Project 1: Linear Panel Data and Production Technology

By Astrid Sophie Fugleholm, Jacob Strabo and Jørgen Høst

Characters: Approx 12,000 (including spaces and math).

1 Introduction

The production of goods and services is the central element in any economy. Workers supply labor and capital as inputs and gain utility from consuming outputs. Understanding the relationship between in- and outputs is crucial to profit maximizing firms. The decision to scale production up or down relies on the necessary change in inputs - is this more, less, or exactly proportional to the change output?

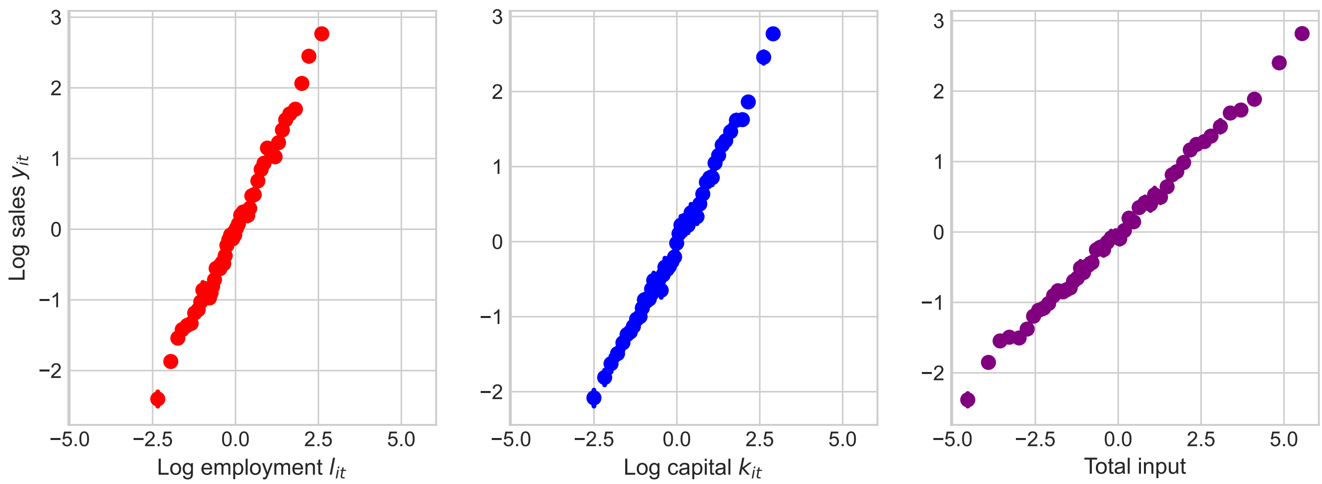
We attempt to answer this by investigating French manufacturing firms using the fixed effects (FE), first differences (FD), and Pooled Ordinary Least Squares (POLS) estimators. The results suggest there is not constant but rather decreasing returns to scale. This is in line with Szpiro and Cette (1994) who also reject the notion of constant returns to scale in the French manufacturing industry.

Crucially, our estimators turn out to be ill-suited for inference. POLS is poor as consistency requires that firm-specific fixed effects and regressors are uncorrelated. Althoug accounting for this, FD and FE require strict exogeneity, which does not hold in our setting. We propose future work alleviate this issue by using the Arellano-Bond estimator instead.

2 Data

We use panel data for $N = 441$ French manufacturing firms observed for $T = 12$ years during 1967-1978. The data contains observations on firms' deflated sales, number of employees, and adjusted capital stock, all in logs and all time-demeaned implying that any effect specific to calendar time are discarded. *Figure 1* shows a particularly striking linear relationship between growth in production in- and output. A slope less than one suggests decreasing rather than constant returns to scale.

Figure 1: *The link between production in- and output*



Note: Bin size = 50 observations. Total input is the sum of log capital and log employment.

