Applied Production Analysis

SOK-3011—Part 1

Estimation of production functions

Overview

- Estimation: Ordinary Least Squares (OLS)
- Production Function Properties & Firm Performance
- Functional Forms:
  - Linear Technology
  - Cobb–Douglas
- Model Comparison

OLS

- In this course, we will use **linear model estimation techniques**. When can we appropriately estimate a production function using OLS?

- Key Conditions:

  - **Similar conditions:** Firms operate under comparable settings, or differences are modeled.

  - **Representative sample:** No selection bias; firms must represent the industry fairly.

  - **Frontier assumption:** Firms produce at maximum feasible output; deviations = random noise.

  - **No perfect multicollinearity:** Inputs should not be exact linear combinations.

  - **Exogeneity:** Input quantities must be uncorrelated with error terms.

Evaluating an estimated production function

- When we estimate a specific form of the production function with our data, we will check the properties of the estimated model and evaluate:

- **Goodness of fit**

- **Theoretical consistency** of the estimated model

- **Properties of the model**
  – Productivity and output elasticity
  – Returns to scale

- **Firm behavior**
  – Efficient employment of inputs
  – Profit-maximizing behavior
  – Cost-minimizing behavior

Linear technology

Linear technology

$$y = \beta_0 + \sum_{i=1}^{N} \beta_i x_i + \epsilon$$

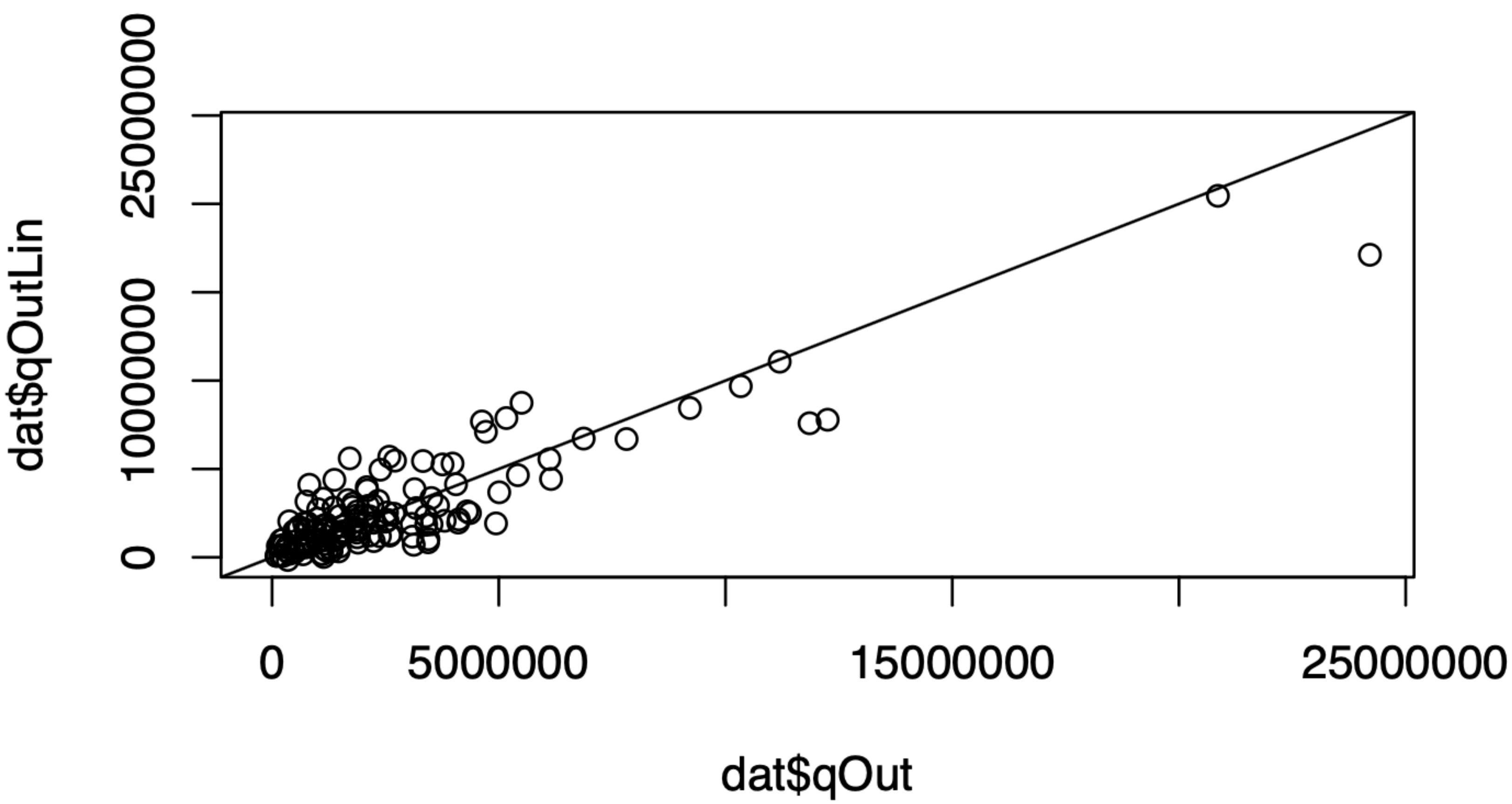
```
# Fitting a linear model
prodlinear <- lm( qOut ~ qCap + qLab + qMat, data = dat )
summary( prodlinear )
```

```
Call:
lm(formula = qOut ~ qCap + qLab + qMat, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-3888955  -773002    86119   769073  7091521

Coefficients:
                Estimate   Std. Error t value             Pr(>|t|)
(Intercept) -1615978.639   231771.709  -6.972 0.000000000123189025 ***
qCap               1.788        1.995   0.896                0.372
qLab              11.831        1.272   9.300 0.000000000000000315 ***
qMat              46.668       11.234   4.154 0.000057378294854380 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1541000 on 136 degrees of freedom
Multiple R-squared:  0.7868,    Adjusted R-squared:  0.7821
F-statistic: 167.3 on 3 and 136 DF,  p-value: < 0.00000000000000022
```

- Labor and materials: coefficients are **positive and significant**; capital: **not significant**.
- Dropping capital may cause **bias** if it is truly an essential input.
- Keeping capital ensures **unbiased estimates**, but reduces efficiency (less precise estimates).

Assessing model fit

- Linear production function may or may not be a good fit.
  R^2 value gives a rough indication of fit.



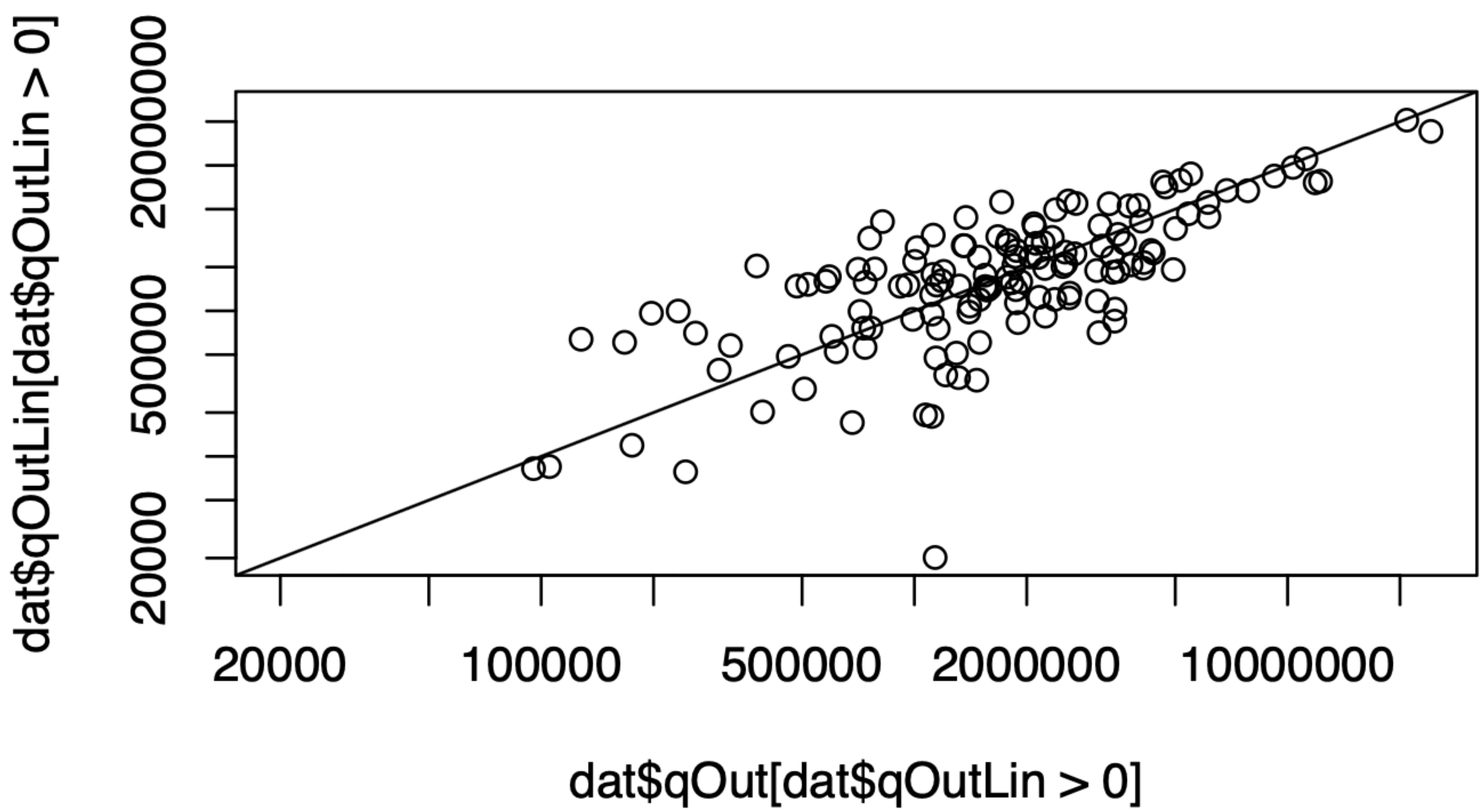
- Scatter plot of predicted vs. observed values:

  Good fit → points cluster around the 45° line.
  Curved patterns → linear model may be inappropriate.
  Outliers may distort results.

Practical consideration

- Plot issues: many points cluster near the origin due to extreme values.
- Possible solution: rescale axes (e.g., logarithmic), but exclude negative values.
- Scatter plots suggest deviations are mostly random → acceptable fit.

```
compPlot( dat$qOut[ dat$qOutLin > 0 ], dat$qOutLin[ dat$qOutLin > 0 ], log = "xy" )
```



Theoretical consistency of the estimated model

- **Essentiality**

  —*Weak:* each input contributes positively when others are held constant.
  —*Strict:* output = 0 if any input is missing.
  —Labor & materials: positive coefficients → essential.
  —Negative intercept → violates weak essentiality.

- **Monotonicity**

  —Output should not fall when an input increases.
  —Labor & materials: positive coefficients → monotonicity satisfied.

- **Quasi-concavity**

  —Requires convex isoquants (convex input requirement set).
  —Linear function → isoquants are straight lines → **trivially satisfied**.

- **Non-negativity**

  —Output ≥ 0 for non-negative inputs.
  —Negative intercept and one negative predicted output → **violation**.

Productivity and output elasticity

- Linear production function
  - Marginal productivity = coefficient (same across firms).
  - Cannot capture variation in productivity across firms.
  - To account for variation → compute **output elasticity**:

$$\varepsilon_i = \frac{\hat{\beta}_i}{\text{Average Product of input } i}$$

- Cobb–Douglas production function
  - Regression coefficients = elasticities directly.
  - Captures firm-level variation in marginal products.

Empirical findings

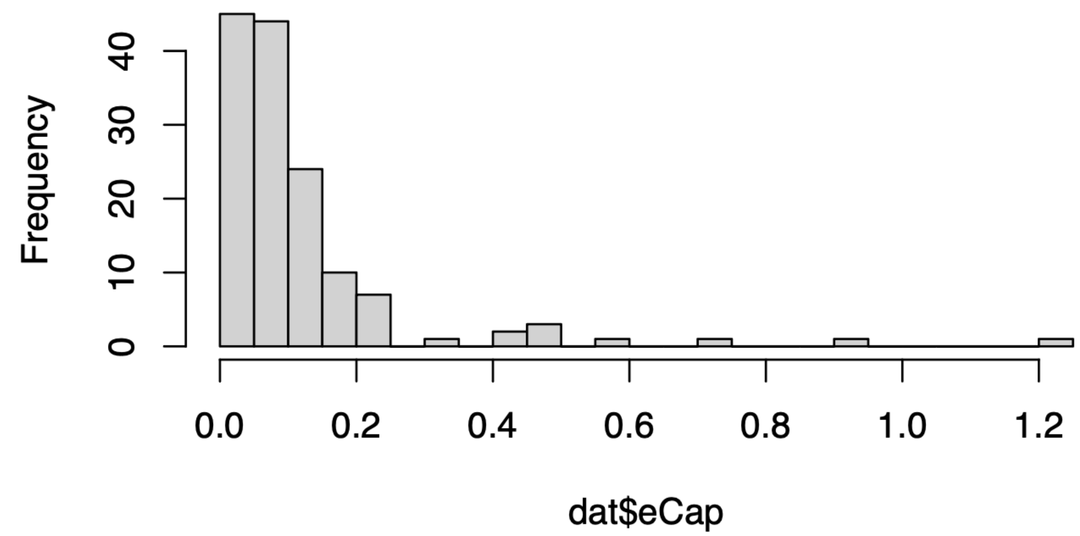
- Elasticity interpretation: % change in output from % change in input.
- **Capital:** marginal effect small (≈ 0–0.2%).
- **Labor:** elasticities often 0.5–3 (many > 1).
- **Materials:** elasticities mostly 0.2–1.2 (many > 1).
- Implausibly high elasticities: 124 out of 140 firms show elasticity measures above 1.

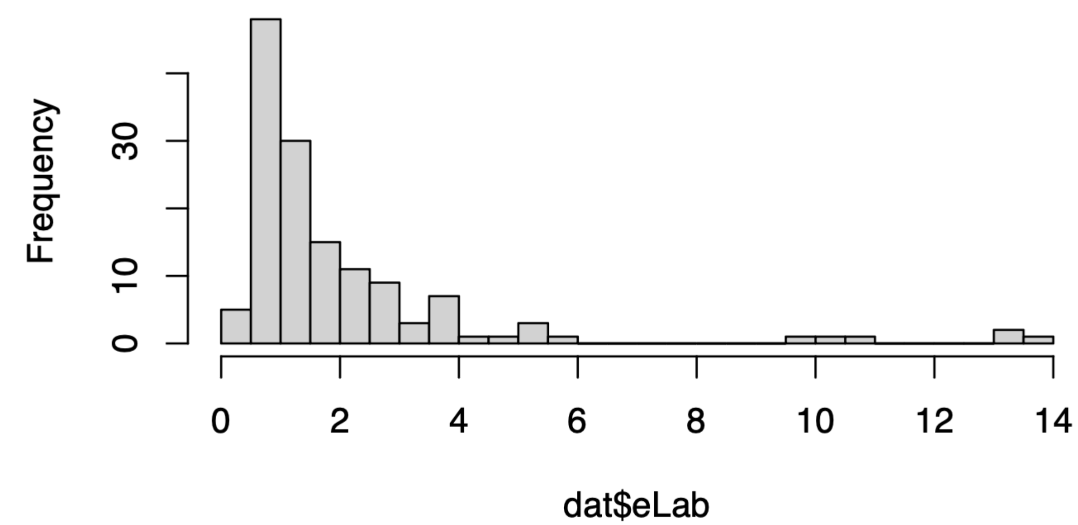
```
colMeans( subset( dat, , c( "eCap", "eLab", "eMat" ) ) )
```

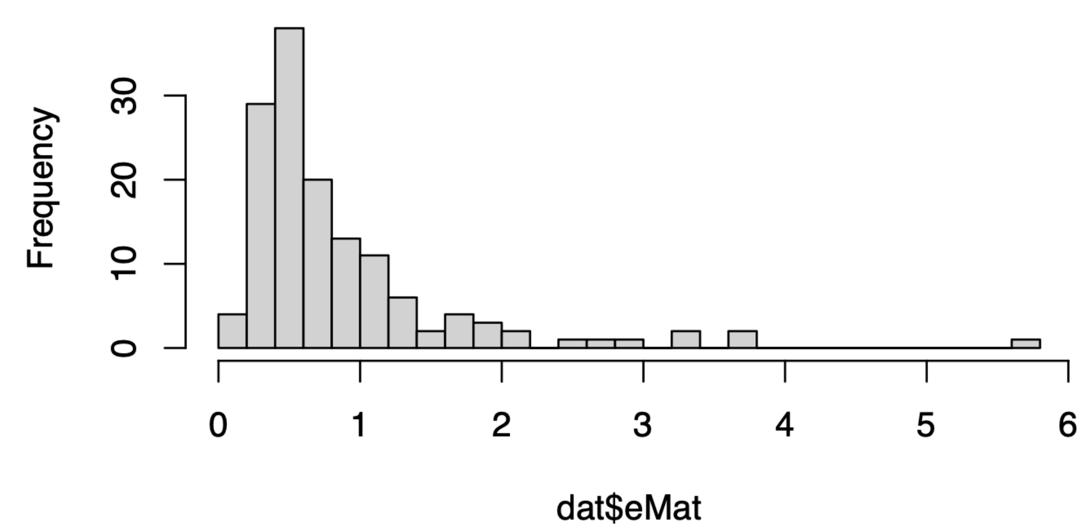
```
     eCap      eLab      eMat 
0.1202721 2.0734793 0.8631936 
```

```
colMedians( subset( dat, , c( "eCap", "eLab", "eMat" ) ) )
```

```
      eCap       eLab       eMat 
0.08063406 1.28627208 0.58741460 
```







