Assignment 1

Mathias Porsgaard, Mikkel Reich, Noah Carelse

Faculty of Social Sciences
University of Copenhagen

1 Cobb-Douglas production with CRS

In this assignment we want to test whether the Cobb-Douglas production function (CD) exhibits constant returns to scale (CRS) at a firm level using real data and panel data methods. The Cobb-Douglas production function is given by,

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

where K is capital, L is labor, A is total factor productivity (TFP) and β_K, β_L are the output elasticities of capital and labor, respectively. It follows that for CD to have CRS, it must be homogeneous of degree one, i.e. $\beta_K + \beta_L = 1$, as seen from,

$$F(\lambda K, \lambda L) = A(\lambda K)^{\beta_K} (\lambda L)^{\beta_L} = A\lambda^{\beta_K+\beta_L} K^{\beta_K} L^{\beta_L} = \lambda F(K, L) \iff \beta_K + \beta_L = 1.$$

which forms the linear hypothesis we want to test i.e. $H_0 : \beta_K + \beta_L = 1$ vs. $H_A : \beta_K + \beta_L \neq 1$. While K and L are certainly important production inputs, we cannot hope to observe **all** input factors that make up production. To investigate the functional form of the production function, we estimate a simple (two-way) fixed effects (FE) model, to remove the time-invaring and time-varying unobserved heterogeneity. Note, that these factors with control for much more than just TFP, in fact all other unobserved factors that affect production.

We start by taking logs of the production function to linearize it,

$$\begin{aligned} \ln Y_{it} &= \ln A_{it} + \beta_K \ln K_{it} + \beta_L \ln L_{it} \\ y_{it} &= \beta_K k_{it} + \beta_L l_{it} + u_{it}, \quad u_{it} = \ln A_{it} \end{aligned}$$

Suppose that $\ln A_{it}$ can be decomposed into time-invariant firm-specific effects, α_i, a time-varying effects, γ_t, and add an idiosyncratic error, ε_{it}, i.e. $u_{it} = \alpha_i + \gamma_t + \varepsilon_{it}$. The FE model is then given by,

$$\begin{aligned} u_{it} &= \alpha_i + \gamma_t + \varepsilon_{it}, \\ y_{it} &= \alpha_i + \gamma_t + \beta_K k_{it} + \beta_L l_{it} + \varepsilon_{it}, \end{aligned}$$

where $y_{it} = \ln Y_{it}$, $k_{it} = \ln K_{it}$ and $l_{it} = \ln L_{it}$. Collect the RHS variables in the vector $\boldsymbol{x}_{it} = (k_{it}, l_{it})$ and the parameters in the vector $\boldsymbol{\beta} = (\beta_K, \beta_L)$. The FE model can then be written on compact form as,

$$y_{it} = \alpha_i + \gamma_t + \boldsymbol{x}_{it}\boldsymbol{\beta} + \varepsilon_{it}.$$

Due to α_i and γ_t being unobserved, we cannot estimate them directly, and if we suspect that they are correlated with $\boldsymbol{x}_{it}$, we need to control for them to obtain consistent estimates of the parameters. This can be done by transforming the data to remove these effects, and then estimate the parameters using OLS on the transformed data.

The transformation used to remove α_i is the within transformation, which subtracts the individual specific mean from each observation, i.e. $\tilde{y}_{it} = y_{it} - \bar{y}_i$, where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$ is the individual specific mean. Similarly to remove the time variant fixed effect, we can use the time within transformation, which subtracts the time specific mean from each observation, i.e. $\dot{y}_{it} = y_{it} - \bar{y}_t$, where $\bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$ is the time specific mean. However, since we have both individual and time fixed effects, we need to use a two-way transformation to remove both effects. This can be done by applying the within transformation twice, first over the individual dimension and then over the time dimension (or vice versa).

Denote the transformation matrices,

$$Q_T = I_T - \frac{1}{T}\iota_T \iota_T', \quad Q_N = I_N - \frac{1}{N}\iota_N \iota_N',$$

The same transformation is applied to $\boldsymbol{x}_{it}$ and ε_{it}, i.e. $\tilde{\boldsymbol{x}}_{it} = \boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_i$ and $\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_i$. The time fixed effects, γ_t, can be removed by including time dummies in the regression (or by a similar transformation over the cross-sectional dimension). The transformed model is then given by,

$$\tilde{y}_{it} = \tilde{\boldsymbol{x}}_{it}\boldsymbol{\beta} + \tilde{\varepsilon}_{it}.$$

Gameplan:

Write stuff about assumptions, and then transformations, prob go from inconsistent pooled OLS to consistent FE. Right the estimator, list assumptions for identification and consistency (strict/cont exog, non-generacy) -¿ show consistency by LLN convergence in plim. Next, talk about asymptotic normality, CLT, variance estimation (efficiency - robust, cluster). Finally, hypothesis testing (Wald). Then proceed to empirical application (include wald-test code, since it is not given or done in any exercises).