3 Cobb-Douglas Producton Function

We assume firms produce according to a Cobb-Douglas production function

$$Y = F(K, L) = AK^{\beta_K} L^{\beta_L}, \tag{1}$$

where Y is output generated by total factor productivity (TFP) $A > 0$, capital K, and labor L. Output elasticities of capital and labor are denoted β_K and β_L, resp. Capital and labor are observed, while TFP is unobserved and composed by time-varying and time-invariant factors. Taking logs and expressing this as a unobserved linear panel data model for firm i at time t yields

$$y_{it} = \beta_K k_{it} + \beta_L l_{it} + \eta_{it}, \tag{2}$$

where y_{it}, k_{it}, l_{it}, and ν_{it} are the log of deflated sales, adjusted capital stock, employment, and TFP, resp. The composite error term $\eta_{it} \equiv c_i + u_{it}$ consists of a time-invariant component c_i and an idiosyncratic u_{it} component. The location of a factory and labor strikes are examples of time-invariant and -varying unobservable factors, resp.

Our parameters of interest are β_K and β_L, capturing the returns to scale, with $\beta_K + \beta_L > 1$ implying increasing, $\beta_K + \beta_L < 1$ implying decreasing, and $\beta_K + \beta_L = 1$ implying constant returns to scale. Testing the assumption of constant returns to scale thus amounts to testing the null and alternative hypotheses

$$\mathcal{H}_0 : \beta_K + \beta_L = 1 \quad \text{and} \quad \mathcal{H}_A : \beta_K + \beta_L \neq 1. \tag{3}$$

Throughout the paper, we employ the following notation: $K = 2$ is the number of regressors in our model; $\mathbf{x}_{it} = \{k_{it}, l_{it}\}$ is a (1×2) vector of regressors; and $\beta = \{\beta_K, \beta_L\}$ is a (1×2) vector of parameters.

4 Econometric Model

4.1 The Pooled Ordinary Least Squares (POLS) Estimator

A natural starting point is to estimate (3) using POLS.[1] POLS is consistent, i.e. converges in probability to the true parameter (Angrist and Pischke (2009)), if OLS.1-OLS.2 hold, cf. Wooldridge (2010b)

OLS.1: $\mathbb{E}[\mathbf{x}_{it} u_{it}] = 0 \quad \forall i \in N, \quad \forall t \in T$, i.e. regressors are uncorrelated with the idiosyncratic error.

OLS.2: Rank $\left[\mathbb{E}[\mathbf{X}_i' \mathbf{X}_i]\right] = K \quad \forall i \in N$, i.e. the matrix of regressors has full rank.

[1] As the time-specific means have all ready been subtracted in our data, our POLS is in fact a time fixed effects model.

As firm-specific fixed effects are unobserved, c_i is effectively included in a composite error term $\eta_{it} = c_i + u_{it}$. OLS.1 thus implies that inputs are uncorrelated with both the idiosyncratic error and firm-specific fixed effects: $\mathbb{E}\left[u_{it}|\mathbf{x}_{it}, c_i\right] = 0$. Since more productive firms likely employ different levels of capital and labor than less productive firms, OLS.1 is unlikely to hold, and POLS is inconsistent. The random effects (RE) estimator also relies on firm-specific fixed effects being orthogonal to the regressors, it is also assumed to be inconsistent. We therefore turn to FD and FE with different approaches to removing firm-specific fixed effects.

4.2 The Fixed Effects (FE) Estimator

The FE estimator removes firm-specific fixed effects by subtracting from each variable its mean:

$$y_{it} - \bar{y}_i = \beta_K(k_{it} - \bar{k}_i) + \beta_L(l_{it} - \bar{l}_i) + c_i - \bar{c}_i + u_{it} - \bar{u}_i \Leftrightarrow$$

$$\ddot{y}_i t = \beta_K \ddot{k}_{it} + \beta_L \ddot{l}_{it} + \ddot{u}_{it}, \quad t = 1, ..., T, \tag{4}$$

where c_i vanishes as it is constant over time. FE is consistent under FE.1 and FE.2, cf. Wooldridge (2010a):

> **FE.1** $\mathbb{E}[\ddot{\mathbf{x}}_{it}\ddot{u}_{is}] = 0 \quad \forall i \in N$ and $s, t \in T$, i.e. idiosyncratic errors are uncorrelated with regressors in period t and all other periods (strictly exogeneous).
>
> **FE.2**: Rank $\left[\mathbb{E}\left[\ddot{\mathbf{X}}_i{}'\ddot{\mathbf{X}}_i\right]\right] \quad \forall i \in N$, i.e. the matrix of regressors $\ddot{\mathbf{X}}_i$ has full rank ruling out time-invariant regressors.

