

Applied Production Analysis

SOK-3011—Part 1

Dual approach:

Estimation of cost and profit functions

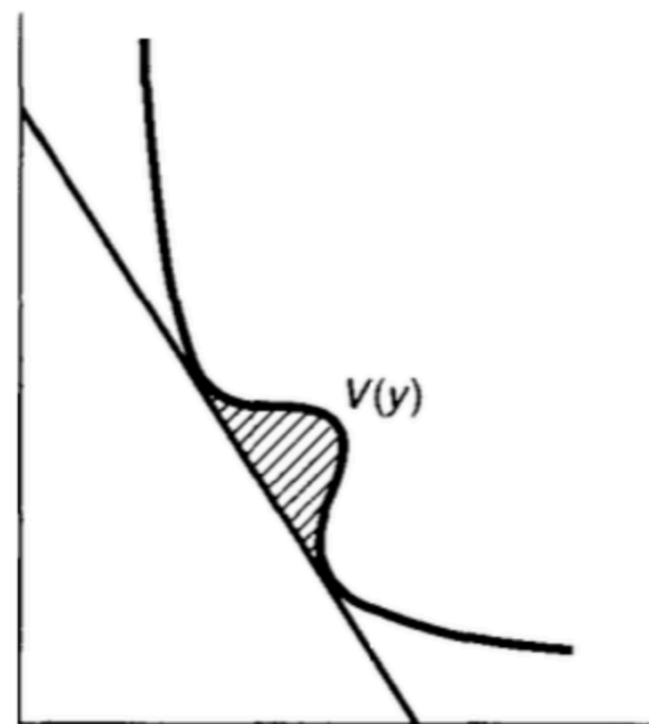
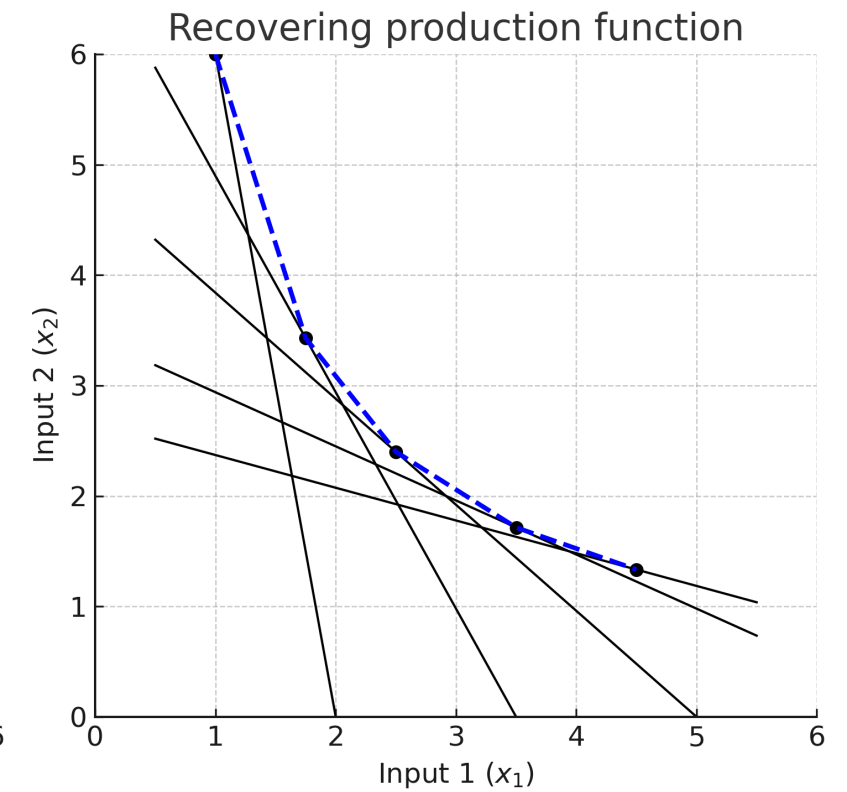
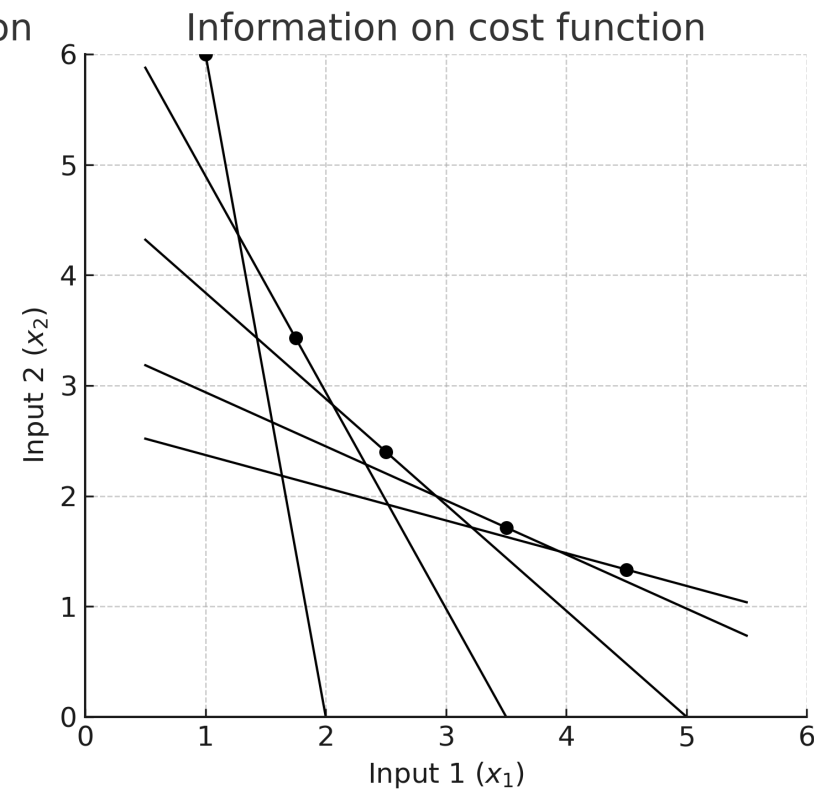
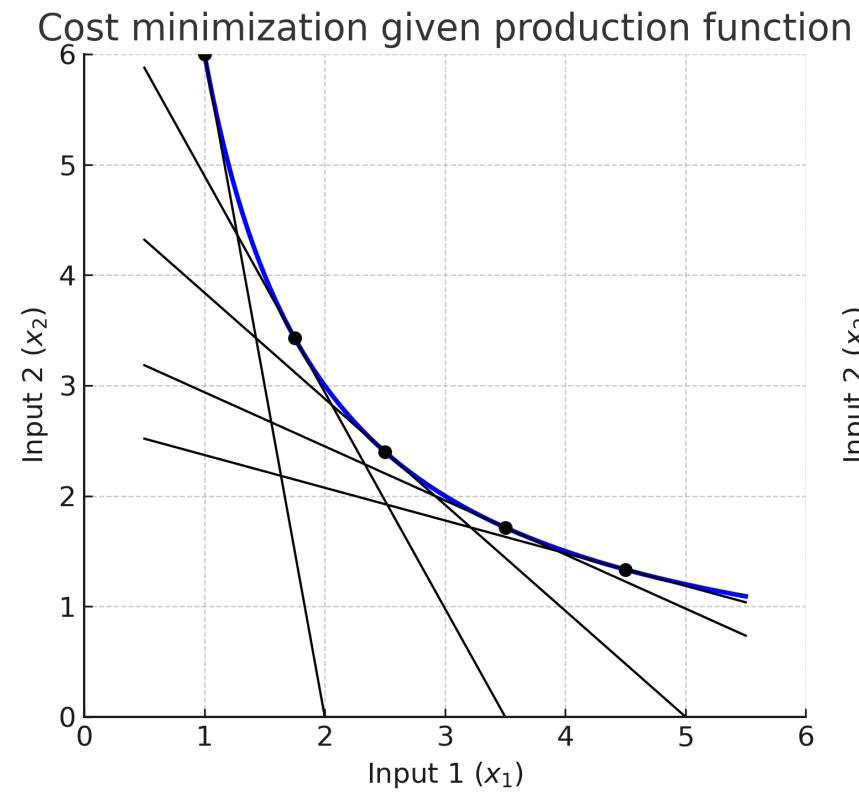
Overview