- Mikkel: Requirements for FE and why pooled OLS is inconsistent (identification, assumptions, consistency, exogeneity).
- Mikkel: Show consistency (LLN $\rightarrow$ plim convergence).
- Mikkel: Show Asymptotic normality.
- Mathias: Show Efficiency (Robust or not).
- Noah: Seriel Corr, Rank test, Homoskedasticity (maybe).
- Noah: Wald-Test (one sided as well)? Check form of test when using robust std. err.
- Mathias: Potentially the above, for time variant fixed effects.
- Noah: Present empirical results.
- Noah: Estimation table.
- Noah: Misspecification test table.
- Noah: Test table for linear hypothesis (Wald).
- Missing references for eq's

Plan:

1. Alle prøver at nå at lave hver deres til/inden torsdag
2. Alle læser hinandens arbejde torsdag eftermiddag
3. Zoom møde fredag eftermidag, hvor vi får den sendt afsted

2 Methodology

2.1 Efficiency

It follows, that for the FE estimator to be efficient - i.e. have the smallest possible (asymptotic) variance -, we need to make some assumptions regarding the variance of the idiosyncratic error term, ε_{it}. Specifically, we need that the variance is homoscedastic and not serially correlated, i.e. $\text{Var}(\varepsilon_{it}) = \sigma^2$ and $\text{Cov}(\varepsilon_{it}, \varepsilon_{is}) = 0$ for $t \neq s$, or in the stacked notation, that, $\Omega_{\boldsymbol{\varepsilon}} = \sigma^2 I_T$. This results in the following (ref Mikkels eq for asymptotic normality),

$$\begin{aligned}\mathbb{E}\left[\ddot{\boldsymbol{X}}_i'\boldsymbol{u}_i\boldsymbol{u}_i'\ddot{\boldsymbol{X}}_i\right] &= \mathbb{E}\left[\ddot{\boldsymbol{X}}_i'\mathbb{E}[\boldsymbol{u}_i\boldsymbol{u}_i'|\boldsymbol{x}_i\alpha_i]\ddot{\boldsymbol{X}}_i\right] \\ &= \mathbb{E}[\ddot{X}_i'\sigma_u^2\boldsymbol{I}_T\ddot{\boldsymbol{X}}_i]\end{aligned}$$

such that the asymptotic variance is,

$$\text{Avar}(\hat{\beta}_{FE}) = \sigma_u^2\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i]^{-1}/N$$

Using the analogy principle, we have,

$$\hat{\text{Avar}}(\hat{\beta}_{FE}) = \hat{\sigma}_u^2 \left(\sum_{i=1}^{N} \ddot{\boldsymbol{X}}_i' \ddot{\boldsymbol{X}}_i \right)^{-1}, \quad \text{where} \quad \hat{\sigma}_u^2 = \frac{1}{N(T-1)-K} \sum_{i=1}^{N} \hat{\boldsymbol{u}}_i' \hat{\boldsymbol{u}}_i,$$
$$\text{and} \quad \hat{\boldsymbol{u}}_i = \ddot{\boldsymbol{y}}_i - \ddot{\boldsymbol{X}}_i \hat{\beta}_{FE}$$

Note that we need to correct the degrees of freedom for the implicit estimation of the fixed effects, which is why we subtract N from the denominator. If these assumptions do not hold, we can still obtain consistent estimates of the standard errors using robust or clustered standard errors, though often at the cost of efficiency. In which case we estimate the middle part of the sandwich from (Mikkel ref normality) using the analogy principle i.e. using the estimated residuals, such that,

$$\hat{\text{Avar}}(\hat{\beta}_{FE}) = \left(\sum_{i=1}^{N} \ddot{\boldsymbol{X}}_i' \ddot{\boldsymbol{X}}_i \right)^{-1} \left(\sum_{i=1}^{N} \ddot{\boldsymbol{X}}_i' \hat{\boldsymbol{u}}_i \hat{\boldsymbol{u}}_i' \ddot{\boldsymbol{X}}_i \right) \left(\sum_{i=1}^{N} \ddot{\boldsymbol{X}}_i' \ddot{\boldsymbol{X}}_i \right)^{-1}$$

Which is consistent under just the strict exogeneity assumption as well as the rank assumption - because consistent estimation of the error, as well as the outer matrices being invertible? (remove outer sums possibly and go full stacked like a maniac).