FE is asymptotically efficient under FE.3, cf. Wooldridge (2010a):

> **FE.3** $\mathbb{E}\left[\mathbf{u}_i\mathbf{u}_i'|x_{it}c_i\right] = \sigma_u^2\mathbf{I}_T \quad \forall i \in N$, i.e. the time-demeaned idiosyncratic errors are homoskedastic with constant variance.

In our setting, FE is consistent if capital and labor are unaffected by omitted variables observed in any other time period, and capital and labor are time-variant. FE is efficient if time-demeaned capital and labor explain (most of) the variation in time-demeaned output across firms. Following derivations in Wooldridge (2010a), we define the estimator as

$$\widehat{\beta}_{FE} = (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\ddot{\mathbf{X}}'\ddot{\mathbf{y}} \tag{5}$$

and its variance-covariance estimator under FE.3 as

$$\widehat{Avar}(\widehat{\beta}_{FD}) = \hat{\sigma}_u^2 \left(\ddot{\mathbf{X}}'\ddot{\mathbf{X}}\right)^{-1}, \tag{6}$$

where $\hat{\sigma}_u^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{N}\hat{\ddot{u}}_{it}^2}{NT-N-K}$, and $\ddot{\mathbf{X}}$ and $\ddot{\mathbf{y}}$ are resp. a $441 * 12 \times 2$ matrix and $441 * 12 \times 1$ vector of stacked $\ddot{\mathbf{x}}_{it}$ and $\ddot{\mathbf{y}}_{it}$ from (4).

4.3 The First Difference (FD) Estimator

The FD estimator removes firm-specific fixed effects by subtracting the value from the previous period

$$y_{it} - y_{i,t-1} = \beta_K(k_{it} - k_{i,t-1}) + \beta_L(l_{it} - l_{i,t-1}) + c_i - c_i + u_{it} - u_{i,t-1} \Leftrightarrow$$

$$\Delta y_{it} = \beta_K \Delta k_{it} + \beta_L \Delta l_{it} + \Delta u_{it}, \quad t = 2, ..., T, \tag{7}$$

where c_i vanishes as it is constant over time. FD is consistent under FD.1 and FD.2, cf. Wooldridge (2010a):

FD.1: $\mathbb{E}[\Delta x_{it} \Delta u_{is}] = 0, \quad \forall i \in N \quad \text{and} \quad s, t \in T$, i.e. strict exogeneity as in FE.1. It is sufficient to assume that errors are uncorrelated with regressors in period $t-1$, t, and $t+1$ $\mathbb{E}[u_{it}|x_{i,t-1}, x_{it}, x_{i,t+1}] = 0$.

FD.2: Rank $\left[E[\Delta \mathbf{X}_i' \Delta \mathbf{X}_i]\right] = K \quad \forall i \in N$, i.e. the matrix of regressors $\Delta \mathbf{X}_i$ has full rank ruling out time-invariant regressors.

FE is asymptotically efficient under FD.3, cf. Wooldridge (2010a):

FD.3: $\mathbb{E}[\mathbf{e}_i \mathbf{e}_i' | x_{it} c_i] = \sigma_e^2 I_{T-1} \quad \forall i \in N$, where $\mathbf{e}_i = \Delta \mathbf{u}_{it}$, i.e. the differenced idiosyncratic errors are homoskedastic and thus have constant variance.

