Lecture 7: Linear Panel Data Model III
Dynamic Panel

February 11, 2025

Production Function Estimation

We focus on the simple 2-factor Cobb-Douglas production function:

$$Y_j = A_j K_j^{\beta_k} L_j^{\beta_l},$$

or, in natural logarithms,

$$y_j = \beta_0 + \beta_k k_j + \beta_l l_j + \varepsilon_j,$$

where

$$\ln A_j = \beta_0 + \varepsilon_j$$

is log TFP. β_0 is a constant, interpretable as the mean of log TFP, while ε_j measures the deviation in productivity from the mean, for firm j. TFP is typically assumed unobserved (at least partially).

Production Function Estimation

Suppose we have micro data on output, capital and labour. How can the parameters of this equation be estimated?

- As you know, for OLS to consistently estimate the β-parameters, the error term must have zero mean and be uncorrelated with the explanatory variables:

$$
\begin{aligned}
E(\varepsilon_j) &= 0, \\
\mathrm{Cov}(k_j, \varepsilon_j) &= 0, \qquad (1) \\
\mathrm{Cov}(l_j, \varepsilon_j) &= 0 \qquad (2)
\end{aligned}
$$

The zero mean assumption is innocuous, as the intercept β_0 would pick up a non-zero mean in ε_j.

Production Function Estimation

- The crucial assumption is zero covariance. Is this likely to hold in the present context?
- No - because it seems quite possible that the firm's capital and labour decisions are influenced by factors that are observed to the firm's manager but unobserved to the econometrician, i.e. by ε_j. This would set up a correlation between the regressors and the residuals, rendering the OLS estimates biased and inconsistent.

Illustration

Assumption 1

1. Firms operate in perfectly competitive input and output markets (so that input and output prices are not affected by the actions of firm j)
2. Capital is a fixed input (decided upon one period in advance, say) rented at rate r
3. Firms observe ε_j before hiring labour (at rate W), and labour is a 'flexible input' that can be altered without dynamic implications.

Illustration

The firm's profit is given by

$$\pi_j = pY_j - WL_j - rK_j$$

$$\pi_j = p\left(A_j K_j^{\beta_k} L_j^{\beta_l}\right) - wL_j - rK_j,$$

where p is the output price. Assuming the firm maximizes profits, it will choose labour such the following first-order condition is fulfilled:

$$\beta_l p A_j K_j^{\beta_k} L_j^{\beta_l - 1} = W,$$

which implies

$$L_j = \left(\frac{\beta_l p A_j}{W}\right)^{\frac{1}{1-\beta_l}} K_j^{\frac{\beta_k}{1-\beta_l}},$$

or, in logs,

$$l_j = \frac{1}{1-\beta_l}\left[\ln \beta_l + \ln p - \ln W + \ln \beta_0 + \varepsilon_j + \beta_k k_j\right].$$

Illustration

- Clearly in this case l_j depends on unobserved TFP (which is the interpretation assigned to the residual ε_j) and so estimating the production function

$$y_i = \beta_0 + \beta_k k_j + \beta_l l_j + \varepsilon_j$$

  by means of OLS will give biased and inconsistent results.

- Note that, since the first-order condition for labour implies a positive correlation between l_j and ε_j, we would expect the OLS estimate of β_l to be upward biased.

Illustration

- **Measurement errors** In general, we expect measurement errors in inputs to lead to downward bias (**attenuation bias**) in the estimated coefficients. Recall the attenuation bias formula:

  $$y_{it} = \beta x^*_{it} + v_{it},$$

  where x^*_{it} is the true but unobserved value of the explanatory variable, and v_{it} is a non-autocorrelated, homoskedastic error term with zero mean. We observe an imperfect measure of x^*_{it}, namely x_{it} such that

  $$x_{it} = x^*_{it} + u_{it},$$

  where u_{it} is a random measurement error uncorrelated with x^*_{it}. Our estimable equation is

  $$y_{it} = \beta x_{it} + (v_{it} - \beta u_{it}),$$

  so the regressor x_{it} is correlated with the error term $(v_{it} - \beta u_{it})$. It can be shown that this will lead to a downward bias in the OLS estimate of β.