- Dual Approach
- Cost function
 - Recovery of production function
 - Properties
 - Estimation
- Profit function
 - Properties
 - Estimation

Dual Approach (vs. Primal Approach)

- **Primal approach:** Start with the **production function** → analyze how firms maximize profit or minimize cost given technology.
- **Dual approach:** Start with the **cost function** → analyze how much it costs to produce a given output at given input prices.
- Key duality insight:
 - From production function → we can derive the cost function.
 - From cost function → we can recover the production function (if technology is well-behaved, i.e., quasi-concave).

Recovery of production function



Recovery of production function

Theorem.

Let $c(y, w)$ be a differentiable function satisfying the regular properties (stated later) of a cost function.

Define a function f as

$$f(x) \equiv \max\{y \geq 0 \mid w \cdot x \geq c(y, w), \forall w \gg 0\}.$$

Then:

- f is increasing, unbounded above, and quasiconcave.
- Moreover, the cost function generated by f is precisely c .

Dual Approach (vs. Primal Approach)

- In the context of empirical estimation of technology, we can address it by studying the same problem from two mirrors:
 - *Primal*: What output can I get with these inputs?
 - *Dual*: What is the minimum cost of producing this output?
- In practice, the dual approach is often more useful because **prices and expenditures are visible in markets**, while internal technology is hidden.
- Applied advantage:
 - Easier to observe **input prices and costs** than to measure technical production details.
 - Researchers can estimate cost functions directly from market data and then infer the production side.

