Solutions to Problem Set - 1

Quantifying Markets: Estimating Production, Demand, and Competition

Anubhav Jha

October 15, 2025

1 Question 1

1.1 Part A: Theoretical Model

We consider demand $P = \beta_0 - \beta_1 Q$, constant marginal cost $MC(q) = \gamma_{MC,0}$, and a Cournot firm that chooses q taking $\widetilde{Q}$ (other firms' aggregate output) as given, with $Q = q + \widetilde{Q}$.

1. **Variable profit function.** The firm's variable profit (ignoring fixed cost) is

$$\pi(q) = P(Q) \cdot q - \int_0^q MC(s)\, ds = (\beta_0 - \beta_1 Q)\, q - \gamma_{MC,0}\, q,$$

with $Q = q + \widetilde{Q}$.

2. **First-order condition (FOC).** Maximize $\pi(q)$ w.r.t. q:

$$\frac{d\pi}{dq} = \frac{\partial P}{\partial q} q + P(Q) - MC(q) = \frac{\partial P}{\partial q} q + P(Q) - \gamma_{MC,0} = 0.$$

Because $P(Q) = \beta_0 - \beta_1 Q$ and $Q = q + \widetilde{Q}$, we have $\frac{\partial P}{\partial q} = \frac{dP}{dQ} \cdot \frac{dQ}{dq} = -\beta_1 \cdot 1 = -\beta_1$. Thus the FOC becomes

$$-\beta_1 q + P(Q) - \gamma_{MC,0} = 0.$$

3. **Symmetric equilibrium outputs.** By symmetry in the Cournot game all firms choose the same output in equilibrium, so if there are N identical firms then

$$q = \frac{Q}{N}.$$

4. **Solve for q^* and P.** Use the FOC and the inverse demand $P = \beta_0 - \beta_1 Q$. Substitute P into the FOC:

$$-\beta_1 q + (\beta_0 - \beta_1 Q) - \gamma_{MC,0} = 0.$$

Since $Q = Nq$,

$$-\beta_1 q + \beta_0 - \beta_1 (Nq) - \gamma_{MC,0} = 0 \quad \Longrightarrow \quad \beta_0 - \gamma_{MC,0} - \beta_1 (1 + N) q = 0.$$

Hence

$$\boxed{q^* = \frac{\beta_0 - \gamma_{MC,0}}{\beta_1 (1 + N)}}, \tag{1}$$

$$\boxed{P^* \;=\; \beta_0 - \beta_1 N q^* \;=\; \beta_0 - \frac{\beta_1 N(\beta_0 - \gamma_{MC,0})}{\beta_1(1+N)} = \gamma_{MC,0} + \frac{\beta_1(\beta_0 - \gamma_{MC,0})}{\beta_1(1+N)}} \tag{2}$$

(we will use a simpler form for P^* below).

5. **Sign of dq^*/dN.** Differentiate $q^*(N) = \dfrac{\beta_0 - \gamma_{MC,0}}{\beta_1(1+N)}$ w.r.t. N:

$$\frac{dq^*}{dN} = -\frac{\beta_0 - \gamma_{MC,0}}{\beta_1(1+N)^2} < 0$$

(provided $\beta_0 > \gamma_{MC,0}$ so interior solution exists). Thus $dq^*/dN < 0$.

6. **Variable profit at equilibrium.** Variable profit of the firm is

$$\pi_{\text{var}} = Pq - \gamma_{MC,0} q.$$

Use the FOC in the form $P - \gamma_{MC,0} = \beta_1 q$ (recall $\partial P/\partial q = -\beta_1$ so $P - \gamma_{MC,0} = -(\partial P/\partial q)q = \beta_1 q$). Then

$$\pi_{\text{var}} = (P - \gamma_{MC,0})q = (\beta_1 q)\, q = \beta_1 q^2.$$

Therefore

$$\boxed{\pi_{\text{var}} = \beta_1 (q^*)^2}.$$

7. **π_{var} is strictly decreasing in N.** Substitute q^* into π_{var}:

$$\pi_{\text{var}}(N) = \beta_1 \left(\frac{\beta_0 - \gamma_{MC,0}}{\beta_1(1+N)}\right)^2 = \frac{(\beta_0 - \gamma_{MC,0})^2}{\beta_1(1+N)^2}.$$

Differentiate w.r.t. N:

$$\frac{d\pi_{\text{var}}}{dN} = -2\frac{(\beta_0 - \gamma_{MC,0})^2}{\beta_1(1+N)^3} < 0,$$

so variable profit strictly decreases as N increases.