Returns to scale

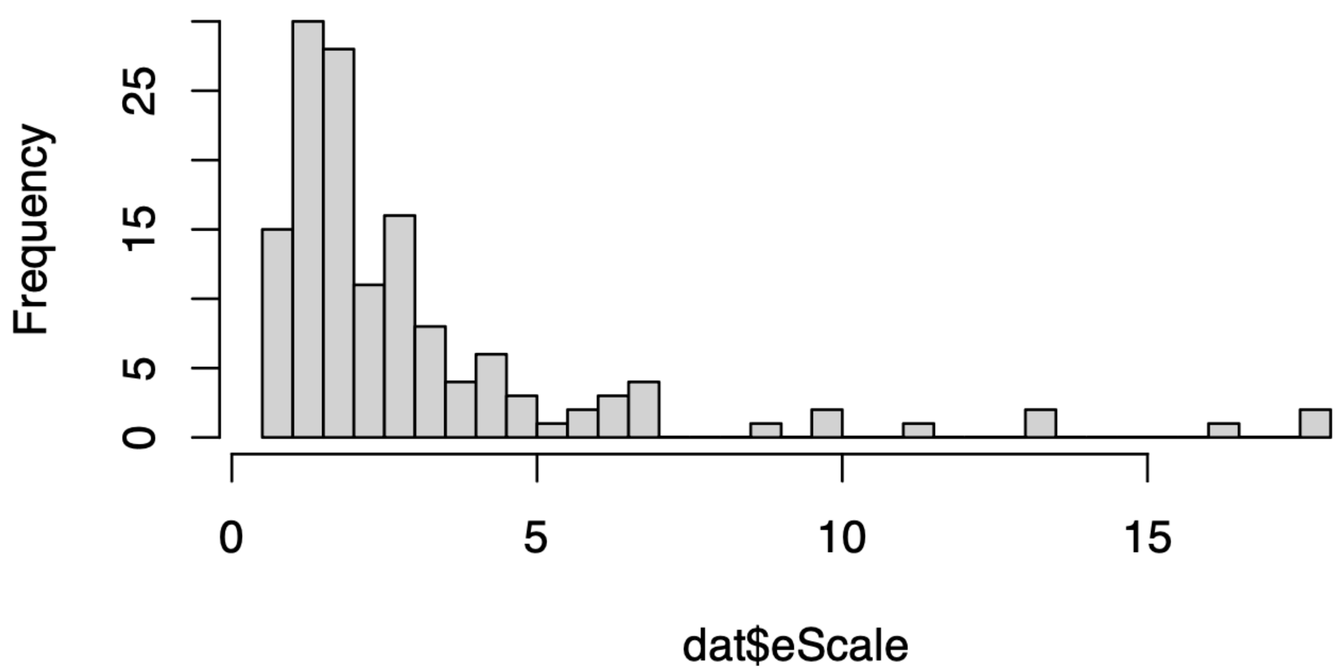
- Recall that adding the output elasticities gives us the elasticity of scale, reflecting the returns to scale.
- Median values: Most firms exhibit increasing returns to scale (IRS).
- **Distribution:** Majority of firms have elasticity of scale between **1 and 2**.
- Outliers: 67 firms show implausibly high IRS (ε>2).

```
colMeans( subset( dat, , c( "eScale" ) ) )
```

```
  eScale
3.056945
```

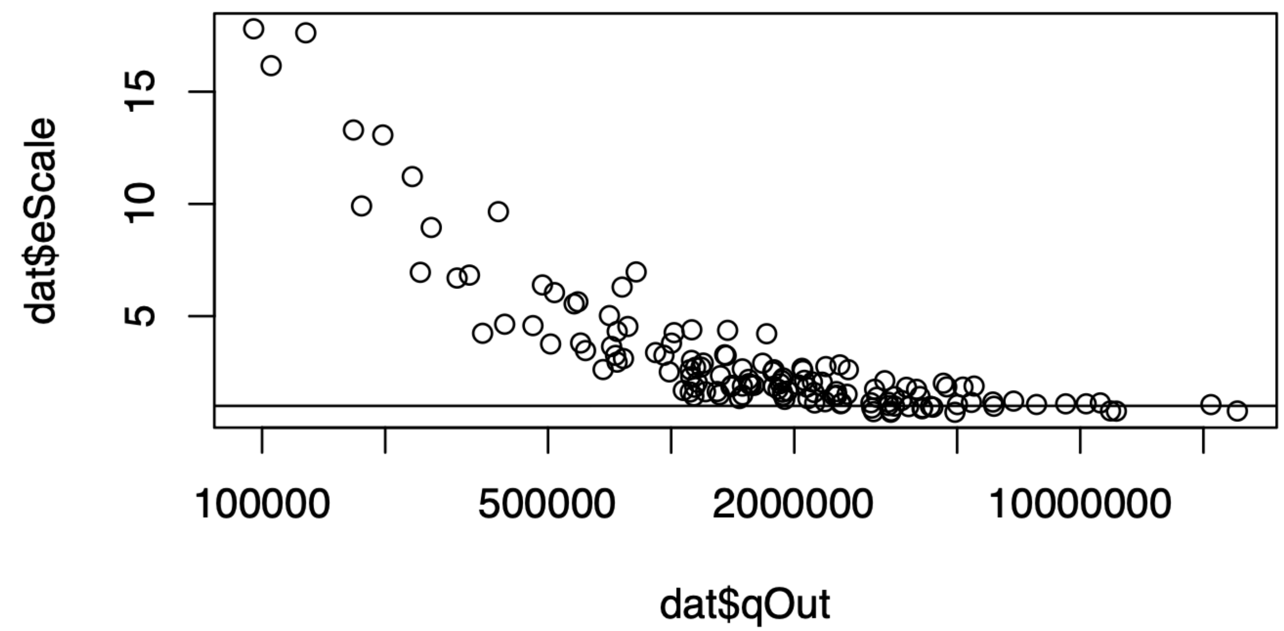
```
colMedians( subset( dat, , c( "eScale" ) ) )
```

```
  eScale
1.941536
```

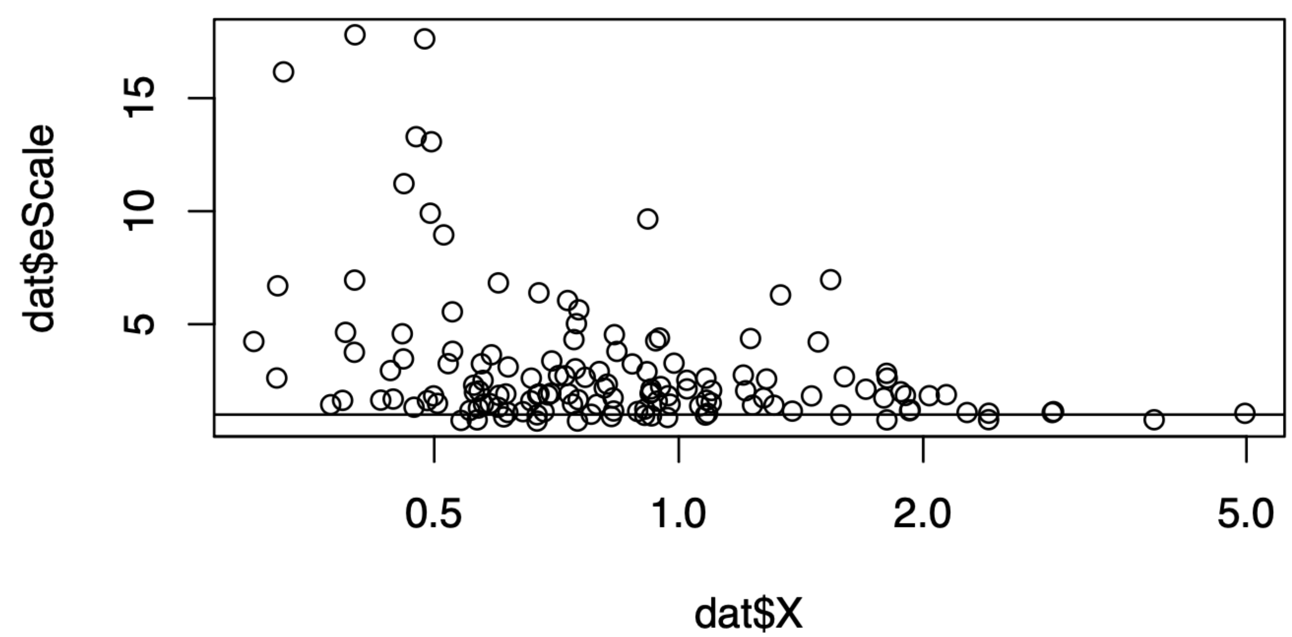


Returns to scale

- Recall from our discussion of factor productivity: Firms with more inputs → higher average productivity (consistent with earlier findings).