Illustration

- That is, estimated $\hat{\beta}$ is **lower** than true β. To give you an idea of what the bias looks like, consider the following formula showing the bias caused by measurement errors:

$$\text{plim}\hat{\beta}^{OLS} = \beta \left(\frac{\sigma^2_{x^*}}{\sigma^2_{x^*} + \sigma^2_u} \right),$$

where $\sigma^2_{x^*}$ is the variance of the true, unobserved explanatory variable, and σ^2_u is the variance of the measurement error. The operator plim can be thought of as showing the value of estimated β in a large sample. Loosely speaking, this is what we can expect to get if there are measurement errors in the explanatory variable. Clearly the higher the variance of the measurement error, the more severe is the bias.

Illustration

- What happens if we take first differences? Clearly,

$$\text{plim}\hat{\beta}^{FD} = \beta \left(\frac{\sigma^2_{dx^*}}{\sigma^2_{dx^*} + \sigma^2_{de}} \right),$$

where d indicates that the variance refers to the differenced variable. Assuming that the variance is constant over time and that the mean of the error is zero, it can be shown that

$$\text{plim}\hat{\beta}^{FD} = \beta \left(\frac{\sigma^2_{x^*}}{\sigma^2_{x^*} + \sigma^2_{e}(1-\rho_e)/(1-\rho_{x^*})} \right),$$

where ρ_e is the serial correlation of the measurement errors and ρ_{x^*} is the serial correlation of the true values of the regressors.

Illustration

$$\text{plim}\hat{\beta}^{OLS} = \beta\left(\frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_e^2}\right)$$
$$\text{plim}\hat{\beta}^{FD} = \beta\left(\frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_e^2(1-\rho_e)/(1-\rho_{x^*})}\right).$$

Which one has the most severe bias?

- The bias of the FD estimator will be more severe than that of the levels estimator if
$$\frac{(1-\rho_e)}{(1-\rho_{x^*})} > 1, \text{i.e., if } \rho_{x^*} > \rho_e.$$
- This is an important result. In most applications we assume that the serial correlation of the measurement errors typically is quite small or zero, while the serial correlation of the true unobserved explanatory variable is positive. In this case, first differencing the data is bound to exacerbate the measurement error bias, and OLS estimation of the levels equation would be preferable to the FD model.
- In practice, estimating the coefficient on the capital stock whilst controlling for fixed effects has proved difficult - see Söderbom and Teal

Potential Solutions

The two traditional solutions to endogeneity problems are **instrumental variables** and **fixed effects**. We are now going to write the production function as

$$y_{jt} = \beta_k k_{jt} + \beta_l l_{jt} + \omega_{jt} + \eta_{jt},$$

where:

- ω_{jt} represents the part of TFP observable to the firm but not to the econometrician - hence this is the source of endogeneity problems. You can think of ω_{jt} as a measure of the managerial quality of the firm. From now on, we will refer to ω_{jt} as ‘unobserved productivity’.
- η_{jt} on the other hand is assumed not to impact on the firm’s input decisions. You can think of η_{jt} as representing measurement errors in output, for example (other interpretations are possible too; see Section 2.2 in ABBP). What’s important is that η_{jt} is not a source of endogeneity bias.

Instrumental Variable

Our problem: We want to estimate

$$y_{jt} = \beta_k k_{jt} + \beta_l l_{jt} + \omega_{jt} + \eta_{jt},$$

but we cannot use OLS, since

$$\text{Cov}(l_{jt}, \omega_{jt}) \neq 0.$$

(It is likely, of course, that capital is endogenous too, but we abstract from that possibility for the moment.)
Suppose an instrument z_{jt} is available, that fulfills the following conditions:

1. The instrument is **valid** (or exogenous):

$$\text{cov}(z_{jt}, \omega_{jt}) = 0.$$

Instrumental Variable

- This is an **exclusion restriction** - z_{jt} is excluded from the structural equation (the production function).
- The instrument is **informative (or relevant)**. This means that the instrument z_{jt} must be correlated with the endogenous regressor (labour in the current example), conditional on all exogenous variables in the model (i.e., capital, if this is thought exogenous). That is, if we assume there is a linear relationship between l_{jt} and z_{jt} and k_{jt},

$$l_{jt} = \delta_0 + \delta_1 k_{jt} + \theta_1 z_{jt} + r_{jt},$$

  where r_{jt} is mean zero and uncorrelated with the variables on the right-hand side, we require $\theta_1 \neq 0$.