In our setting, FD is consistent if inputs are unaffected by omitted variables observed in the previous, current, and future time period, and capital and labor are time-variant. FD is efficient if first-differenced capital and labor explain (most of) the variation in first-differenced output across firms. Following derivations in Wooldridge (2010a), we define the estimator as

$$\widehat{\beta}_{FD} = \left(\Delta \mathbf{X}' \Delta \mathbf{X}\right)^{-1} \Delta \mathbf{X}' \Delta y \tag{8}$$

and its variance-covariance under FD.3 as

$$\widehat{Avar}(\widehat{\beta}_{FD}) = \hat{\sigma}_u^2 \left(\sum_{i=1}^{N} \Delta \mathbf{X}_i' \Delta \mathbf{X}\right)^{-1}, \tag{9}$$

where $\hat{\sigma}_u^2 = \frac{1}{NT-N-K}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\Delta u}_{it}^2$, and $\Delta \mathbf{X}$ and $\Delta \mathbf{y}$ are resp. a $441 * 11 \times 2$ matrix and $441 * 11 \times 1$ vector of stacked $\Delta \mathbf{x}_{it}$ and $\Delta \mathbf{y}_{it}$ from (7).

5 Results

We proceed by estimating the three above mentioned panel data models in a static setting. The results are reported in *table 1*. In any case, capital and labor have positive effects on output. A 1% increase in capital and labor leads to a 0.70-0.55% and 0.30-0.06% increase in output, resp.

Table 1: *Effects of Employment and Capital on Firms' Production*

	POLS	FE	FD
Constant	0.0	.	.
	(0.005)	.	.
Log employment, β_L	0.6778***	0.6964***	0.5462***
	(0.0101)	(0.0144)	(0.0177)
Log capital stock, β_K	0.3028***	0.1427***	0.0645***
	(0.009)	(0.0124)	(0.0182)
N	441	441	441
T	12	12	11

Note: Standard errors in parentheses, $p < 0.05$*, $p < 0.01$**, $p < 0.001$***.

Comparing POLS with FE and FD estimates, the effect of capital on output is remarkably lower. With the two latter accounting for unobserved time-invariant heterogeneity, this is unsurprising; firms with an advantageous geographic location can produce more output and devote more resources to investments in capital stock creating upward bias in β_K. Hence, OLS is deemed inconistent. We thus turn to the FE and FD estimators and evaluate the validity of assumptions required for them to be consistent and efficient.

5.1 First Differences versus Fixed Effects

5.1.1 Efficiency

The FD and FE estimators both eliminate omitted variable bias by accounting for firm-specific fixed effects. When $T = 2$, they are numerically identical. When $T > 2$, the choice between them hinges on the level of serial correlation in the idiosyncratic errors. FE is more efficient if errors are serially uncorrelated. FD is more efficient if the error term follows a random walk; the errors themselves are correlated over time, but the differenced errors are not. When errors are serially correlated, first-differencing leaves only white noise, whereas FE simply demeans errors. Hence, to inform our choice between FE and FD, we perform a test for serial correlation by running the regression

$$\widehat{\epsilon}_{it} = \rho\widehat{\epsilon}_{i,t-1} + u_{it}, \tag{10}$$

where $\widehat{\epsilon}_{it} \equiv \Delta u_{it}$ comes from the auxiliary FD-AR(1) regression $\Delta y_{it} = \Delta y_{it-1} + \Delta u_{it}$. The results are reported in *table 2*.

Table 2: *Test for serial correlation*

	ϵ_{it}
$\epsilon_{i,t-1}$, ρ	-0.2141***
	(0.0145)
N	441
T	10

Note: (S.E.), $p < 0.05$*, < 0.01**, < 0.001***.

We formulate a null and alternative hypothesis for the presence of serial correlation following Wooldridge (2016): $\mathcal{H}_0 : \rho = -0.5$ and $\mathcal{H}_A : \rho \neq -0.5$. The t-statistic is significantly greater than the critical value on a 1% level, so we reject $\mathcal{H}_0$ and conclude that the errors are serially correlated. This suggests that in our setting, FD is more efficient than and thus preferred to FE.

5.1.2 Consistency

To examine consistency of FD and FE, we conduct a test for strict exogeneity. This is done by including a lead of employment in (4).