- Elasticity of scale vs. size (output, input index):
  - Small firms → typically **increasing returns to scale**.
  - A few firms show **decreasing returns to scale**, both at low and high input levels.



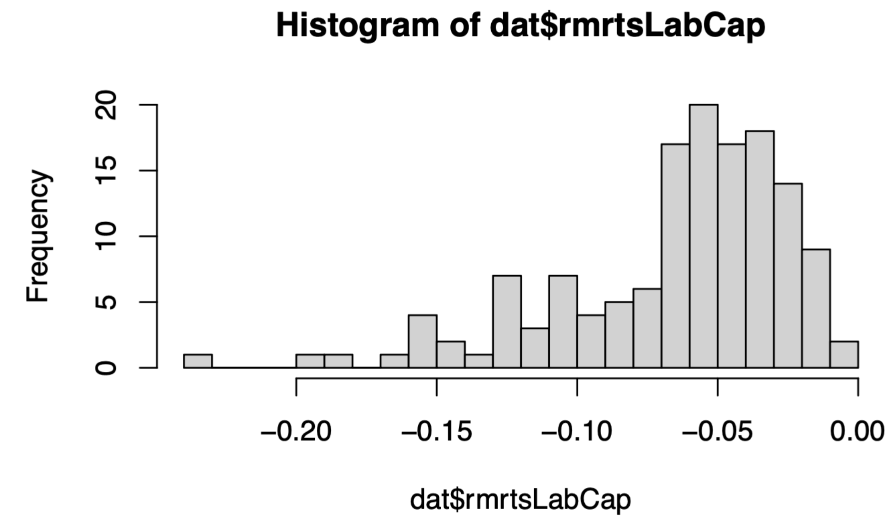
Firm behavior—Efficient employment of inputs

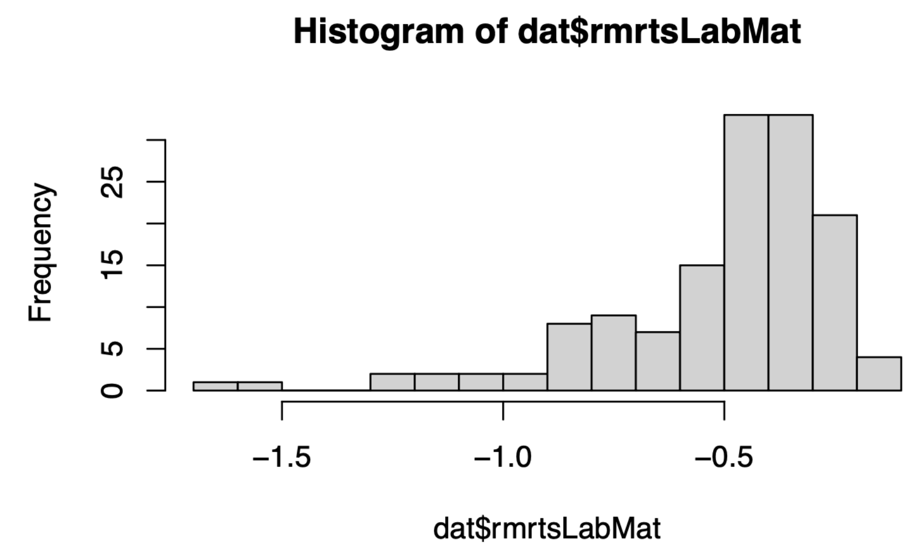
- MRTS vs. RMRTS in a Linear Production Function

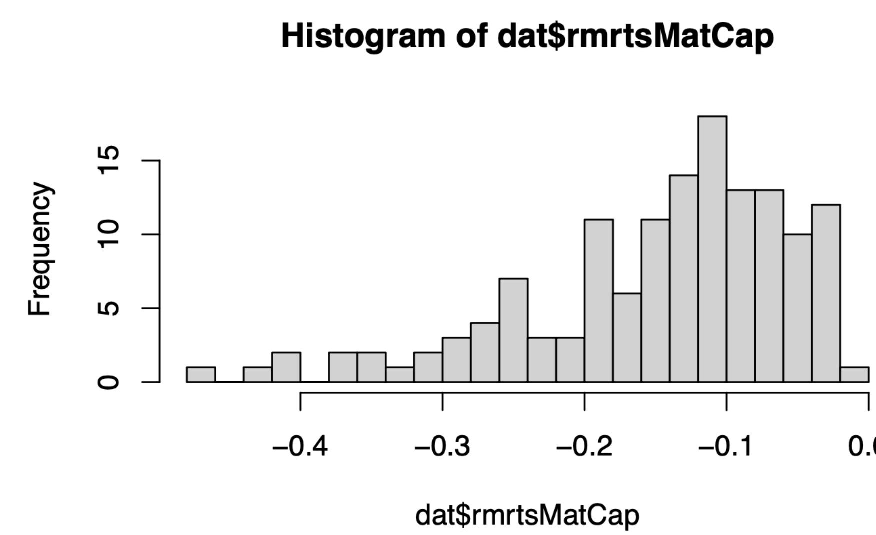
  MRTS = constant, ratio of input coefficients.
  Not meaningful here (depends on input measurement units).

  RMRTS = ratio of output elasticities → **unit-free and more useful**.
  RMRTS varies across firms due to differences in elasticities.

- In our dataset, most firms need about **20% more capital** or **2% more materials** to offset a **1% reduction in labor**.



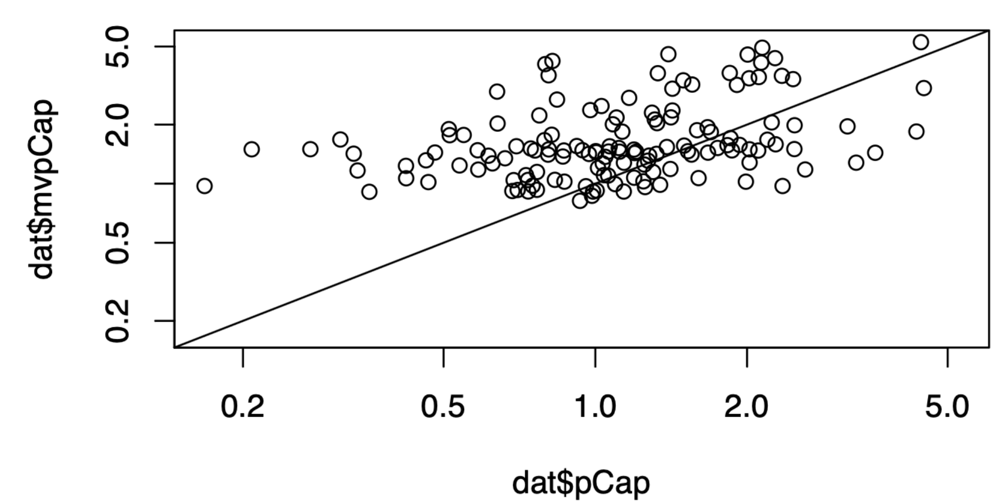




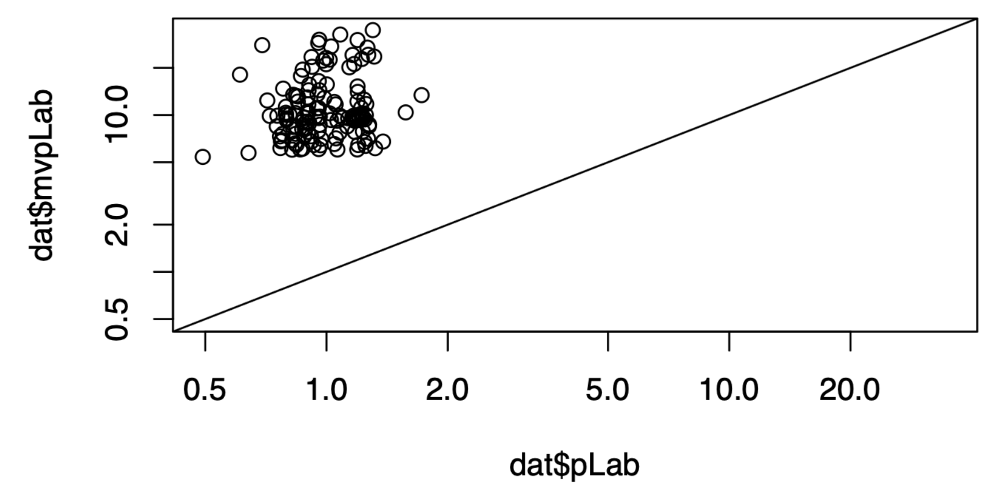
Profit-maximizing behavior

- According to the profit-maximizing principle, the **marginal value product (MVP)** of each input—output price times marginal product—should equal its input price at the optimum.
- Scatter plot (with scaled axes for clarity) shows most firms could raise profits by increasing labor and materials, and some by adding capital —consistent with **increasing returns to scale**.
- Firms may not expand inputs because of **input market imperfections**, where prices don't fully reflect marginal contributions.