Many economists take the view that, for instrumental variable estimation to be convincing, the instruments used must be motivated by theory.

Instrument

Given the first order condition of the labour derived in the previous page:

$$l_{jt} = \frac{1}{1-\beta_l}\left[\ln \beta_l + \ln p - \ln W + \beta_k k_{jt} + \omega_{jt}\right].$$

- This suggests the wage rate W might be a useful instrument:
  - Our theory says it is (negatively) correlated with labor.
  - The wage rate also must be uncorrelated with ω_{jt}. This may not be an entirely innocuous assumption to make. While the wage rate does not directly enter the production function, wages might be correlated with unobserved productivity for other reasons - e.g., if more productive firms have stronger market power in input markets - in which case the wage will not be a valid instrument.

Instrument

- It also follows from the first-order condition above that the output price is a potential instrument. However, it has been used less often in the literature. Why might we be concerned about using the output price as an instrument?
- A similar way of reasoning can be applied to capital, if it is considered endogenous (i.e., using the cost of capital as an instrument).

Instrument

Five reasons why the IV approach based on prices as instruments has not been very successful.

- **Market power.** Wages and capital prices (and output prices) could well be correlated with unobserved productivity if input (output) markets are not perfectly competitive: e.g., high unobserved productivity gives the firm market power and so enables it to influence the price.
- **Wages and unobserved worker quality.** When labour costs are reported in firm-level datasets, they typically come in the form of average wage per worker, and you may well be concerned that the average wage in the firm is correlated with unobserved quality of the workforce. Since the unobserved quality of the workforce likely impacts on unobserved productivity, this would imply the average wage is an invalid instrument.

Instrument

- **Law of one price.** If, as is typically the case, one wants to include time dummies in the production function, there must be variation in input prices across firms at a given point in time for these to be useful instruments. If input markets are essentially national in scope, this seems unlikely. Average wages varying across firms in most datasets suggest this is at least partly picking up unobserved worker quality.
- **Endogenous unobserved productivity.** Suppose unobserved productivity actually depends on input choices - e.g., investment in modern technology raises productivity. In that case, it will be hard to argue that input prices are valid instruments, since these surely will impact on investment.
- **Attrition.** A different kind of endogeneity problem sometimes discussed in the literature is posed by endogenous attrition, i.e., that the firm's exit decision depends on unobserved productivity as well as input prices (after all, these jointly determine the profitability of the firm). In such a case, we will have a Heckman type selection problem, where all variables determining the exit decision will go into the residual of the production function in the selected sample. Clearly, input prices cannot be used as instruments in this case.

Fixed Effects

A second traditional solution to the endogeneity problem is fixed effects estimation, which as you know requires panel data. One key assumption underlying this approach is that unobserved productivity is constant over time,

$$\omega_{jt} = \omega_j$$

but varies across firms. We would now write the production function as

$$y_{jt} = \beta_k k_{jt} + \beta_l l_{jt} + (\omega_j + \eta_{jt})$$

and use perhaps the within estimator ('fixed effects' estimator) or the first-differenced estimator to estimate the parameters - in the latter case for example we would thus estimate

$$y_{jt} - y_{j,t-1} = \beta_k(k_{jt} - k_{j,t-1}) + \beta_l(l_{jt} - l_{j,t-1}) + (\eta_{jt} - \eta_{j,t-1}),$$

using OLS (probably with firm-clustered standard errors since the differenced residual is likely serially correlated).

Notice that the source of endogeneity bias has been eliminated, thus effectively solving the endogeneity problem (subject of course to strict exogeneity; see earlier lectures in this course).