Table 3: *Test for strict exogeneity*

	FE
Log employment, β_L	0.556***
	(0.0226)
Log capital stock, β_K	0.1323***
	(0.0129)
Log employment +1, β_L for $t+1$	0.1688***
	(0.0218)
N	441
T	12

Note: Standard errors in parentheses, $p < 0.05$*, $p < 0.01$**, $p < 0.001$***.

In our setting, FE.1/FD.1 imply that output at time t is uncorrelated with employment in any other period, controlling for firm-specific fixed effects. Result from *table 3* show that this is unfortunately the case. As such, we are left with inconsistent estimates from FD and FE in *table 1*.

5.2 Testing Hypothesis of Constant Returns to Scale

We test the phenomenon of constant returns to scale despite our inconsistent estimates. We define the Wald statistics of the linear hypothesis $\mathbf{R}\beta = \mathbf{r}$

$$(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{R}\boldsymbol{\beta})' \left[\mathbf{R}(\mathbf{V}/N)\mathbf{R}'\right]^{-1} (\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{R}\boldsymbol{\beta}) \overset{d}{\approx} \chi^2_Q$$

With $\mathcal{H}_0 : \beta_K + \beta_L = 1$, $\mathbf{R} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ with $\mathbf{r} = \mathbf{1}$. For FD and FE, the test statistics are $\chi^2(1) = 122,656$ and $\chi^2(1) = 74,256$, resp., with $\mathbf{Q} = rank(\mathbf{R}) = 1$ degree of freedom. They are above the critical value on a 1 % significance level with p-values equal to zero. We therefore reject the null of constant of returns to scale. However, as established before, we emphasize the questionable validity of this test as the strict exogeneity assumption is violated.

6 Discussion

To account for the problem of inconsistent FD and FE estimates, we turn to dynamic panel methods, e.g., a pooled instrumental variable estimation which is consistent under the assumption of sequential exogeneity. That is, *previous* realisations of inputs are uncorrelated with *current* output.

Assuming a panel data model on the form $y_{it} = \rho y_{it-1} + \mathbf{x}_{it}\beta + c_i + u_{it}$. By taking the first difference, $\Delta y_{tt} = \rho\Delta y_{tt-1} + \Delta\mathbf{x}'_{tt}\beta + \Delta u_{tt}$. The idea is using output from two years prior as instruments ($z_{it} = y_{t-2}$) for the difference in previous output (Δy_{t-1}) as the instruments likely satisfy requirements of being exogenous ($E\left[\mathbf{z}_{it}u_{it}\right] = E\left[y_{t-2}\Delta u_{it}\right] = E\left[y_{t-2}(u_t - u_{it-1})\right] = 0$) and relevant ($E\left[y_{t-2}\Delta y_{t-1}\right] = E\left[y_{t-2}(y_{t-1} - y_{t-2})\right] \neq 0$). Extending this idea and defining a set of moment conditions from instrumental variables can be used in a GMM-estimator such as Arellano-Bond. This, however, goes beyond the scope of this paper.

7 Conclusion

In this paper, we examine the relationship between production in- and outputs. More specifically, we investigate whether French manufacturing firms exhibit increasing, decreasing, or constant returns to scale in production. The result suggest not constant but rather decreasing returns to scale. This is in line with the conclusion in Szpiro and Cette (1994). Unobserved time-invariant heterogeneity, such as geographic advantages specific to certain firms, renders POLS a poor estimator. While FD and FE account for this, the assumption of strict exogeneity nedded for consistency is violated. Hence, we propose using the Arellano-Bond estimator instead in future work.

References

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Szpiro, D. and Cette, G. (1994). Returns to scale in the french manufacturing industry. *European Economic Review*, 38(7):1493–1504.

Wooldridge, J. M. (2010a). *Basic Linear Unobserved Effects Panel Data Methods*, chapter 10, pages 281–334. The MIT Press.

Wooldridge, J. M. (2010b). *Single-Equation Linear Model and Ordinary Least Squares Estimation*, chapter 4, pages 51–88. The MIT Press.

Wooldridge, J. M. (2016). *Multiple Regression Analysis: OLS Asymptotics.* Cengage Learning.