8. **Free entry: equilibrium number of firms.** Let fixed cost be $FC = \gamma_{FC,0}$. Free entry implies firms enter until total profit (variable profit minus fixed cost) is zero:

$$\pi_{\text{var}}(N) - \gamma_{FC,0} = 0.$$

Using the expression for $\pi_{\text{var}}(N)$,

$$\frac{(\beta_0 - \gamma_{MC,0})^2}{\beta_1(1+N)^2} - \gamma_{FC,0} = 0.$$

Solve for N (assume continuous N and take the positive root for $1+N$):

$$(1+N)^2 = \frac{(\beta_0 - \gamma_{MC,0})^2}{\beta_1\, \gamma_{FC,0}} \quad\Longrightarrow\quad 1+N = \frac{\beta_0 - \gamma_{MC,0}}{\sqrt{\beta_1\, \gamma_{FC,0}}}.$$

Thus the equilibrium number of firms is

$$\boxed{N^* \;=\; \frac{\beta_0 - \gamma_{MC,0}}{\sqrt{\beta_1\, \gamma_{FC,0}}} - 1\;.}$$

(Feasibility requires the right-hand side to be nonnegative; otherwise no interior positive entry equilibrium exists.)

Remarks. The expressions above are valid under $\beta_0 > \gamma_{MC,0} > 0$ and $\beta_1 > 0$. The free-entry formula gives an interior positive N^* only if $\beta_0 - \gamma_{MC,0} \geq \sqrt{\beta_1 \gamma_{FC,0}}$.

1.2 Part B: Empirical Model

From Part A, the Cournot model gives the structural relations:

$$q^* = \frac{\beta_0 - \gamma_{MC,0}}{\beta_1(1+N)}, \qquad \pi_{\text{var}} = \frac{(\beta_0 - \gamma_{MC,0})^2}{\beta_1(1+N)^2},$$
$$N^* = \frac{\beta_0 - \gamma_{MC,0}}{\sqrt{\beta_1 \gamma_{FC,0}}} - 1.$$

To estimate these empirically, we express equilibrium relationships in estimable reduced-form equations and instrument potentially endogenous variables.

1.2.1 Data variables

The dataset `df` contains:

`P, Q, N, X_D, X_MC, X_FC`

where `P` = market price, `Q` = aggregate output, `N` = number of active firms, `X_D` = demand shifter, `X_MC` = marginal-cost shifter, and `X_FC` = fixed-cost shifter.

1.2.2 Empirical steps

We compute per-firm output and its square:

$$q_i = \frac{Q_i}{N_i}, \qquad q_i^2 = \left(\frac{Q_i}{N_i}\right)^2.$$

```
df <- df %>%
  mutate(q = Q/N,
         q_sq = (Q/N)^2)
```

Listing 1: Data preparation

1.2.3 (3) Entry equation

From free entry:

$$\pi_{\text{var}} = \gamma_{FC,0} \quad \Rightarrow \quad q^2 = \theta_0 + \theta_1 X_{FC} + \varepsilon.$$

We estimate by OLS:

```
entry_ols <- feols(q_sq ~ X_FC, data = df)
```

Listing 2: Entry OLS regression

Interpretation: $\hat{\theta}_1 < 0$ implies higher fixed costs lead to smaller equilibrium firm sizes.