Fixed Effects

Three reasons why the fixed effects approach has not been very successful:

1. **Time invariant unobserved productivity.** The assumption that unobserved productivity is fixed over time is quite restrictive, especially in longer panels.
2. **Differencing may exacerbate measurement error bias.** When there are measurement errors in inputs, the fixed effects estimator may well be more severely biased than the OLS estimator. Discuss.
3. **Poor performance in practice.** Fixed effects estimates of the capital coefficient are often implausibly low, and estimated returns to scale is often (severely) decreasing ($\beta_k + \beta_l << 1$).

Dynamic Panel Approach (Bond 2002)

$$y_{it} = \alpha y_{i,t-1} + (\eta_i + v_{it}); \quad |\alpha| < 1; \quad i = 1, 2, ..., N; \quad t = 2, 3, ..., T$$

We treat the individual effects (η_i) as being stochastic, which here implies that they are necessarily correlated with the lagged dependent variable $y_{i,t-1}$ unless the distribution of the η_i is degenerate. Initially we further assume that the disturbances (v_{it}) are serially uncorrelated. These jointly imply that the Ordinary Least Squares (OLS) estimator of α in the levels equations (2.1) is inconsistent.

Fixed Effect (Within) Estimator

The transformed lagged dependent variable is:

$$y_{i,t-1} - \frac{1}{T-1}(y_{i1} + \ldots + y_{it} + \ldots + y_{iT-1})$$

The transformed error term is:

$$v_{it} - \frac{1}{T-1}(v_{i2} + \ldots + v_{i,t-1} + \ldots + v_{iT})$$

- The correlation between transformed lagged dependent variable and the transformed error term can be shown to be negative.
- Within estimator is biased downwards.
- Key – *strict exogeneity* assumption is violated.

First Difference

The dynamic panel data model can be written as:

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta v_{it}, \quad |\alpha| < 1; \quad i = 1, 2, \ldots, N; \quad t = 3, 4, \ldots, T \qquad (1)$$

- The dependence of Δv_{it} on $v_{i,t-1}$ implies that OLS estimates of α in the first-differenced model are inconsistent, with the direction of the inconsistency being downward and typically greater than that found for the Within Groups estimator.
- Consistent estimates of α can now be obtained using 2SLS with instrumental variables that are both correlated with $\Delta y_{i,t-1}$ and orthogonal to Δv_{it}.
- The only assumption required on the initial conditions y_{i1} is uncorrelated with the subsequent disturbances v_{it} for $t = 2, 3, \ldots, T$.
- Disturbances v_{it} are serially uncorrelated.

GMM Estimator

The instrument matrix Z_i is defined as:

$$Z_i = \begin{bmatrix} y_{i1} & 0 & \cdots & 0 & \\ 0 & y_{i1} & y_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & y_{i1} & y_{i,T-2} \end{bmatrix},$$

where rows correspond to the first-differenced equations for periods $t = 3, 4, ...T$, for individual i, and exploit the moment conditions

$$E[Z_i' \Delta v_i] = 0 \quad \text{for} \quad i = 1, 2, \ldots, N$$

where $\Delta v_i = (\Delta v_{i3}, \Delta v_{i4}, \ldots, \Delta v_{iT})'$.

GMM Estimator

In general, the asymptotically efficient GMM estimator based on this set of moment conditions minimises the criterion

$$J_N = \left(\frac{1}{N} \sum_{i=1}^{N} \Delta v_i' Z_i \right) W_N \left(\frac{1}{N} \sum_{i=1}^{N} Z_i' \Delta v_i \right),$$

using the weight matrix

$$W_N = \left[\frac{1}{N} \sum_{i=1}^{N} \left(Z_i' \hat{\Delta} v_i \hat{\Delta} v_i' Z_i \right) \right]^{-1},$$

where the $\hat{\Delta} v_i$ are consistent estimates of the first-differenced residuals obtained from a preliminary consistent estimator. Hence this is known as a two-step GMM.