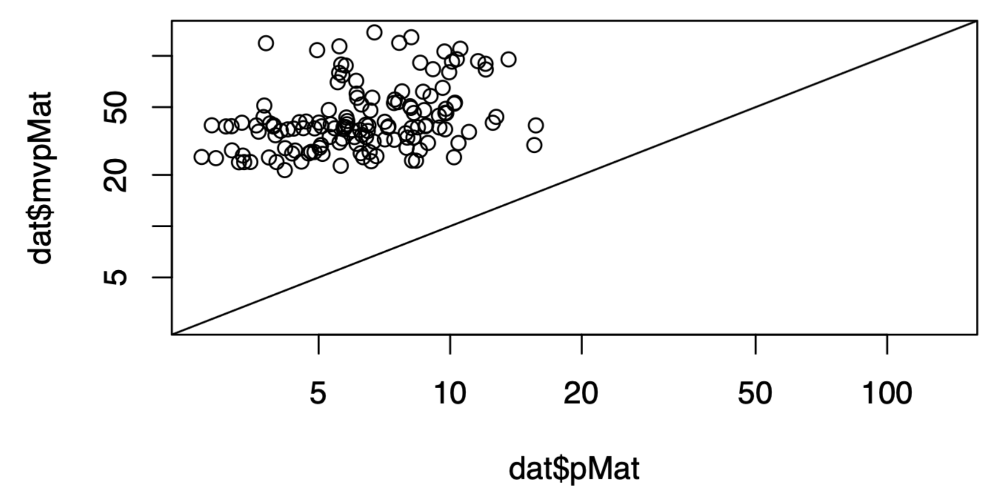
```
compPlot( dat$pCap, dat$mvpCap, log = "xy" )
```




```
compPlot( dat$pLab, dat$mvpLab, log = "xy" )
```

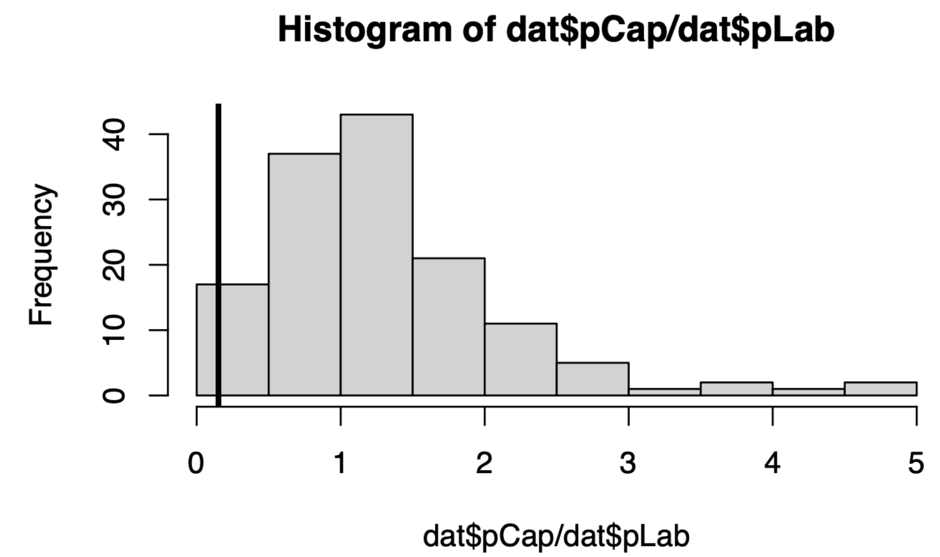


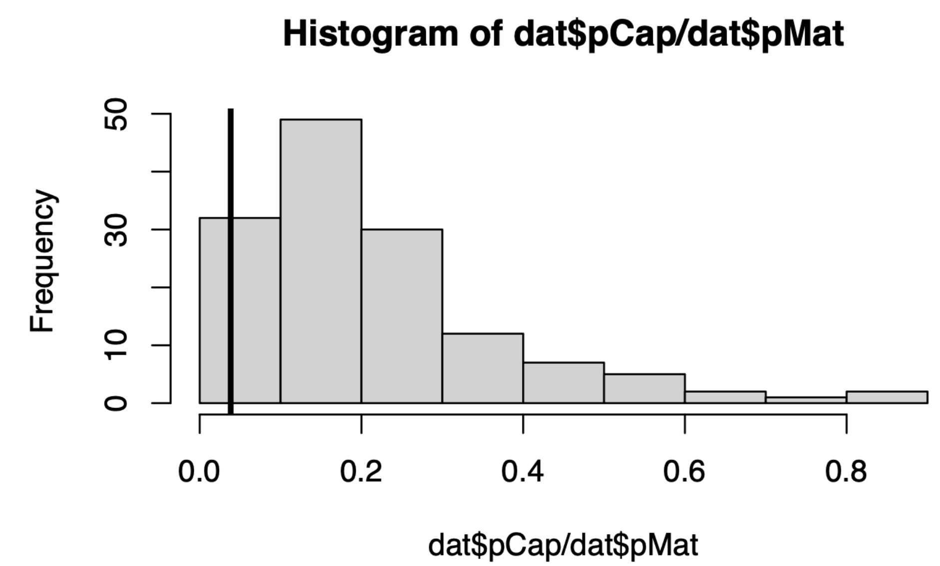

```
compPlot( dat$pMat, dat$mvpMat, log = "xy" )
```

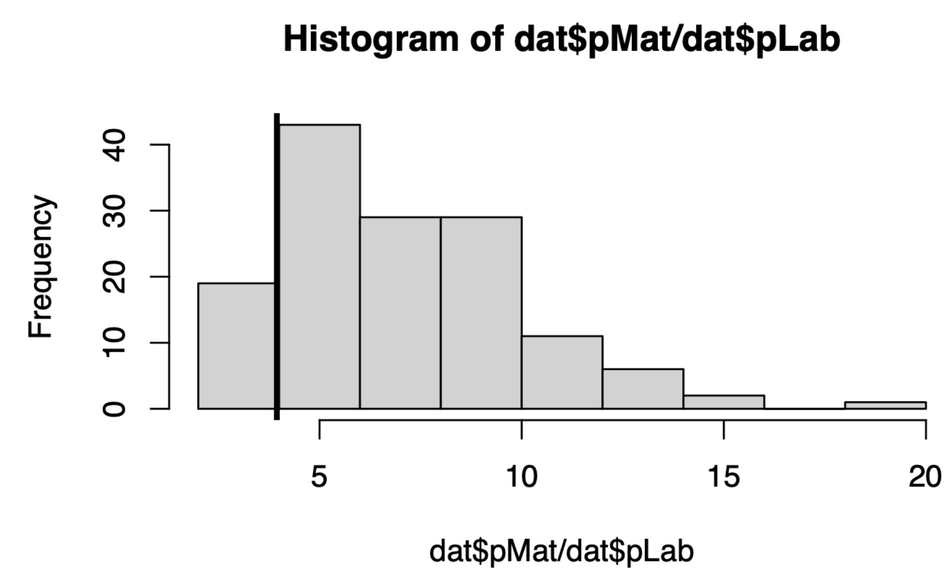


Cost-minimizing behavior

- **Cost-minimizing principle:** Input price ratio = absolute MRTS (constant in linear functions, given by coefficient ratios).
- Histogram of input price ratios vs. MRTS shows most firms could gain by substituting **capital with labor** or **capital with materials**.
- Between labor and materials, most firms would benefit from substituting **materials with labor**—consistent with profit-maximization results.