Cost function

Cost minimization and cost function

- **Cost Minimization Problem:**

$$\min_x \sum_{i=1}^n w_i x_i \quad \text{s.t.} \quad f(x) \geq y$$

- **Cost Function:**

$$c(y, w) = \sum_{i=1}^n w_i x_i(y, w)$$

where $x_i(y, w)$ is the cost-minimizing demand for input i .

- **Interpretation:** Minimum cost of producing output y given input prices w .

Properties of cost function

If production function f is continuous and strictly increasing:

1. $c(0, w) = 0$
2. Continuous on domain
3. Strictly increasing & unbounded in y (for $w \gg 0$)
4. Increasing in input prices w
5. Homogeneous of degree one in w
6. Concave in w

Properties of cost function

- An important result is the relationship between the cost function and the optimal input demand function

- **Shephard's Lemma**

$$\frac{\partial c(y, w)}{\partial w_i} = x_i(y, w)$$

- This is a direct application of the envelope theorem for constrained optimization
- We can use this result to check if firms' choices of inputs are consistent with cost minimization principle

Estimation of cost function

A Cobb-Douglas cost function

- We will estimate a Cobb-Douglas cost function using our data.
- Note that a Cobb-Douglas production function generates a cost function that exhibits a Cobb-Douglas form in input prices.
- If the original production technology exhibits constant returns to scale (CRS), the cost function is typically linear in y .
- We consider the following specification:

$$c = A \left(\prod_{i=1}^N w_i^{\alpha_i} \right) y^{\alpha_y}$$

and in its linearized form:

$$\ln c = \alpha_0 + \sum_{i=1}^N \alpha_i \ln w_i + \alpha_y \ln y$$

Estimation

- Coefficients of log-input prices are **non-negative** → cost function is non-decreasing in input prices.

- Coefficient of log-output quantity is **non-negative** → cost function is non-decreasing in output.

- With a **positive intercept**, the cost function satisfies the **non-negativity condition**.

```
lm(formula = log(cost) ~ log(pCap) + log(pLab) + log(pMat) +
    log(qOut), data = dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.77663	-0.23243	-0.00031	0.24439	0.74339

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.75383	0.40673	16.605	< 0.0000000000000002 ***
log(pCap)	0.07437	0.04878	1.525	0.12969
log(pLab)	0.46486	0.14694	3.164	0.00193 **
log(pMat)	0.48642	0.08112	5.996	0.0000000174 ***
log(qOut)	0.37341	0.03072	12.154	< 0.0000000000000002 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3395 on 135 degrees of freedom

Multiple R-squared: 0.6884, Adjusted R-squared: 0.6792

F-statistic: 74.56 on 4 and 135 DF, p-value: < 0.00000000000000022

Does it satisfy homogeneity?

- The **F-statistic** tests whether the null hypothesis (sum of coefficients = 1) holds.

```
# Linear homogeneity condition check
linearHypothesis( costCD, "log(pCap) + log(pLab) + log(pMat) = 1" )
```

- It compares the fit of a **restricted model** (with the null imposed) and an **unrestricted model**.

Linear hypothesis test:

$\log(\text{pCap}) + \log(\text{pLab}) + \log(\text{pMat}) = 1$

Model 1: restricted model

Model 2: $\log(\text{cost}) \sim \log(\text{pCap}) + \log(\text{pLab}) + \log(\text{pMat}) + \log(\text{qOut})$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	136	15.563				
2	135	15.560	1	0.0025751	0.0223	0.8814

- A small difference in residuals → null likely true → **large Pr(>F)** (as in this case).

Imposing homogeneity

- We can directly impose homogeneity of degree one by modifying the model.
- This involves setting one input's coefficient as **1 minus the sum of the others**, which changes the regression specification.

$$\ln c = \alpha_0 + \sum_{i=1}^{N-1} \alpha_i \ln w_i + (1 - \sum_{i=1}^{N-1} \alpha_i) \ln w_N + \alpha_y \ln y$$

or, equivalently,

$$\ln \frac{c}{w_N} = \alpha_0 + \sum_{i=1}^{N-1} \alpha_i \ln \frac{w_i}{w_N} + \alpha_y \ln y$$

Imposing homogeneity

```
lm(formula = log(cost/pMat) ~ log(pCap/pMat) + log(pLab/pMat) +  
    log(qOut), data = dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.77096	-0.23022	-0.00154	0.24470	0.74688

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.75288	0.40522	16.665	< 0.00000000000000002 ***
log(pCap/pMat)	0.07241	0.04683	1.546	0.124
log(pLab/pMat)	0.44642	0.07949	5.616	0.000000106 ***
log(qOut)	0.37415	0.03021	12.384	< 0.00000000000000002 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3383 on 136 degrees of freedom

Multiple R-squared: 0.5456, Adjusted R-squared: 0.5355

F-statistic: 54.42 on 3 and 136 DF, p-value: < 0.000000000000000022

- Coefficients change slightly but keep the **same interpretation**; the “materials” coefficient is derived as $\alpha_{MAT} = 0.48$.