1.2.4 (4) Cournot condition (pricing equation)

The firm's first-order condition implies a relationship between price and own output:

$$P = \beta_0 - \beta_1 N q + \gamma_{MC,0} + \eta.$$

We estimate both OLS and IV versions:

- Instrument for q using X_{FC} (affects output via entry, not directly price)
- Control for cost shifter X_{MC}

```
cournot_ols <- feols(P ~ q + X_MC, data = df)
cournot_iv  <- feols(P ~ X_MC | q ~ X_FC, data = df)
```

Listing 3: Cournot pricing equations

Interpretation: $\hat{\beta}_1 > 0$ indicates the slope of marginal revenue (or the inverse demand parameter) consistent with Cournot competition.

1.2.5 (5) Market demand estimation

Demand equation:

$$Q = \alpha_0 + \alpha_1 P + \alpha_2 X_D + \nu.$$

To correct for simultaneity between P and Q, we instrument P using both supply-side shifters X_{MC} and X_{FC}:

```
demand_ols <- feols(Q ~ P + X_D, data = df)
demand_iv  <- feols(Q ~ X_D | P ~ X_MC + X_FC, data = df)
```

Listing 4: Demand estimation

1.2.6 (6) Consolidated results table

Finally, we present all results in a single regression summary table:

```
etable(
  entry_ols,
  cournot_ols,
  cournot_iv,
  demand_ols,
  demand_iv,
  tex=TRUE,
  headers = c(
    "Entry␣OLS",
    "Cournot␣OLS␣(IV:␣X_FC)",
    "Cournot␣IV␣(IV:␣X_FC)",
    "Demand␣OLS",
    "Demand␣(IV:␣X_MC␣+␣X_FC)"
  )
)
```

Listing 5: Export regression results

Table 1: Q1B: Estimates for Cement Example

Dependent Variables:	q_sq	P		Q	
	Entry OLS	Cournot OLS (IV: X_FC)	Cournot IV (IV: X_FC)	Demand OLS	Demand (IV: X_MC)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	0.6200***	-13.15***	0.8693	40.84***	48.72***
	(0.1128)	(0.3985)	(1.074)	(0.2821)	(0.9408)
X_FC	0.2379***				
	(0.0111)				
q		10.04***	1.755***		
		(0.1648)	(0.5819)		
X_MC		0.5447***	0.5517***		
		(0.0280)	(0.0423)		
P				0.4798***	-0.3499***
				(0.0117)	(0.0848)
X_D				2.001***	1.989***
				(0.0253)	(0.0474)
Fit statistics					
Observations	1,991	1,991	1,991	1,991	1,991
R^2	0.18910	0.67366	0.25839	0.79897	0.29270
Adjusted R^2	0.18870	0.67333	0.25764	0.79877	0.29199

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

1.2.7 Interpretation Summary

(1) Entry Equation. Column (1) shows that X_{FC} has a positive and highly significant effect on q^2. This implies that markets with higher values of the fixed-cost shifter support larger firm sizes, consistent with lower effective entry costs or better operating conditions.

(2) Cournot Pricing. Columns (2)–(3) estimate the Cournot pricing condition. In OLS, q has a large positive coefficient (10.04), but this likely reflects endogeneity with unobserved demand. When q is instrumented by X_{FC}, the coefficient drops to 1.76 and remains significant, aligning with theoretical predictions. The positive effect of X_{MC} confirms that higher marginal costs increase equilibrium prices.

(3) Demand. In Column (4), the OLS estimate of price on Q is positive, violating the law of demand. After instrumenting price with X_{MC} and X_{FC} (Column 5), the coefficient becomes negative (-0.35), restoring the expected downward-sloping demand. The demand shifter X_D remains positive and significant in both models.

(4) Overall. Overall, the results are consistent with the Cournot entry framework: cost shifters influence firm size and pricing as predicted, and the IV estimates correct the simultaneity bias in both pricing and demand equations.