Cobb-Douglas technology

Cobb-Douglas technology

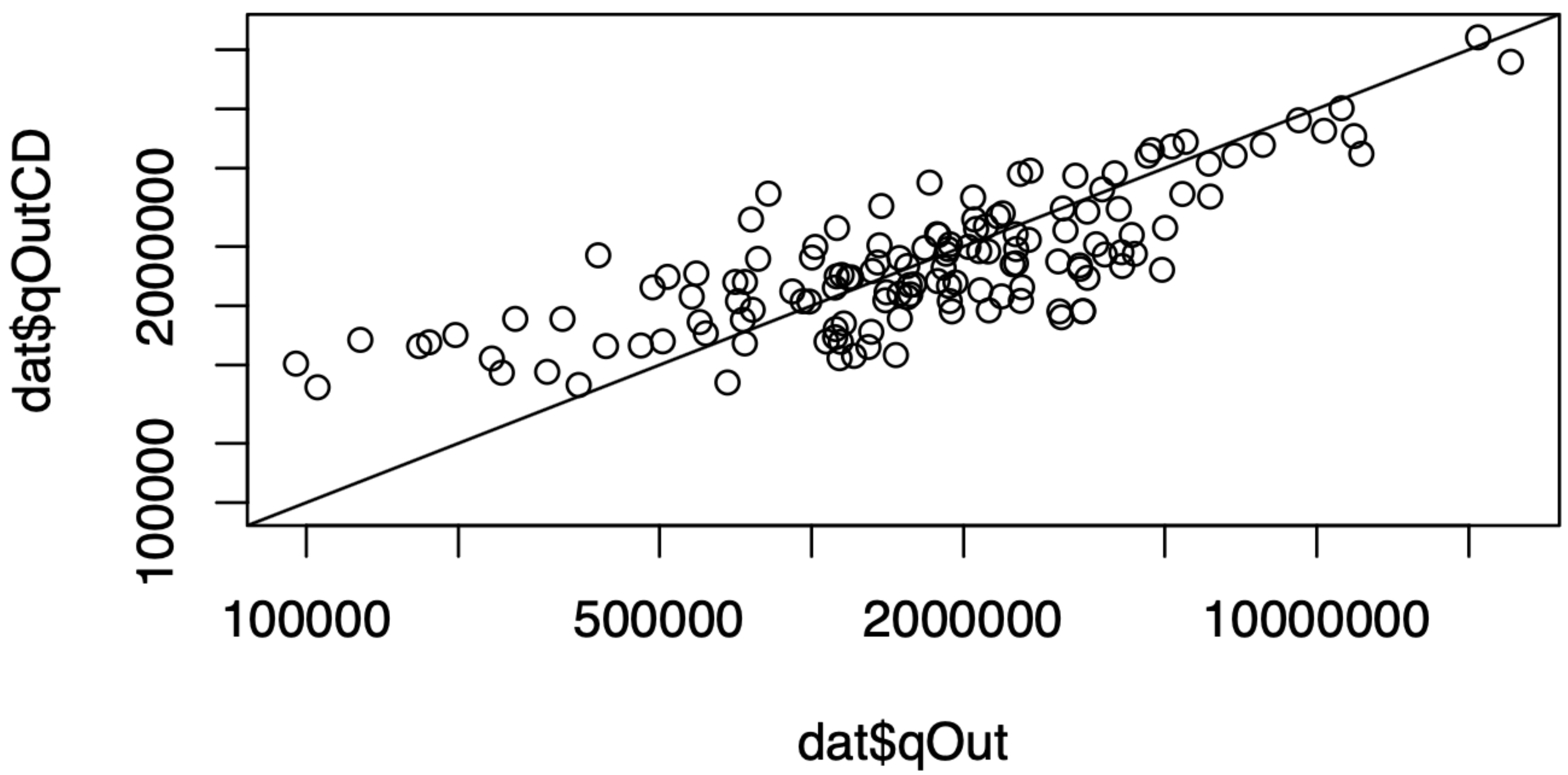
$$y = A \prod_{i=1}^{N} {x_i}^{\alpha_i} \qquad \ln y = \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln x_i$$

```
Call:
lm(formula = log(qOut) ~ log(qCap) + log(qLab) + log(qMat), data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-1.67239 -0.28024  0.00667  0.47834  1.30115

Coefficients:
            Estimate Std. Error t value   Pr(>|t|)
(Intercept) -2.06377    1.31259  -1.572     0.1182
log(qCap)    0.16303    0.08721   1.869     0.0637 .
log(qLab)    0.67622    0.15430   4.383 0.00002327 ***
log(qMat)    0.62720    0.12587   4.983 0.00000187 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.656 on 136 degrees of freedom
Multiple R-squared:  0.5943,    Adjusted R-squared:  0.5854
F-statistic: 66.41 on 3 and 136 DF,  p-value: < 0.0000000000000022
```

- Labor and materials: coefficients for ln(labor) and ln (materials) are positive and significant, while coefficient for ln(capital) is not.
- However, keeping capital ensures **unbiased estimates**, but reduces efficiency (less precise estimates).

Assessing model fit

- R-square is 0.59, but it is not directly comparable to the R-square from a linear model.



- Scatter plot of predicted vs. observed values:

  For **low output (y)**: predicted values are **too high**.
  For **high output (y)**: predicted values are **too low**.
  Suggests the **Cobb–Douglas model is a poor fit**, with systematic bias across ranges of y.

Theoretical consistency of the estimated model

- **Essentiality**

  —Theoretically, all inputs are essential in Cobb-Douglas form (if we drop any, the output becomes zero).
  —Since ln(zero) is undefined, the estimated form trivially satisfies essentiality.

- **Monotonicity**

  —The positive and significant coefficients of labour and materials satisfy the monotonicity condition.

- **Non-negativity**

  —Although the intercept term is negative, $exp(\alpha_0)$ is positive, and so predicted output remains positive.

Productivity and output elasticity

- Cobb–Douglas production function
  - Regression coefficients = elasticities directly.
  - It will be useful to capture firm-level variation in marginal products.

$$MP_i = \frac{\partial y}{\partial x_i} = A\frac{\partial \prod_{i=1}^{N} {x_i}^{\alpha_i}}{\partial x_i} = A\alpha_i x_i^{\alpha_i - 1}\prod_{j\neq i}^{N} {x_j}^{\alpha_j} = \frac{\alpha_i y}{x_i}$$

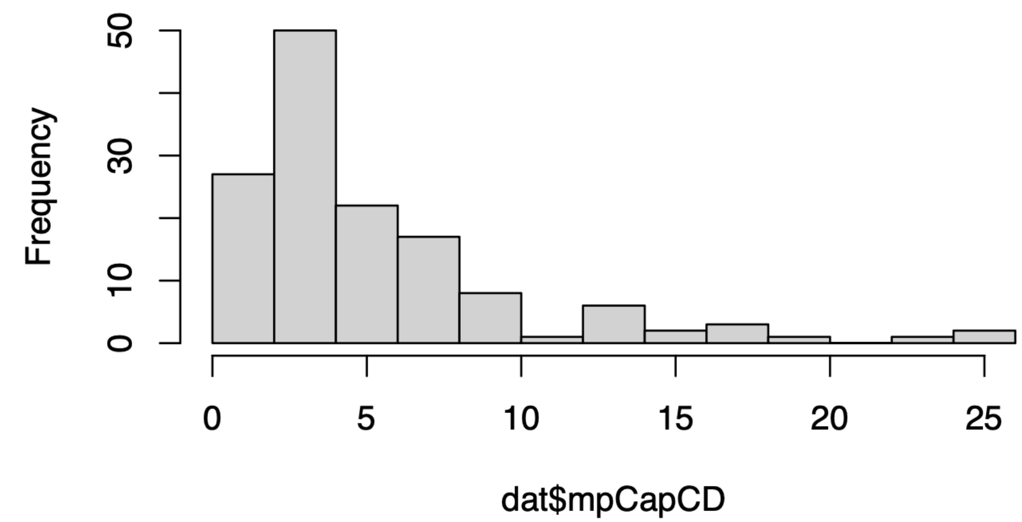
$$\varepsilon_i = \frac{\partial y}{\partial x_i}\cdot\frac{x_i}{y} = \frac{\alpha_i y}{x_i}\cdot\frac{x_i}{y} = \alpha_i$$

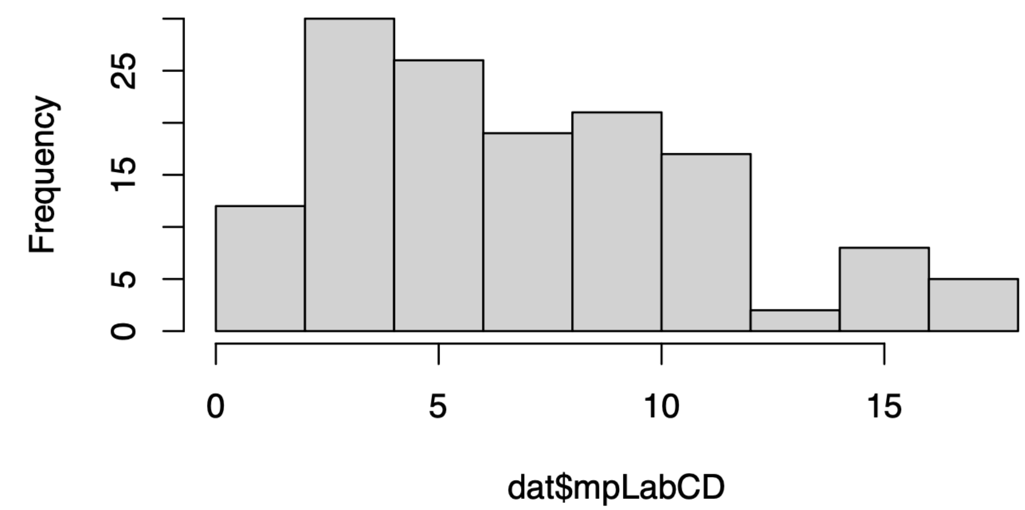
  - The input elasticities equal the linear model coefficients:
    0.16 (capital), 0.67 (labor), 0.62 (materials).