Homogeneity-imposed model

Likelihood ratio test

Model 1: $\log(\text{cost}/\text{pMat}) \sim \log(\text{pCap}/\text{pMat}) + \log(\text{pLab}/\text{pMat}) + \log(\text{qOut})$

Model 2: $\log(\text{cost}) \sim \log(\text{pCap}) + \log(\text{pLab}) + \log(\text{pMat}) + \log(\text{qOut})$

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	5	-44.878			
2	6	-44.867	1	0.0232	0.879

- A Likelihood-ratio (LR) test compares unrestricted vs. restricted models; here, the large $\text{Pr}(>\text{ChiSq})$ means the null holds \rightarrow restricted model fits well.
- We adopt the homogeneity-imposed model for further analysis.
- Advantage: for Cobb-Douglas cost functions, **non-decreasing + linear homogeneity** automatically imply **concavity**.

Are firms' input choices consistent with cost-minimization principle?

- An implication of Shephard's lemma:

$$\alpha_i = \frac{\partial \ln c(y, \mathbf{w})}{\partial \ln w_i} = \frac{\partial c(y, \mathbf{w})}{\partial w_i} \cdot \frac{w_i}{c(y, \mathbf{w})} = \frac{w_i \cdot x_i(y, \mathbf{w})}{c(y, \mathbf{w})} = \text{cost share of input } i$$

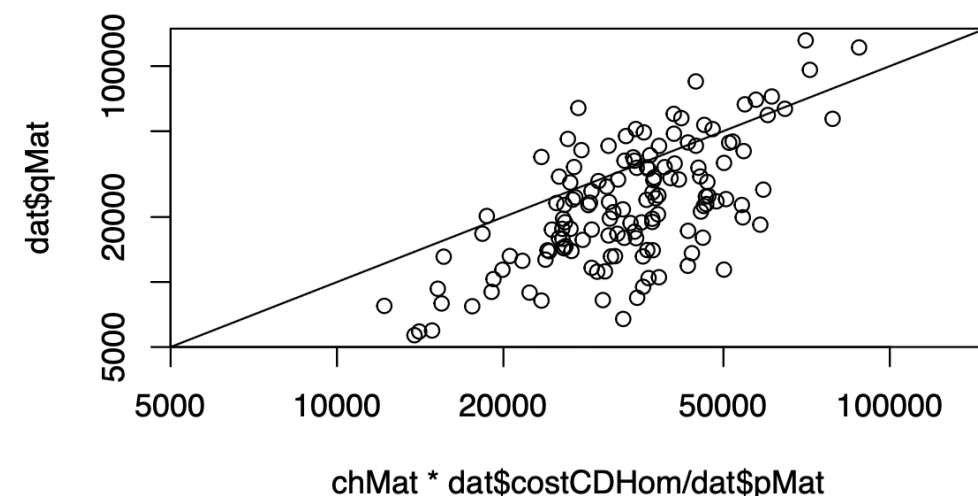
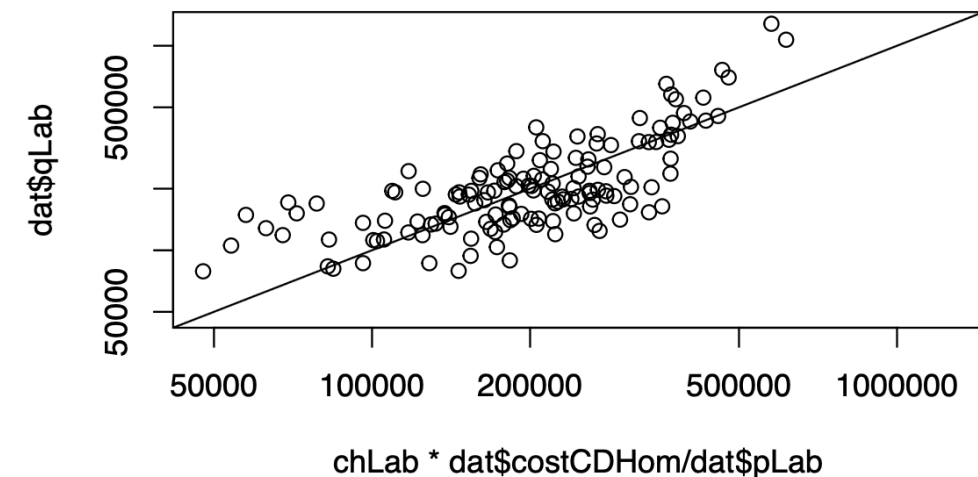
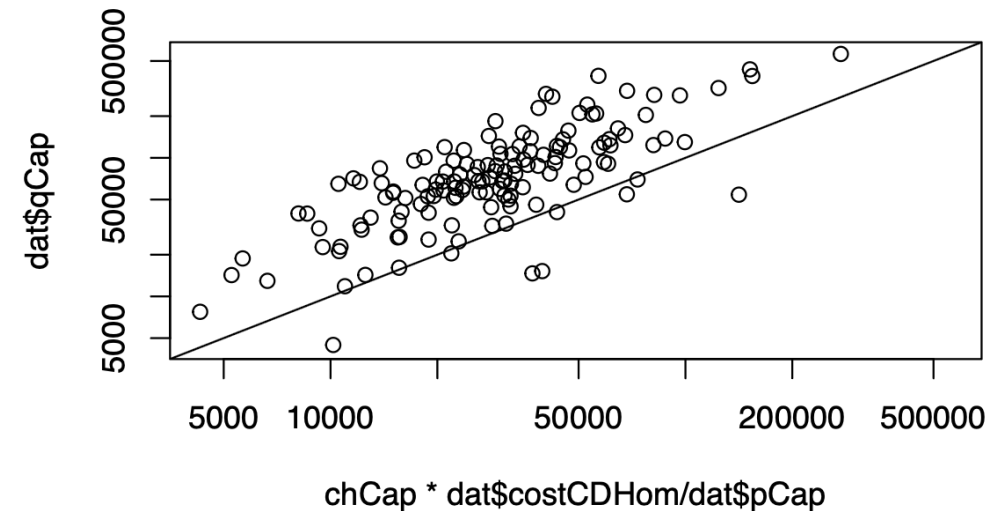
where the first equality follows from Cobb-Douglas specification and the third follows from the Shephard's lemma.