Example: Corporate Investment Rate

Variation across firms in depreciation rates provides one motivation for suspecting the presence of individual-specific effects in this context. The sample contains 703 publicly traded UK firms for which we have consecutive annual data from published company accounts for a minimum of 4 years between 1987 and 2000. Further details of the sample and the data are provided in Bond et al. (2002).

$$\left(\frac{I_{it}}{K_{it}}\right) = c_t + \alpha\left(\frac{I_{i,t-1}}{K_{i,t-1}}\right) + (\eta_i + v_{it}).$$

Example: Corporate Investment Rate

Table 1. Alternative Estimates of the $AR(1)$ Specification for Company Investment Rates

Dependent variable: $(I/K)_t$

	OLS LEVELS	WITHIN GROUPS	2SLS DIF	GMM DIF	GMM DIF
$(I/K)_{t-1}$	0.2669 (.0185)	-0.0094 (.0181)	0.1626 (.0362)	0.1593 (.0327)	0.1560 (.0318)
m1 m2 Sargan	-4.71 2.52	-11.36 -2.02	-10.56 0.61	-10.91 0.52 .36	-11.12 0.46 .43
Instruments			$(I/K)_{t-2}$	$(I/K)_{t-2}$ $(I/K)_{t-3}$	$(I/K)_{t-2}$ $(I/K)_{t-3}$ $\vdots$ $(I/K)_1$

Sample: 703 firms; 4966 observations; 1988-2000

Persistent Series

The GMM estimators for autoregressive models outlined in the previous section extend in a natural way to autoregressive-distributed lag models of the form

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + (\eta_i + v_{it}); \quad i = 1, 2, \ldots, N; \quad t = 2, 3, \ldots, T$$

where x_{it} can be a vector of current and lagged values of additional explanatory variables. An attractive feature of this approach is that it does not require models for the x_{it} series to be specified in order to estimate the parameters (α, β).

System GMM

- Instruments based on x

$$(y_{i1}, \ldots, y_{i,t-2}, x_{i1}, \ldots, x_{i,t-2})$$

- Instrument s based on contemporaneous exogeneity

$$(y_{i1}, \ldots, y_{i,t-2}, x_{i1}, \ldots, x_{i,t-2}, x_{i,t-1})$$

- If Δx is uncorrelated with individual permanent effect
  - An intermediate case occurs when we are not willing to assume that the level of the x_{it} variable is uncorrelated with the individual effects, but we are willing to assume that the first-differences Δx_{it} are uncorrelated with η_i. In this case suitably lagged values of Δx_{it} can be used as instrumental variables in the levels equation for period t; see Arellano and Bover (1995) for details.

System GMM

This last observation raises the possibility that lagged differences $\Delta y_{i,t-1}$ may also be valid instruments for the levels equations in autoregressive models. This turns out to depend on the validity of a stationarity assumption about the initial conditions y_{i1}, which is discussed in detail in Blundell and Bond (1998). Specifically it requires $E\left[\left(y_{i1} - \left(\frac{\eta_i}{1-\alpha}\right)\right)\eta_i\right] = 0$ for $i = 1, \ldots, N$, so that the initial conditions do not deviate systematically from the value $\left(\frac{\eta_i}{1-\alpha}\right)$ which the model specifies that the y_{it} series for individual i converges towards from period 2 onwards.

Revisit Production Function (Blundel et al. 2000)

Blundell and Bond (2000) consider the Cobb-Douglas production function

$$\begin{aligned} y_{it} &= \beta_n n_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + v_{it} + m_{it}) \\ v_{it} &= \alpha v_{i,t-1} + e_{it}, \quad |\alpha| < 1 \\ e_{it}, m_{it} &\sim MA(0) \end{aligned}$$

Interest is in the consistent estimation of the parameters $(\beta_n, \beta_k, \alpha)$ when the number of firms (N) is large and the number of years (T) is fixed. It is assumed that both employment (n_{it}) and capital (k_{it}) are potentially correlated with the firm-specific effects (η_i), and with both productivity shocks (e_{it}) and measurement errors (m_{it}).