2 Question 2

2.1 Part A: Solution (Cost–Minimization Model)

Let the production technology be

$$Y = A\,K^{\alpha_K} L^{\alpha_L} M^{\alpha_M}, \qquad \alpha \equiv \alpha_K + \alpha_L + \alpha_M > 0.$$

1) Cost minimization problem. The firm minimizes total expenditure subject to producing output Y:

$$\min_{K,L,M} \; C = W_K K + W_L L + W_M M \quad \text{s.t.} \quad AK^{\alpha_K} L^{\alpha_L} M^{\alpha_M} = Y.$$

2) Lagrangian. Introduce multiplier λ:

$$\mathcal{L} = W_K K + W_L L + W_M M + \lambda\big(Y - AK^{\alpha_K} L^{\alpha_L} M^{\alpha_M}\big).$$

3) First-order conditions (FOCs). Differentiate w.r.t. each input and set equal to zero:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial K} &: \quad W_K - \lambda A \alpha_K K^{\alpha_K - 1} L^{\alpha_L} M^{\alpha_M} = 0, \\
\frac{\partial \mathcal{L}}{\partial L} &: \quad W_L - \lambda A \alpha_L K^{\alpha_K} L^{\alpha_L - 1} M^{\alpha_M} = 0, \\
\frac{\partial \mathcal{L}}{\partial M} &: \quad W_M - \lambda A \alpha_M K^{\alpha_K} L^{\alpha_L} M^{\alpha_M - 1} = 0.
\end{aligned}$$

Rearrange each to the form

$$W_K = \lambda \alpha_K \frac{AK^{\alpha_K} L^{\alpha_L} M^{\alpha_M}}{K} = \lambda \alpha_K \frac{Y}{K},$$

and analogously for L and M.

4) Solve FOCs for K, L, M. From the FOCs we get

$$K = \lambda \frac{\alpha_K Y}{W_K}, \qquad L = \lambda \frac{\alpha_L Y}{W_L}, \qquad M = \lambda \frac{\alpha_M Y}{W_M}.$$

(These relations also imply that expenditure shares equal the Cobb–Douglas exponent shares: $W_K K / C = \alpha_K / \alpha$, etc.)

5) Determine λ using the production constraint. Plug the expressions for K, L, M into the production constraint:

$$A\left(\lambda\frac{\alpha_K Y}{W_K}\right)^{\alpha_K}\left(\lambda\frac{\alpha_L Y}{W_L}\right)^{\alpha_L}\left(\lambda\frac{\alpha_M Y}{W_M}\right)^{\alpha_M} = Y.$$

Collect terms:

$$A\lambda^{\alpha}Y^{\alpha}\Big(\prod_{i\in\{K,L,M\}}(\alpha_i)^{\alpha_i}\Big)\Big(\prod_{i\in\{K,L,M\}}W_i^{-\alpha_i}\Big) = Y.$$

Solve for λ:

$$\lambda^{\alpha} = \frac{Y^{1-\alpha}}{A}\cdot\left(\prod_i \alpha_i^{-\alpha_i}\right)\left(\prod_i W_i^{\alpha_i}\right),$$

so

$$\lambda = Y^{\frac{1-\alpha}{\alpha}}\ A^{-1/\alpha}\left(\prod_i \alpha_i^{-\alpha_i}\right)^{1/\alpha}\left(\prod_i W_i^{\alpha_i}\right)^{1/\alpha}.$$

6) Input demand functions $K(Y,W), L(Y,W), M(Y,W)$. Using $K = \lambda\alpha_K Y/W_K$ (and analogously for L, M) we obtain compactly:

$$K(Y,W) = \frac{\alpha_K}{\alpha}\frac{C(Y,W)}{W_K},\quad L(Y,W) = \frac{\alpha_L}{\alpha}\frac{C(Y,W)}{W_L},\quad M(Y,W) = \frac{\alpha_M}{\alpha}\frac{C(Y,W)}{W_M},$$

where $C(Y,W)$ is the (unknown yet) cost function. Equivalently, substituting λ above yields explicit power-form solutions:

$$K(Y,W) = Y^{1/\alpha}\ A^{-1/\alpha}\left(\prod_i \alpha_i^{-\alpha_i}\right)^{-1/\alpha}\alpha_K\ W_K^{\frac{\alpha_K}{\alpha}-1}\ W_L^{\frac{\alpha_L}{\alpha}}\ W_M^{\frac{\alpha_M}{\alpha}},$$

(and analogous expressions for $L(Y,W), M(Y,W)$).