- This can also be rewritten as

$$\alpha_i \frac{c(y, \mathbf{w})}{w_i} = x_i(y, \mathbf{w})$$

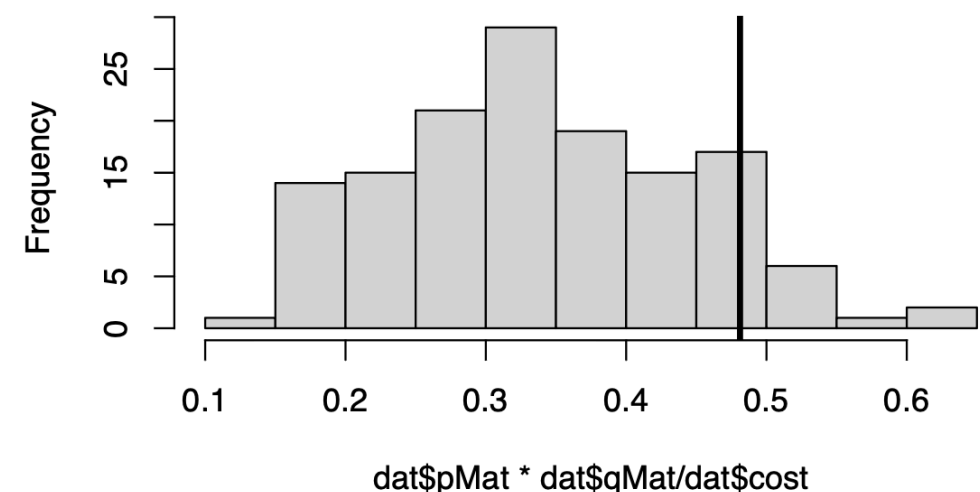
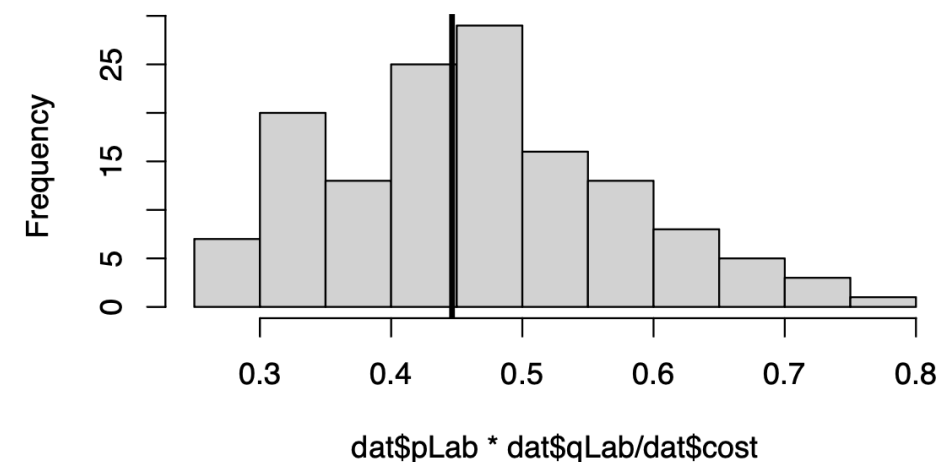
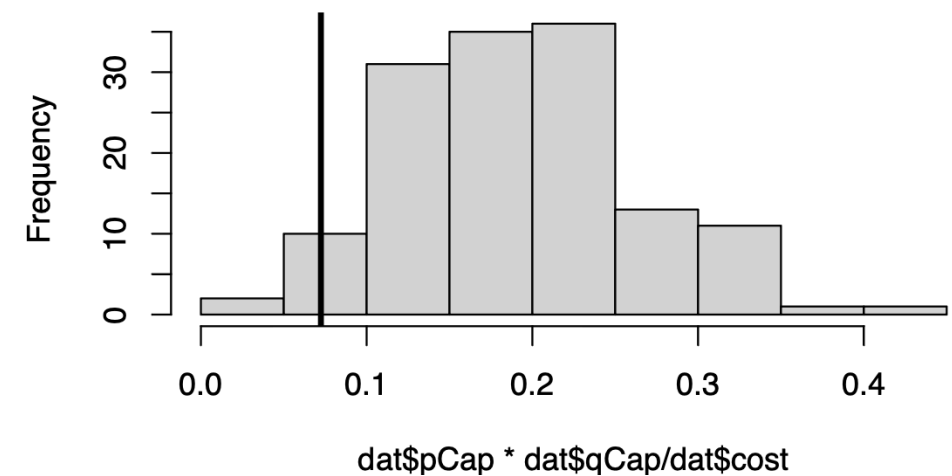
Are firms' input choices consistent with cost-minimization principle?