2.2 (Econometric theory and methodology)

Define $\boldsymbol{x}_{it} = [k_{it}, l_{it}]$ as the regressors of our model and y_{it} as the dependent variable, which are both observable. Further define the composite error term $v_{it} = c_i + u_{it}$, which contains both a time-invariant term c_i and a time-varying term u_{it}. Further define

Pooled OLS

A naive approach to formulate an econometric model would be to use pooled OLS for the case when we define the composite error term $u_{it} = v_{it} + \epsilon_{it}$

$$\begin{aligned} y_{it} &= \boldsymbol{x}_{it}\boldsymbol{\beta} + u_{it} \\ \Leftrightarrow \boldsymbol{x}_{it}' y_{it} &= \boldsymbol{x}_{it}' \boldsymbol{x}_{it} \boldsymbol{\beta} + \boldsymbol{x}_{it}' u_{it} \\ \Leftrightarrow \mathbb{E}[\boldsymbol{x}_{it}' y_{it}] &= \mathbb{E}[\boldsymbol{x}_{it}' \boldsymbol{x}_{it}]\boldsymbol{\beta} + \mathbb{E}[\boldsymbol{x}_{it}' u_{it}] \\ \Leftrightarrow \boldsymbol{\beta} &= (\mathbb{E}[\boldsymbol{x}_{it}' \boldsymbol{x}_{it}])^{-1} (\mathbb{E}[\boldsymbol{x}_{it}' y_{it}] - \mathbb{E}[\boldsymbol{x}_{it}' u_{it}]) \end{aligned}$$

It is evident that identification of $\boldsymbol{\beta}$ requires that the composite error u_{it} is uncorrelated with the regressor. Unpacking the composite error such that $\mathbb{E}[\boldsymbol{x}_{it}' u_{it}] = \mathbb{E}[\boldsymbol{x}_{it}' v_{it}] + \mathbb{E}[\boldsymbol{x}_{it}' \epsilon_{it}]$ shows that logged TFP (time-invariant unobservable) v_{it} must be assumed uncorrelated with logged capital and logged employment (regressors) in addition to $\mathbb{E}[\boldsymbol{x}_{it}\epsilon_{it}] = 0$, to successfully identify $\boldsymbol{\beta}$. An example of violation of $\mathbb{E}[\boldsymbol{x}_{it}' v_{it}] = 0$ would be if firms that have more capital (or employees) are generally more productive than firms with less capital (or less employees). This could be due to structural differences a.o.

If we were willing to assume $\mathbb{E}[\boldsymbol{x}_{it}' v_{it}] = 0$, then to estimate $\boldsymbol{\beta}$ using pooled OLS, we would stack y_{it} and $\boldsymbol{x}_{it}$ over (i, t) into matrices, obtaining $\boldsymbol{X}$ and $\boldsymbol{y}$, both two-dimensional of size (NT, K) and $(NT, 1)$ respectively.

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

However with a DGP as $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{v} + \boldsymbol{\epsilon}$, we would obtain a biased estimate since:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X\beta} + \boldsymbol{X}'\boldsymbol{v} + \boldsymbol{X}'\boldsymbol{\epsilon}) = \boldsymbol{\beta} + \boldsymbol{X}'\boldsymbol{v} +$$

(Ser lige hvad jeg gør med det her afsnit, skal nok skærre fra, strømline)

Fixed Effects

We now proceed to consider another estimator, namely the fixed effects estimator. Firstly, the data is transformed by a within-transformation, which demeans the data. In that regard, we denote the time-demeaned data with double dots (mere elegant), which results in the following equation for our model

$\forall\ t, i.$

$$\begin{aligned} y_{it} - \bar{y}_{it} &= \beta(\boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_{it}) + (c_i - c_i) + (u_{it} - \bar{u}_{it}) \\ \ddot{y}_{it} &= \beta\ddot{\boldsymbol{x}}_{it} + \ddot{u}_{it} \end{aligned}$$

We then consider the identification of β. Firstly, the data is stacked over time time-dimension (why is this smart?= så man slipper for at skrive summer) such that $\ddot{\boldsymbol{y}}_i$ and $\ddot{\boldsymbol{u}}_i$ are $T \times 1$ vectors and $\ddot{\boldsymbol{X}}$ is a $T \times K$ matrix.