Revisit Production Function (Blundel et al. 2000)

In this case there are no valid moment conditions for the static specification in the previous page if the disturbances v_{it} are indeed autoregressive ($\alpha \neq 0$). However, this model has a dynamic (common factor) representation

$$
\begin{aligned}
y_{it} = \beta_n n_{it} - \alpha\beta_n n_{i,t-1} + \beta_k k_{it} - \alpha\beta_k k_{i,t-1} + \alpha y_{i,t-1} \\
+(\gamma_t - \alpha\gamma_{t-1}) + (\eta_i(1-\alpha) + e_{it} + m_{it} - \alpha m_{i,t-1})
\end{aligned}
$$

or

$$
y_{it} = \pi_1 n_{it} + \pi_2 n_{i,t-1} + \pi_3 k_{it} + \pi_4 k_{i,t-1} + \pi_5 y_{i,t-1} + \gamma_t^* + (\eta_i^* + w_{it})
$$

subject to two non-linear (common factor) restrictions $\pi_2 = -\pi_1\pi_5$ and $\pi_4 = -\pi_3\pi_5$. Whilst the error term $(v_{it} + m_{it})$ is serially correlated at all lag lengths, the error term w_{it} is serially uncorrelated if there are no measurement errors ($w_{it} = e_{it}$ with $\text{var}(m_{it}) = 0$), or $w_{it} \sim MA(1)$ if there are transient measurement errors in some of the series.

Revisit Production Function (Blundel et al. 2000)

Table 3. AR(1) Specifications for the Production Series

Labour (n_t)	OLS LEVELS	WITHIN GROUPS	GMM-DIF $t-3$	GMM-SYS $t-3$	GMM-SYS $t-4$
n_{t-1}	0.986 (.002)	0.723 (.022)	0.920 (.062)	0.923 (.033)	
m1 m2 Sargan Dif-Sar	4.16 2.67	-8.51 0.60	-7.62 0.44 .040	-8.99 0.43 .056 .387	
Capital (k_t)					
k_{t-1}	0.987 (.002)	0.733 (.027)	0.768 (.070)	0.925 (.021)	
m1 m2 Sargan Dif-Sar	7.72 2.29	-6.82 -1.73	-5.80 -1.73 .563	-6.51 -1.81 .627 .562	
Sales (y_t)					
y_{t-1}	0.988 (.002)	0.693 (.025)	0.775 (.063)	0.963 (.048)	0.893 (.063)
m1 m2 Sargan Dif-Sar	5.70 0.97	-7.35 -2.37	-5.95 -2.46 .040	-7.15 -2.53 .025 .134	-6.35 -2.63 .092

Revisit Production Function (Blundel et al. 2000)

Table 4. Production Function Estimates

	OLS LEVELS	WITHIN GROUPS	GMM DIF $t-2$	GMM DIF $t-3$	GMM SYS $t-2$	GMM SYS $t-3$
n_t	0.479	0.488	0.513	0.499	0.629	0.472
	(0.029)	(0.030)	(0.089)	(0.101)	(0.106)	(0.112)
n_{t-1}	-0.423	-0.023	0.073	-0.147	-0.092	-0.278
	(0.031)	(0.034)	(0.093)	(0.113)	(0.108)	(0.120)
k_t	0.235	0.177	0.132	0.194	0.361	0.398
	(0.035)	(0.034)	(0.118)	(0.154)	(0.129)	(0.152)
k_{t-1}	-0.212	-0.131	-0.207	-0.105	-0.326	-0.209
	(0.035)	(0.025)	(0.095)	(0.110)	(0.104)	(0.119)
y_{t-1}	0.922	0.404	0.326	0.426	0.462	0.602
	(0.011)	(0.029)	(0.052)	(0.079)	(0.051)	(0.098)
m1	-2.60	-8.89	-6.21	-4.84	-8.14	-6.53
m2	-2.06	-1.09	-1.36	-0.69	-0.59	-0.35
Sargan	-	-	.001	.073	.000	.032
Dif-Sar	-	-	-	-	.001	.102
β_n	0.538	0.488	0.583	0.515	0.773	0.479
	(0.025)	(0.030)	(0.085)	(0.099)	(0.093)	(0.098)
β_k	0.266	0.199	0.062	0.225	0.231	0.492
	(0.032)	(0.033)	(0.079)	(0.126)	(0.075)	(0.074)
α	0.964	0.512	0.377	0.448	0.509	0.565
	(0.006)	(0.022)	(0.049)	(0.073)	(0.048)	(0.078)
Comfac	.000	.000	.014	.711	.012	.772
CRS	.000	.000	.000	.006	.922	.641