- If $(\alpha_i C)/w_i < x_i$, the firm is **overusing the input** relative to cost-minimization.
- The scatter plots show input use across firms.
- Findings: most firms **overuse capital** and **underuse materials**.



Are firms' input choices consistent with cost-minimization principle?

- We can also check the distribution of $x_i w_i / c$ across firms and compare it against α_i
- If α_i is sufficiently low, it implies firms are overusing the input.
- Findings: most firms **overuse capital** and **underuse materials**.



Profit function

Profit maximization and profit function

- **Profit Maximization Problem:**

$$\max_{x,y} \pi = p \cdot y - \sum_{i=1}^n w_i x_i \quad \text{s.t. } f(x) \geq y$$

- **Profit Function:**

$$\pi(p, w) = p \cdot y(p, w) - \sum_{i=1}^n w_i x_i(p, w)$$

where $x_i(p, w)$ is the profit-maximizing demand for input i .

- **Interpretation:** Maximum profit attainable given output price p and input prices w .

Properties of profit function

If f is continuous, strictly increasing, and strictly quasiconcave:

1. Continuous
2. Increasing in output price p
3. Decreasing in input prices w
4. Homogeneous of degree one in (p, w)
5. Convex in (p, w)

Properties of profit function

- Like the Shephard's lemma, there is also an important result on the relationship between the profit function and the optimal output function

Hotelling's Lemma

- If $\pi(p, w)$ is differentiable at $(p, w) \gg 0$:

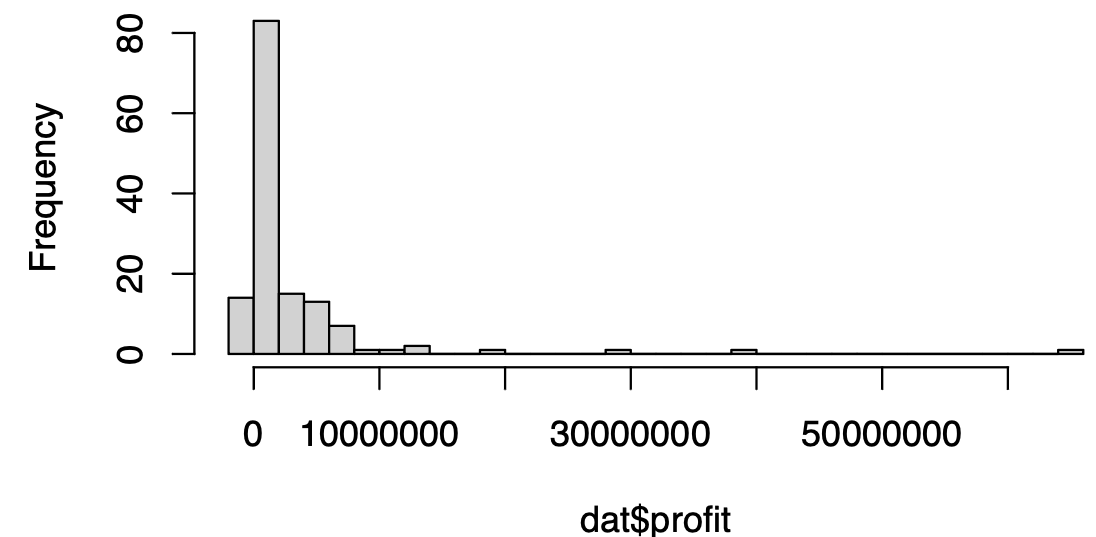
$$\frac{\partial \pi(p, w)}{\partial p} = y(p, w), \quad -\frac{\partial \pi(p, w)}{\partial w_i} = x_i(p, w)$$

- The output and input demand functions can be recovered from the profit function. These can be derived via **envelope theorem** (analogous to Shephard's lemma).
- We can use this result to check if firms' choices of inputs and output volume are consistent with profit maximization principle

Estimation of profit function

A Cobb-Douglas profit function

- We define **Profit = Revenue – Costs**, assuming full information on both.
- In our data set, the constructed profit measure shows **negative profits** for several firms.
- Negative profits may reflect **short-run conditions**, where not all inputs can be adjusted (e.g., **capital fixed**).
- Even after treating capital costs as fixed, **8 observations** still show negative profits.
- For estimation, we **exclude 14 observations** with negative profits (see Henningsen 2024 for discussion).