7) Cost function $C(Y,W)$. Plug the input demands K, L, M written as $(\alpha_i/\alpha)C/W_i$ into the production constraint to solve for C:

$$A\prod_i\left(\frac{\alpha_i}{\alpha}\frac{C}{W_i}\right)^{\alpha_i} = Y \quad\Longrightarrow\quad C^{\alpha} = Y\ A^{-1}\prod_i\left(\frac{\alpha_i}{\alpha}\right)^{-\alpha_i}\prod_i W_i^{\alpha_i}.$$

Therefore the cost function is

$$\boxed{C(Y,W) = Y^{1/\alpha}\ A^{-1/\alpha}\left(\prod_{i\in\{K,L,M\}}\left(\frac{\alpha_i}{\alpha}\right)^{-\alpha_i}\right)^{1/\alpha}\prod_{i\in\{K,L,M\}}W_i^{\alpha_i/\alpha}}.$$

This expression is the familiar homothetic (power) cost function for Cobb–Douglas technology. Suppose the constant κ collects all parameters and input-price terms that do not depend on output Y. From the full expression of $C(Y,W)$ we have:

$$\boxed{\kappa = A^{-1/\alpha}\left(\prod_{i\in\{K,L,M\}}\left(\frac{\alpha_i}{\alpha}\right)^{-\alpha_i}\right)^{1/\alpha}\prod_{i\in\{K,L,M\}}W_i^{\alpha_i/\alpha}.}$$

and $C(Y,W) = \kappa\, Y^{1/\alpha}$.

Marginal cost and curvature. Differentiate $C(Y, W) = \kappa\, Y^{1/\alpha}$ (where κ is the constant factor above) to obtain

$$MC(Y) = \frac{dC}{dY} = \kappa \frac{1}{\alpha} Y^{1/\alpha - 1} = \frac{1}{\alpha} \frac{C(Y, W)}{Y}.$$

The second derivative is

$$C''(Y) = \kappa \frac{1}{\alpha}\Big(\frac{1}{\alpha} - 1\Big) Y^{1/\alpha - 2} = \frac{C(Y, W)}{Y^2} \cdot \frac{1}{\alpha}\Big(\frac{1}{\alpha} - 1\Big).$$

Hence the sign of $C''(Y)$ depends only on $\alpha = \alpha_K + \alpha_L + \alpha_M$:

- If $\alpha < 1$ (decreasing returns to scale) then $\frac{1}{\alpha} - 1 > 0$, so $C''(Y) > 0$: *cost is convex* and marginal cost is increasing in Y.
- If $\alpha = 1$ (constant returns) then $C(Y)$ is linear in Y, $C''(Y) = 0$ and marginal cost is constant.
- If $\alpha > 1$ (increasing returns) then $\frac{1}{\alpha} - 1 < 0$, so $C''(Y) < 0$: *cost is concave* and marginal cost is decreasing in Y.

Remarks. The derivation above used (i) the Cobb–Douglas functional form and (ii) interior solutions. The compact expressions using expenditure shares, $W_i x_i / C = \alpha_i / \alpha$, are often the easiest way to obtain the conditional demands and the cost function for Cobb–Douglas technology.

2.2 Part B: Empirical Implementation

2.2.1 Model Setup

We consider the production and revenue equations:

$$\log(Y_{it}) = \alpha_K \log K_{it} + \alpha_L \log L_{it} + \alpha_M \log M_{it} + \omega_{it} + \varepsilon_{it},$$
$$\log(R_{it}) = \beta_K \log K_{it} + \beta_L \log L_{it} + \beta_M \log M_{it} + \omega_{it} + \nu_{it},$$

where ω_{it} represents firm productivity and $P_{\text{industry},t}$ is the industry-level price index. We estimate several versions of these equations depending on whether ω_{it} and $P_{\text{industry},t}$ are observed.