$$\begin{aligned} \ddot{\boldsymbol{y}}_i &= \ddot{\boldsymbol{X}}_i\boldsymbol{\beta} + \ddot{\boldsymbol{u}}_i \\ \Leftrightarrow \ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{y}}_i &= \ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i\boldsymbol{\beta} + \ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i \\ \Leftrightarrow \mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{y}}_i] &= \mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i]\boldsymbol{\beta} + \mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i] \\ \Leftrightarrow \boldsymbol{\beta} &= (\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i])^{-1}\left(\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{y}}_i] - \mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i]\right) \end{aligned}$$

, i.e. we see that $\boldsymbol{\beta}$ is identified if we can assume the following:

1. Strict exogenity $\mathbb{E}[\ddot{\boldsymbol{u}}_i|\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{iT}, c_i] = 0$, which implies that the time-demeaned regressors and error terms are uncorrelated $\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i] = 0$ (derive + interpret?).
2. the matrix $\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i]$ fulfills a (full) rank condition such that is has full rank, i.e. $rank(\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i]) = K$ and can be inverted. Interpret

which would imply:

$$\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i] = 0 \Rightarrow \boldsymbol{\beta} = (\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{X}}_i])^{-1}\mathbb{E}[\ddot{\boldsymbol{X}}_i'\boldsymbol{y}_i]$$

Using the analogy principle, estimation of $\boldsymbol{\beta}$ would result in the following estimator $\hat{\boldsymbol{\beta}}_{FE}$, where we use that the expectation is over the individuals i.

$$\hat{\boldsymbol{\beta}}_{FE} = \left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\right)^{-1}\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{y}}_i$$

With $\ddot{\boldsymbol{X}}'\ddot{\boldsymbol{X}}$, we obtain a $(K \times K)$ matrix, which, with a $(K \times 1)$ vector from $\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{y}}_i$ results in a $(K \times 1)$ vector of the parameters.

Consistency of Fixed Effects estimator

Now, we turn to the consistency of the FE-estimator. Inserting $\ddot{\boldsymbol{y}}_i$ in the expression for $\hat{\boldsymbol{\beta}}_{FE}$, we obtain the following:

$$\begin{aligned} \Rightarrow \hat{\boldsymbol{\beta}}_{FE} &= \left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\boldsymbol{\beta} + \ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{u}}_i\right) \\ &= \boldsymbol{\beta} + \left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i\right) \end{aligned}$$

The Law of Large Numbers (LLN) implies that $p\text{-lim}\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\right) = \mathbb{E}[\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i]$ and $p\text{-lim}\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{u}}_i\right) = \mathbb{E}[\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{u}}_i]$. Then we use Slutskys theorem: $p\text{-lim}\left(\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i\right)^{-1}\right) = (\mathbb{E}[\ddot{\boldsymbol{X}}'_i\ddot{\boldsymbol{X}}_i])^{-1}$, which is allowed since the inverse matrix is defined, due to the rank assumption, and the inverse operator is a continuous function. Altogether, it implies that the FE-estimator $\hat{\boldsymbol{\beta}}_{FE}$ is consistent, which is a first-order concern.

2.2.1 Asymptotic normality of Fixed Effects estimator

When analysing our consistent FE-estimator, we need to know the asymptotic distribution of the parameter. Rearranging the equation from the former section, we get the following:

$$\sqrt{N}(\hat{\boldsymbol{\beta}}_{FE}-\boldsymbol{\beta}) = \left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}\boldsymbol{\prime}_i\ddot{\boldsymbol{X}}_i\right)^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i\right)$$

From the former section, we saw that the Law of Large Numbers and Slutskys Theorem combined with our assumptions (strict exogeneity and full rank condition) implied that asymptotically $\mathbb{E}[\sqrt{N}(\hat{\boldsymbol{\beta}}_{FE}-\boldsymbol{\beta})]=\mathbf{0}$. This means that the sum of real-valued i.i.d. r.v.s converges accordingly to the Central Limit Theorem such that $\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i\right) \to_d \mathrm{N}(\mathbf{0},\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i(\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i)']) =^d \mathrm{N}(\mathbf{0},\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i\ddot{\boldsymbol{u}}_i'\ddot{\boldsymbol{X}}_i])$.