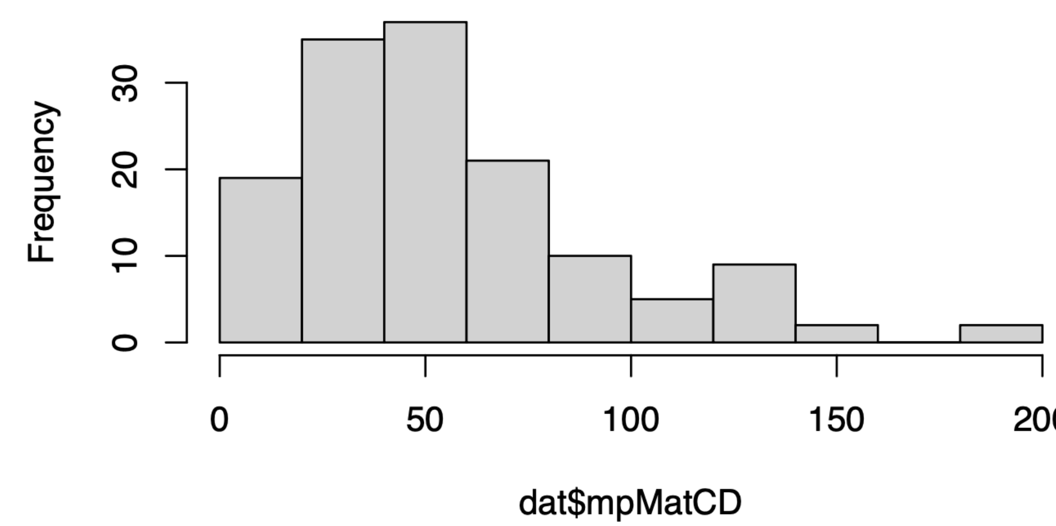
    Example: a **1% increase in labor** raises output by about **0.67%** on average.
  - Compared to the linear model, the Cobb–Douglas estimates show lower elasticities for labor and materials, but a higher elasticity for capital.

Marginal productivity

- **Marginal productivity (MP)** is calculated at each firm's observed output y (can also be done at predicted y).
- **Capital:** +1 unit → output rises by **0–8 units** (most firms).
- **Labor:** +1 unit → output rises by **2–12 units**.
- **Materials:** +1 unit → output rises by **20–80 units**.







Returns to scale

- The elasticity of scale is the **sum of output elasticities**:

$$\varepsilon = \sum_i \alpha_i = 1.466$$

- In the Cobb–Douglas model, elasticity of scale is **constant across all firms**.

- The estimate (ε>1) indicates **increasing returns to scale**, suggesting that firms could benefit from expanding production.

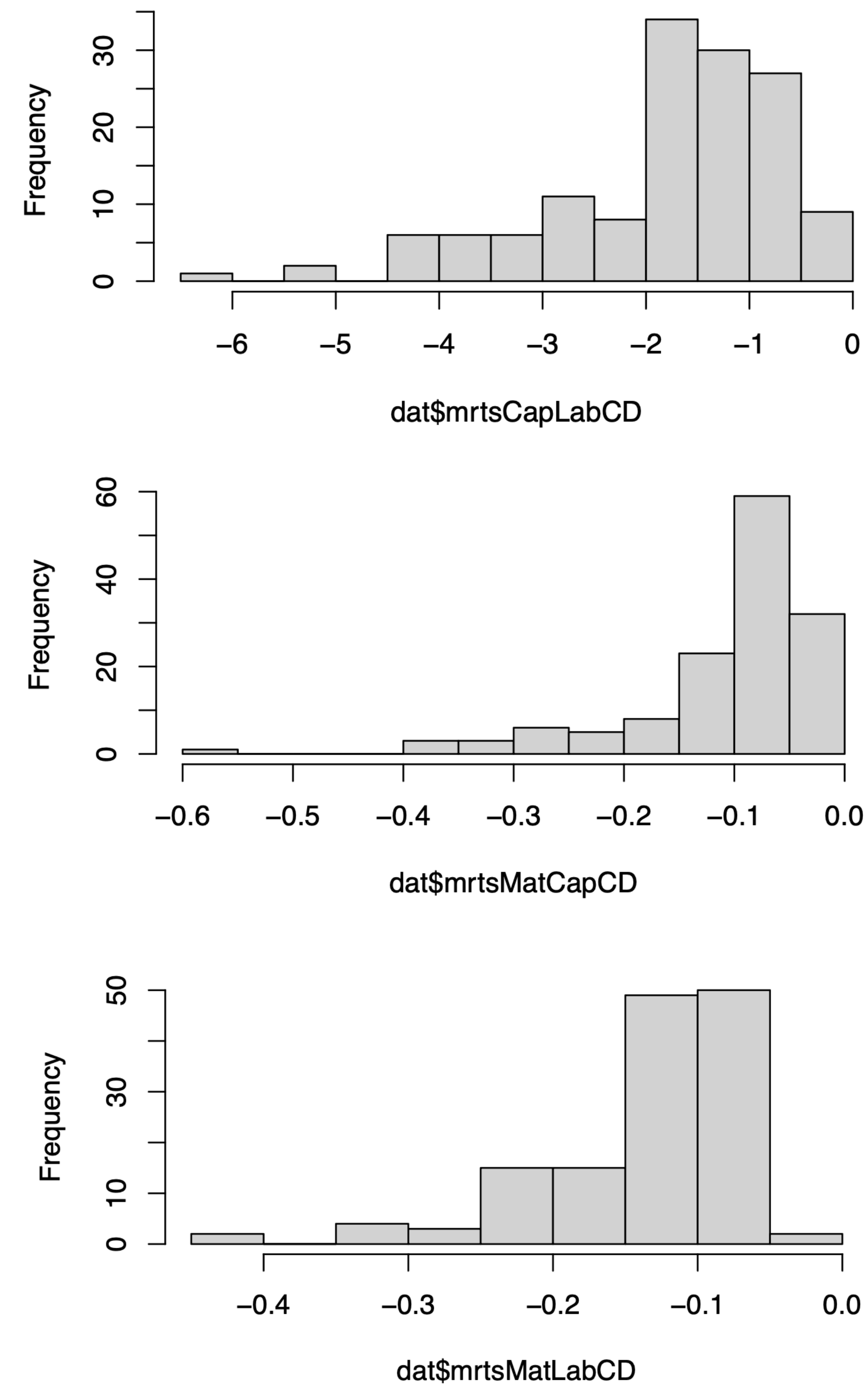
Firm behavior—Efficient employment of inputs

- MRTS vs. RMRTS
  —**MRTS** varies across firms → shows differences in input substitution, but interpretation depends on input units..

  —**RMRTS** is unit-free, but in Cobb–Douglas it is **constant across firms**.