2.2.2 (1) Data preparation and summary

We begin by inspecting the data and creating log-transformed variables. This prepares the dataset for log-linear estimation of the Cobb–Douglas production and revenue functions.

We begin by computing detailed summary statistics (mean, standard deviation, quartiles, and range) for all main variables. The code below uses a robust summarization approach with a double underscore separator and formats the output using `kable()` for LaTeX.

```r
library(dplyr)
library(tidyr)
library(knitr)

# Define variables to include in summary
vars_to_keep <- c("Y", "R", "K", "L", "M", "omega", "P_industry")

# Compute summary stats with clear column naming
summary_stats <- df %>%
  select(all_of(vars_to_keep)) %>%
  summarise(across(
    everything(),
    list(
      Mean   = ~mean(., na.rm = TRUE),
      StdDev = ~sd(., na.rm = TRUE),
      Median = ~median(., na.rm = TRUE),
```

```
      P25    = ~quantile(., 0.25, na.rm = TRUE),
      P75    = ~quantile(., 0.75, na.rm = TRUE),
      Min    = ~min(., na.rm = TRUE),
      Max    = ~max(., na.rm = TRUE),
      N      = ~sum(!is.na(.))
    ),
    .names = "{.col}__{.fn}"   # double underscore for safe reshaping
  ))

# Reshape into long format for a clean table
stats_table <- summary_stats %>%
  pivot_longer(
    everything(),
    names_to = c("Variable", ".value"),
    names_sep = "__"
  )

# Display formatted summary table
kable(
  stats_table,
  digits = 3,
  caption = "Summary␣Statistics␣for␣Main␣Variables",
  format = "latex",
  booktabs = TRUE
)
```

Listing 6: Comprehensive summary statistics using knitr::kable

Table 2: Summary Statistics for Main Variables

Variable	Mean	StdDev	Median	P25	P75	Min	Max	N
Y	1.499	1.187	1.193	0.744	1.873	0.076	18.774	4800
R	2.044	2.675	1.170	0.551	2.442	0.024	40.721	4800
K	2.564	5.486	1.035	0.382	2.652	0.005	120.489	4800
L	2.601	5.229	1.038	0.384	2.618	0.008	90.931	4800
M	2.583	5.139	1.009	0.377	2.653	0.007	108.803	4800
omega	0.176	1.023	0.101	-0.563	0.909	-2.604	3.333	4800
P_industry	1.324	1.044	0.915	0.631	1.846	0.214	4.955	4800

2.2.3 (2) Production function with observed productivity

Assuming productivity ω_{it} is observed, we can estimate the true production function:

$$\log Y_{it} = \alpha_K \log K_{it} + \alpha_L \log L_{it} + \alpha_M \log M_{it} + \omega_{it} + \varepsilon_{it}.$$

Including ω_{it} directly yields consistent estimates of the input elasticities.

```
ols_true <- feols(lY ~ lK + lL + lM + omega, data = df)
```

Listing 7: Production function with observed ω_{it}

2.2.4 (3) Revenue function with observed productivity

Next, we estimate the log revenue function when ω_{it} is observable:

$$\log R_{it} = \beta_K \log K_{it} + \beta_L \log L_{it} + \beta_M \log M_{it} + \omega_{it} + \nu_{it}.$$

This version allows us to separate the effects of input quantities and productivity on firm revenue.

```r
rev_w_omega <- feols(lR ~ lK + lL + lM + omega, data = df)
```

Listing 8: Revenue regression with observed ω_{it}

2.2.5 (4) Accounting for industry prices

If the industry price $P_{\text{industry},t}$ is unobserved, we can absorb industry-level heterogeneity using fixed effects. Alternatively, if it is observed, we include it directly in the regression.

```r
# (i) Unknown P_industry,t: control via industry fixed effects
rev_timeFE <- feols(lR ~ lK + lL + lM + omega | factor(industry), data = df)

# (ii) Observed P_industry,t: include log price directly
rev_with_Pind <- feols(lR ~ lK + lL + lM + omega + lP_industry, data = df)
```

Listing 9: Revenue regressions with/without industry price

When $P_{\text{industry},t}$ is unobserved, the fixed effects absorb time-invariant differences across industries. When observed, its inclusion captures time-varying price effects on revenues.

