' = 0$ for some $\boldsymbol{\alpha} \neq \mathbf{0}$, or even simpler: $\boldsymbol{\alpha}'\mathbf{x}_i\mathbf{x}_i'$.

$$\mathbf{x}_i\mathbf{x}_i' = \begin{bmatrix} 1 \\ S_i \\ EXPR_i \\ S_i + EXPR_i \\ MED_i \end{bmatrix} \begin{bmatrix} 1 & S_i & EXPR_i & S_i + EXPR_i & MED_i \end{bmatrix}$$

The cross-product matrix is expressed as:

$$\begin{bmatrix} 1 & S_i & EXPR_i & S_i + EXPR_i & MED_i \\ S_i & S_i^2 & S_iEXPR_i & S_i(S_i + EXPR_i) & S_iMED_i \\ EXPR_i & EXPR_iS_i & EXPR_i^2 & EXPR_i(S_i + EXPR_i) & EXPR_iMED_i \\ S_i + EXPR_i & S_i(S_i + EXPR_i) & EXPR_i(S_i + EXPR_i) & (S_i + EXPR_i)^2 & MED_i(S_i + EXPR_i) \\ MED_i & MED_iS_i & MED_iEXPR_i & MED_i(S_i + EXPR_i) & MED_i^2 \end{bmatrix}$$

Note that for $\boldsymbol{\alpha} = (0, 1, 1, -1, 0)'$, $\boldsymbol{\alpha}'\mathbf{x}_i\mathbf{x}_i'$ becomes a null vector, since all the $(S_i, EXPR_i)$ terms cancel out with $(S_i + EXPR_i)$, namely, the second and third row are equivalent to the fourth one. Thus, $\boldsymbol{\alpha}'\mathbf{g}_i\mathbf{g}_i = \mathbf{0}$, and so $\mathrm{E}[\mathbf{g}_i\mathbf{g}_i']$ is singular.

9 . What if $\mathbf{x}_i = \mathbf{z}_i$, i.e, all regressors are predetermined?

Let us simply follow the MM procedure estimation:

$$\begin{aligned} \mathbf{x}_i(y_i - \mathbf{z}_i'\boldsymbol{\delta}) &= \mathbf{0} \\ \mathbf{x}_iy_i &= \mathbf{x}_i\mathbf{z}_i'\boldsymbol{\delta} \\ \boldsymbol{\delta} &= (\mathbf{x}_i\mathbf{x}_i')^{-1}(\mathbf{x}_iy_i) \end{aligned}$$

This should already remind you of Chapter 2. In case it does not, let us rewrite the expression as:

$$\begin{aligned} \boldsymbol{\delta} &= (\mathbf{x}_i\mathbf{x}_i')^{-1}(\mathbf{x}_iy_i) \\ \boldsymbol{\beta} &= (\mathbf{x}_i\mathbf{x}_i')^{-1}(\mathbf{x}_iy_i) \\ \boldsymbol{\beta} &= \left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_iy_i\right) = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y}) \end{aligned}$$

So the OLS estimator is a Method of Moments estimator when all regressors are exogeneous.