Since we have convergence in probability $p\text{-lim}\left(\left(\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{X}}\boldsymbol{\prime}_i\ddot{\boldsymbol{X}}_i\right)^{-1}\right) = (\mathbb{E}[\ddot{\boldsymbol{X}}\boldsymbol{\prime}_i\ddot{\boldsymbol{X}}_i])^{-1}$ for the first term and convergence in distribution for the second term, the product rule implies that the expression for $\sqrt{N}(\hat{\boldsymbol{\beta}}_{FE}-\boldsymbol{\beta})$ converges only in the distributional sense such that:

$$\Rightarrow \sqrt{N}(\hat{\boldsymbol{\beta}}_{FE}-\boldsymbol{\beta}) \sim \mathrm{N}\left(\mathbf{0},(\mathbb{E}[\ddot{\boldsymbol{X}}\boldsymbol{\prime}_i\ddot{\boldsymbol{X}}_i])^{-1}\mathbb{E}[\ddot{\boldsymbol{X}}_i'\ddot{\boldsymbol{u}}_i\ddot{\boldsymbol{u}}_i'\ddot{\boldsymbol{X}}_i](\mathbb{E}[\ddot{\boldsymbol{X}}\boldsymbol{\prime}_i\ddot{\boldsymbol{X}}_i])^{-1}\right)$$

Wald-test

(We could also test for joint significance for the year dummies)
(Rækkefølgen skal ændres på, da nogle antagelser bruges i de tidligere underafsnit)
We utilise the Wald-test to test for CRS, where the null-hypothesis is $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$. Under weak regularity conditions - consistency of the FE estimator, asymptotic normality and $\mathbf{R}$ has full rank (*refer to consistency, and asymptotic normality equations*), the Wald statistic can be written as:

$$W := (\mathbf{R}\widehat{\boldsymbol{\beta}}-\mathbf{r})[\mathbf{R}\widehat{\mathrm{Avar}(\widehat{\beta})}\mathbf{R}']^{-1}(\widehat{\mathbf{R}\boldsymbol{\beta}-}\mathbf{r})$$

Under the H_0 the Wald statistic follows a χ^2_Q distribution. When testing for CRS we test $H_0 : \beta_K + \beta_K = 1$, where we have $Q = 1$ degrees of freedom and $\mathbf{R}$ is a $[1 \times K]$ matrix and $\mathbf{r} = 1 \times 1$ matrix, where K is the rank $K = E[\mathbf{x}'\mathbf{x}]$ from assumption FE.2. The Walt-statistic is therefore $W := 2$

Serial correlation

Homoskedasticity (maybe)

3 Empirical results

The data consists of $N = 441$ French manufacturing firms over 12 years ($T = 12$) from 1968 to 1979, and is thus a panel data set with $NT = 5292$ observations. The data is balanced.

Maybe variable def.

We perform the two-way or double-demean transformation of our data matrices as described in the methodology section (Ref). The estimates along with standard errors and p-values are reported in Table **??**.

Table 1: Fixed Effects Regression Results

	β	Se	t-values	p-value
ℓ	0.6942	0.0147	47.2447	0.0000
k	0.1546	0.0130	11.9311	0.0000
R^2	0.477			
σ^2	0.018			

Next we proceed to test the null hypothesis of constant returns to scale. From Table **??**, we reject the null at the 1 % significance level with a p-value of ≈ 0. Thus, we conclude that the (Cobb-Douglas) production function does not exhibit constant returns to scale, for French manufacturing firms in the period 1968-1979.

Table 2: Wald Test Results

Wald stat	df	$\chi^2_{1(0.95)}$	p-value
135.190	1	3.842	0.000

In addition one might want to test whether the production function exhibits increasing or decreasing returns to scale. This can be done by testing the null hypothesis of $\beta_K + \beta_L \geq 1$ against the alternative of $\beta_K + \beta_L < 1$ (decreasing returns to scale) or vice versa for increasing returns to scale. Importantly we need to consider that the test includes the boundary i.e. $\beta_K + \beta_L = 1$ such that we cant draw a conclusion if the null is accepted, which is why the tests are read somewhat opposite of what one might expect. One might suspect that this particular sector exhibits increasing returns to scale, as firms in manufacturing often benefit from economies of scale, which indeed is the case as seen from Table **??**.

Table 3: One-Sided Tests for Returns to Scale

Hypothesis	Estimate	t-stat	p-value	Reject H_0
Increasing RTS ($\beta_K + \beta_L > 1$)	0.849	-11.627	1.000	No
Decreasing RTS ($\beta_K + \beta_L < 1$)	0.849	-11.627	0.000	Yes

References

Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data.* MIT press.