2.2.6 (5) Re-estimation without observed productivity

When ω_{it} is unobserved, OLS estimates are biased if productivity is correlated with input choices. We therefore re-estimate all previous models omitting ω_{it} to quantify this bias.

```r
# (a) Production function without productivity
ols_no_omega <- feols(lY ~ lK + lL + lM, data = df)

# (b) Revenue regressions without productivity
rev_no_omega_timeFE <- feols(lR ~ lK + lL + lM | factor(industry), data = df)
rev_no_omega_Pind   <- feols(lR ~ lK + lL + lM + lP_industry, data = df)
```

Listing 10: Models without observed ω_{it}

The comparison between these and the previous specifications reveals the extent of omitted-variable bias due to unobserved productivity.

2.2.7 (6) Combined results table

All specifications are summarized in one comparison table for ease of interpretation.

```r
etable(
  ols_true,                # (2) Production
  rev_w_omega,             # (3) Revenue
  rev_timeFE,              # (4.i) Revenue (unknown P_industry)
  rev_with_Pind,           # (4.ii) Revenue (observed P_industry)
  ols_no_omega,            # (5) Production
  rev_no_omega_timeFE,     # (5) Revenue
  rev_no_omega_Pind,       # (5) Revenue
  tex = TRUE
)
```

Listing 11: Combined regression results for Question 2 Part B

Overall Conclusions. The results are consistent with the theoretical expectations of the production and revenue models.

Including firm productivity (ω_{it}) yields economically meaningful input elasticities that sum close to one, confirming approximate constant returns to scale.

When ω_{it} is omitted, input coefficients fall sharply, indicating strong downward bias from unobserved productivity. Controlling for industry prices or fixed effects restores explanatory power and more plausible

Table 3: Q2B Regression Tables

Dependent Variables:	lY	lR	lR	lR	lY	lR	lR
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Constant	-0.0034	-0.0929***		-0.0051	0.1705***		0.1725***
	(0.0046)	(0.0120)		(0.0055)	(0.0086)		(0.0088)
lK	0.2915***	0.2900***	0.2937***	0.2931***	0.0379***	0.1511***	0.0588***
	(0.0043)	(0.0114)	(0.0067)	(0.0051)	(0.0074)	(0.0082)	(0.0077)
lL	0.4894***	0.4925***	0.4856***	0.4854***	0.2345***	0.3508***	0.2482***
	(0.0043)	(0.0114)	(0.0046)	(0.0052)	(0.0074)	(0.0056)	(0.0077)
lM	0.2059***	0.1913***	0.2052***	0.2060***	-0.0419***	0.0695***	-0.0206***
	(0.0043)	(0.0114)	(0.0047)	(0.0052)	(0.0075)	(0.0065)	(0.0078)
omega	0.9868***	1.366***	0.9888***	0.9856***			
	(0.0083)	(0.0220)	(0.0123)	(0.0103)			
lP_industry				1.010***			1.201***
				(0.0074)			(0.0121)
Fixed-effects							
factor(industry)			Yes			Yes	
Fit statistics							
Observations	4,800	4,800	4,800	4,800	4,800	4,800	4,800
R^2	0.80576	0.47038	0.87509	0.89169	0.24003	0.75238	0.68548
Within R^2			0.71551			0.43600	

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

revenue elasticities, emphasizing that both productivity and price heterogeneity are essential for consistent estimation.

Overall, the evidence shows that accurate measurement of firm performance requires accounting for both productivity shocks and market price variation.