A Cobb-Douglas profit function

- We estimate a Cobb-Douglas form using our data.
- We consider the following specification:

$$\pi = Ap^{\alpha_p} \left(\prod_{i=1}^N w_i^{\alpha_i} \right)$$

and in its linearized form:

$$\ln \pi = \alpha_0 + \alpha_p \ln p + \sum_{i=1}^N \alpha_i \ln w_i$$

Estimation

- Profit function is **decreasing in capital and labor prices**; coefficient for materials is positive but **not statistically significant**.

- Profit function is **increasing in output price** (positive, significant coefficient).

- With a **positive intercept**, the profit function satisfies the **non-negativity condition**.

```
lm(formula = log(profit) ~ log(pOut) + log(pCap) + log(pLab) +
    log(pMat), data = dat_clean)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.6183	-0.2778	0.1261	0.5986	2.0442

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.9380	0.4921	28.321	< 0.0000000000000002 ***
log(pOut)	2.7117	0.2340	11.590	< 0.0000000000000002 ***
log(pCap)	-0.7298	0.1752	-4.165	0.0000586 ***
log(pLab)	-0.1940	0.4623	-0.420	0.676
log(pMat)	0.1612	0.2543	0.634	0.527

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9815 on 121 degrees of freedom

Multiple R-squared: 0.5911, Adjusted R-squared: 0.5776

F-statistic: 43.73 on 4 and 121 DF, p-value: < 0.00000000000000022

Does it satisfy homogeneity?

Linear hypothesis test:

$\log(p_{\text{Out}}) + \log(p_{\text{Cap}}) + \log(p_{\text{Lab}}) + \log(p_{\text{Mat}}) = 1$

Model 1: restricted model

Model 2: $\log(\text{profit}) \sim \log(p_{\text{Out}}) + \log(p_{\text{Cap}}) + \log(p_{\text{Lab}}) + \log(p_{\text{Mat}})$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	122	119.78				
2	121	116.57	1	3.2183	3.3407	0.07005 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- **F-test:** Null hypothesis of homogeneity of degree one is **rejected at 10%**, but **not at 5%** → may be critical if alternative functional forms offer better properties.

Imposing homogeneity

- We impose homogeneity of degree one by modifying the model.
- This involves setting one input's coefficient as **1 minus the sum of the others**, which changes the regression specification.

$$\ln \pi = \alpha_0 + \left(1 - \sum_{i=1}^N \alpha_i\right) \ln p + \sum_{i=1}^N \alpha_i \ln w_i$$

or, equivalently,

$$\ln \frac{\pi}{p} = \alpha_0 + \sum_{i=1}^N \alpha_i \ln \frac{w_i}{p}$$

Imposing homogeneity

```
lm(formula = log(profit/pOut) ~ log(pCap/pOut) + log(pLab/pOut) +  
    log(pMat/pOut), data = dat_clean)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.6045	-0.2724	0.0972	0.6013	2.0385

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	14.27961	0.45962	31.068	< 0.00000000000000002 ***
log(pCap/pOut)	-0.82114	0.16953	-4.844	0.00000378 ***
log(pLab/pOut)	-0.90068	0.25591	-3.519	0.000609 ***
log(pMat/pOut)	-0.02469	0.23530	-0.105	0.916610

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9909 on 122 degrees of freedom

Multiple R-squared: 0.3568, Adjusted R-squared: 0.341

F-statistic: 22.56 on 3 and 122 DF, p-value: 0.00000000001091

Likelihood ratio test

Model 1: log(profit) ~ log(pOut) + log(pCap) + log(pLab) + log(pMat)

Model 2: log(profit/pOut) ~ log(pCap/pOut) + log(pLab/pOut) + log(pMat/pOut)

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	6	-173.88			
2	5	-175.60	-1	3.4316	0.06396 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- As this model is nested within the original model, we can do a likelihood ratio test.
- Chi-square test:** Null rejected at **10%** but not at **5%**, similar to earlier homogeneity test.
- The **homogeneity-imposed model** ensures coefficients sum to one → used for further analysis.

Are firms' choices consistent with profit-maximization principle?

- An implication of Hotelling's lemma:

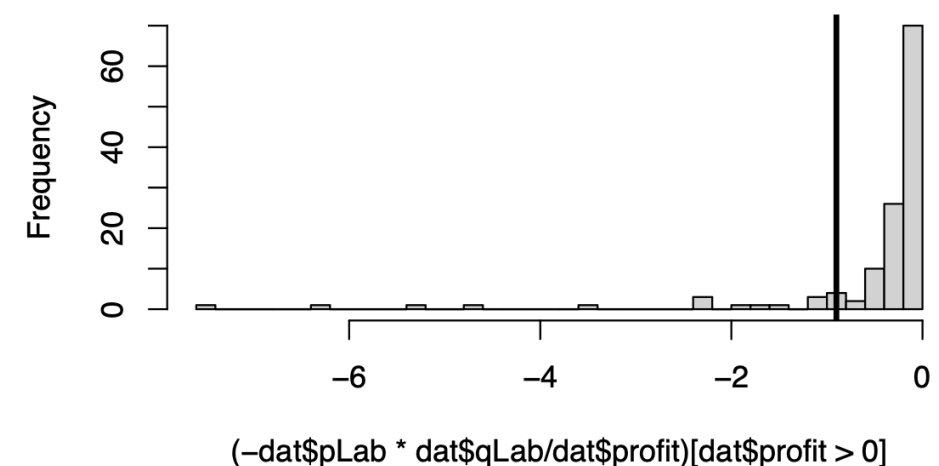
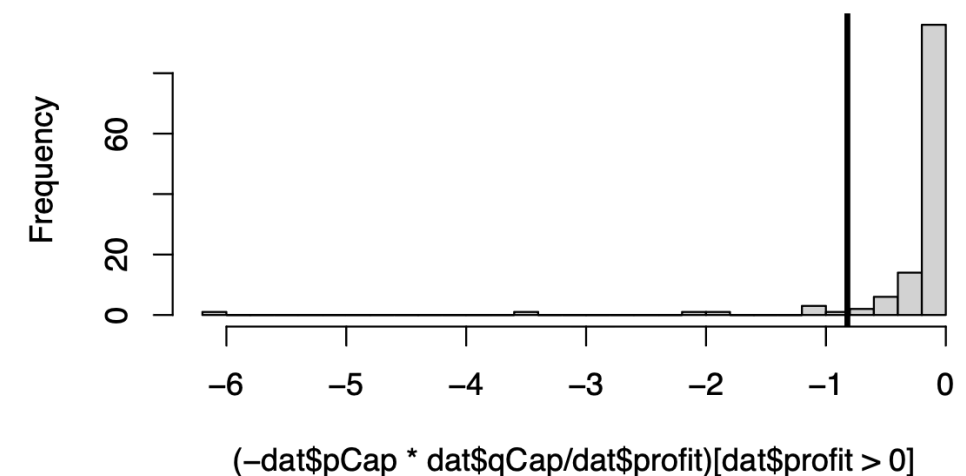
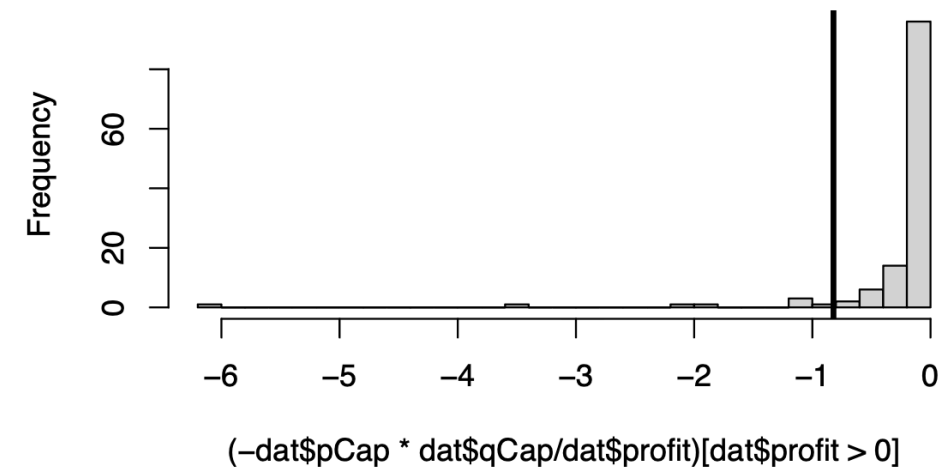
$$\alpha_p = \frac{\partial \ln \pi}{\partial \ln p} = \frac{\partial \pi}{\partial p} \cdot \frac{p}{\pi} = \frac{p \cdot y}{\pi} = \text{profit share of output } (\geq 1)$$

$$\alpha_i = \frac{\partial \ln \pi}{\partial \ln w_i} = \frac{\partial \pi}{\partial w_i} \cdot \frac{w_i}{\pi} = -\frac{w_i \cdot x_i(p, \mathbf{w})}{\pi} = \text{profit share of input } i \ (\leq 0)$$

- This can also be rewritten as **Profit shares** differ from cost shares \rightarrow not bounded between 0 and 1.
- Example: profit share of **2 (output)** = revenue twice the realized profit; profit share of **-1/2 (input)** = input cost equals half of realized profit.
- Comparing actual profit shares with α_i shows whether firms overuse/underuse inputs or deviate from profit-maximizing output.

Are firms' choices consistent with profit-maximization principle?

- Histograms show most firms operate below the optimal profit-maximizing level (similar to scale elasticity findings).
- Input use: firms could benefit from **more capital** and **less materials**, though this contrasts with earlier results.
- Possible reasons: short-run capital constraints or need to test alternative profit function forms (see Henningsen 2024).



Reading materials

- Varian, Chapter 3 and 5
- Henningsen, Chapter 3 and 4