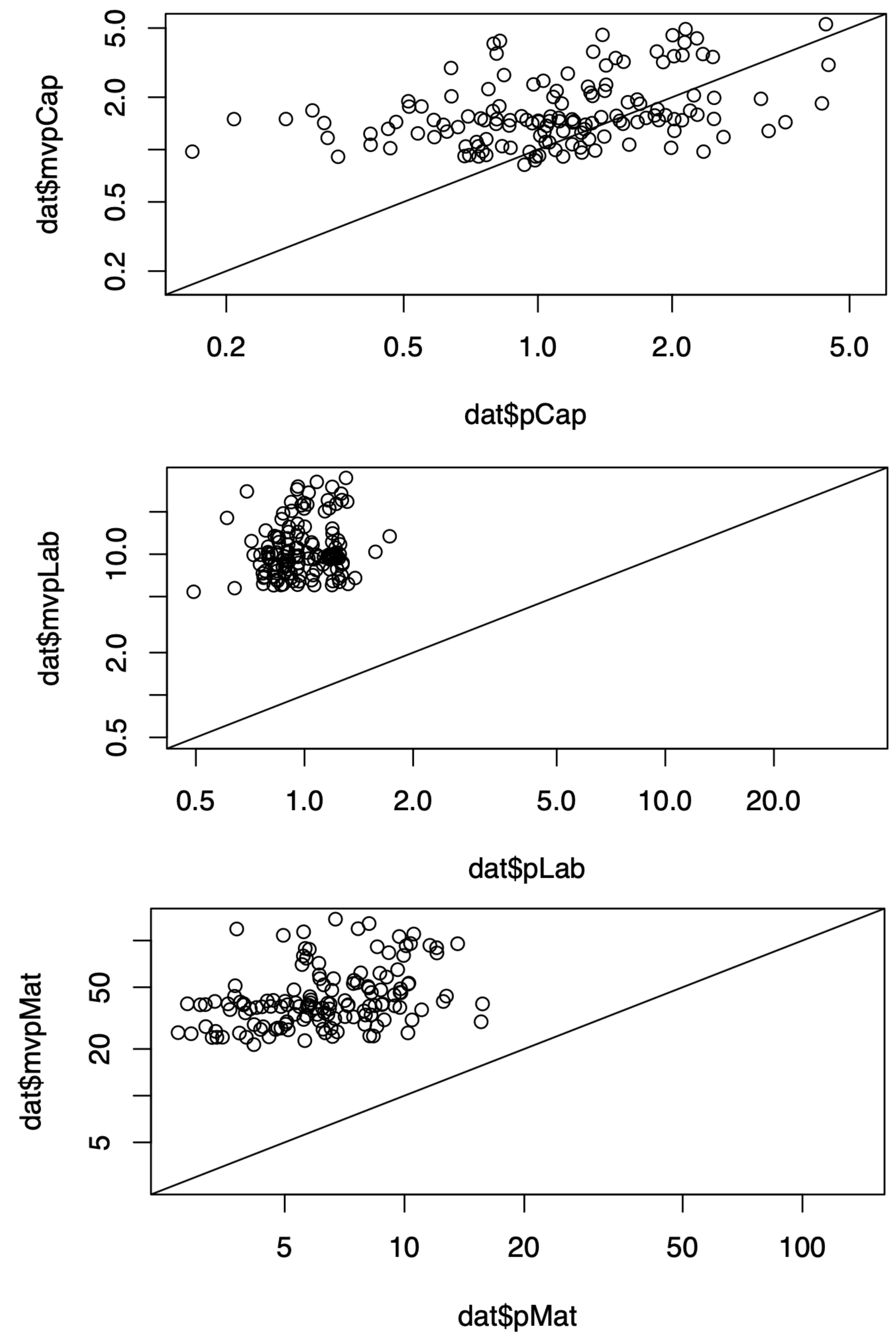
- Results:
  —To replace 1 unit of labor → firms need **0.5–2 units of capital** or **0–0.15 units of materials**.

  —To replace 1% of labor → firms need **+4.15% capital** or **+1.08% materials**.



Profit-maximizing behavior

- At the optimum, marginal value products (MVPs) should equal input prices.
- Plots show MVPs are usually ≥ **input prices** → firms could raise profits by using more inputs.
- This aligns with the **increasing returns to scale** in the Cobb–Douglas model.
- Results are consistent with the **linear production model** findings.
- Firms may not expand input use due to **input market imperfections**.

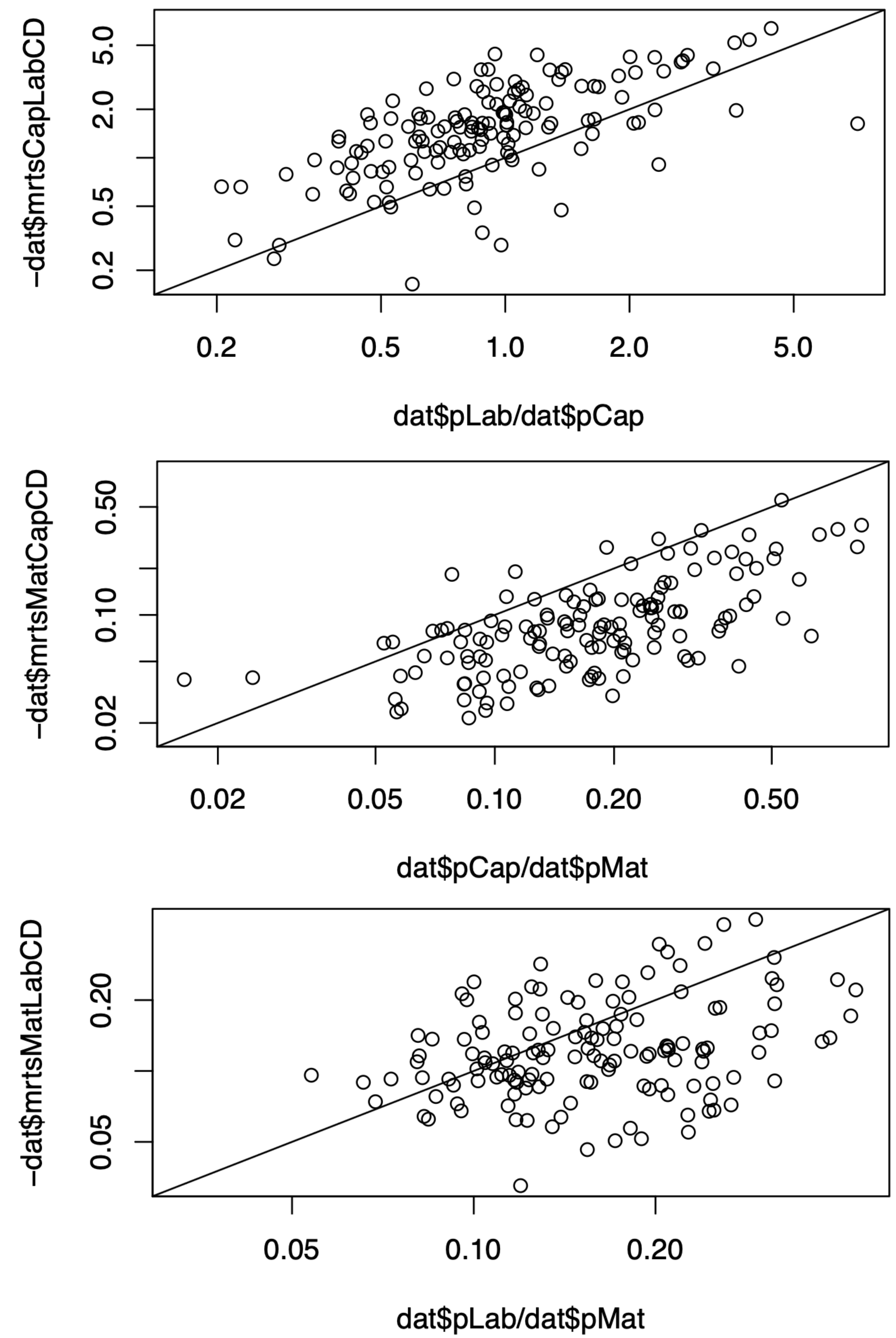


Cost-minimizing behavior

- The cost-minimizing principle suggests that the ratio of input prices must equal the absolute value of the MRTS between two inputs.

- Most firms could benefit from substituting capital with labor, and capital with materials.

- These observations align with our findings based on the profit-maximizing principle.



Model comparison

Comparing production function models

- **Nested models** (e.g., Linear vs. Quadratic, Cobb–Douglas vs. Translog) → can use standard statistical tests.
- Tests often reject Linear in favor of Quadratic.
- Evidence against Cobb–Douglas vs. Translog is weaker.
- **Non-nested models** (Linear vs. Cobb–Douglas) → comparison is more complex.

Model fit: Linear vs. Cobb–Douglas

- Direct R^2 comparison not possible (Linear uses y, Cobb–Douglas uses ln(y).
- Workaround: compute *hypothetical* R^2 by comparing predicted vs. observed values for each model.
- Findings:
  - Similar R^2 when comparing on y.
  - Cobb–Douglas shows much higher R^2 on ln(y).
  - Scatter plots reveal Cobb–Douglas systematically over/underestimates output for many firms.

Model fit: Linear vs. Cobb–Douglas

- Linear model issues:
  - Negative predicted output for one firm. Implausibly high elasticities for labor, materials, and scale.
- Cobb–Douglas issues:
  - Systematic prediction errors in scatter plots.
- **Ramsey RESET test**:

  —A general specification test for regression models. Checks whether the chosen functional form is appropriate.

  –Idea: if the model is correctly specified, adding nonlinear combinations of the fitted values (e.g., y^2, y^3) should not improve the fit.

  —**If significant → model is misspecified** (important variables or nonlinearities missing).
- In our case:
  - Linear model rejected (misspecified). **Cobb–Douglas** passes at 5% level, but marginal at 10%.

Additional models

- Additional specifications—Quadratic and Translog production functions—are estimated in Henningsen (2024), where the Translog production function appears to provide a better fit among the four specifications.

	linear	Cobb-Douglas	quadratic	Translog
R^2 of y	0.79	0.81	0.84	0.77
R^2 of $\ln y$	0.38	0.59	0.55	0.63
visual fit	+	−	−	(+)
RESET (P-value)	0	0.05724	0.00094	0.28127
total monotonicity violations	0	0	41	54
observations with monotonicity violated	0	0	39	48
negative output quantities	1	0	0	0
observations with quasiconcavity violated	0	0	140	77
implausible elasticities of scale	67	0	0	0
implausible output elasticities	124	0	28	56

Reading materials

- Varian, Chapter 2 and 4
- Henningsen, Chapter 1 and 2