3 Question 3: Drawbacks of Observing Physical Output vs. Revenue

3.1 A. Revenue Regression with Isoelastic Demand Assumption

(1) Derivation Starting from the physical-output production function:

$$y_i = \alpha_L \ell_i + \alpha_K k_i + \omega_i + \varepsilon_i.$$

Let $R_i = P_i Y_i$ be firm-level revenue, and assume an isoelastic inverse demand:

$$\ln P_i = a_i - \beta \ln Y_i, \qquad \beta > 1.$$

Then,

$$r_i \equiv \ln R_i = \ln P_i + \ln Y_i = a_i + (1 - \beta) \ln Y_i.$$

Simplifying using $\ln Y_i = y_i$,

$$r_i = a_i + (1 - \beta) y_i = a_i + (1 - \beta)(\alpha_L \ell_i + \alpha_K k_i + \omega_i + \varepsilon_i).$$

Thus, the revenue regression can be expressed as:

$$r_i = \alpha_L^\star \ell_i + \alpha_K^\star k_i + \omega_i^\star + \varepsilon_i^\star,$$

where

$$\alpha_j^\star = \gamma \alpha_j, \qquad \gamma = 1 - \beta.$$

Since $\beta > 1$, $\gamma < 0$, meaning the estimated elasticities from revenue regressions differ in sign and magnitude from true physical-output elasticities.

(2) Empirical Implementation We estimate two specifications using `fixest`: (a) a physical-output regression, and (b) a revenue regression (using $R2$ as the revenue variable and including TFP).

```
# Load data
data <- read.csv("revenue_production.csv")

# Create log variables
data$y <- log(data$Y)
data$r <- log(data$R2)       # R2 is the chosen revenue variable
data$l <- log(data$L)
data$k <- log(data$K)

# (a) Physical-output regression
reg_y <- feols(y ~ l + k, data = data)

# (b) Revenue regression (TFP observed)
reg_r <- feols(r ~ l + k + tfp, data = data)

# Combine results
etable(reg_y, reg_r,
       headers = c("Physical Output", "Revenue (with TFP)"),
       digits = 3,
       title = "Comparison of Physical Output and Revenue Regressions",
       notes = "Dependent variables: log(Y) and log(R2). Dataset: revenue_
           production.csv.")
```

Listing 12: Estimation in `fixest` using `etable`

(3) Interpretation Because prices respond inversely to quantities, the estimated revenue elasticities

$$\alpha_j^\star = \gamma\alpha_j = \alpha_j(1-\beta)$$

are typically flipped in sign when $\beta > 1$, implying apparent increasing returns to scale even if the true production technology exhibits constant returns.

The disturbance term $\omega_i^\star = a_i + (1-\beta)\omega_i$ conflates technological productivity (ω_i) with demand shocks (a_i), so revenue regressions capture both supply and demand heterogeneity simultaneously.

3.2 B. Simultaneity and OLS Bias

Assume the structural system:

$$\begin{aligned} &(\text{log–PF}) \quad y_i = \alpha_0 + \alpha_L \ell_i + \omega_i, \\ &(\text{log–LD}) \quad \ell_i = \gamma_0 + \gamma_1 y_i. \end{aligned}$$

Given that $E[\omega_i] = 0$, substitute the second equation into the first:

$$y_i = \alpha_0 + \alpha_L(\gamma_0 + \gamma_1 y_i) + \omega_i.$$

Rearranging,

$$y_i(1-\alpha_L\gamma_1) = \alpha_0 + \alpha_L\gamma_0 + \omega_i \quad \Rightarrow \quad y_i = \frac{\alpha_0 + \alpha_L\gamma_0}{1-\alpha_L\gamma_1} + \frac{\omega_i}{1-\alpha_L\gamma_1}.$$

Substitute this expression back into the labor demand:

$$\ell_i = \gamma_0 + \gamma_1 y_i = \gamma_0 + \gamma_1\left(\frac{\alpha_0 + \alpha_L\gamma_0}{1-\alpha_L\gamma_1} + \frac{\omega_i}{1-\alpha_L\gamma_1}\right).$$

Covariances.

$$\begin{aligned}
&\text{Cov}(\ell_i, \omega_i) = \mathbb{E}\left[\left(\ell_i - \bar{\ell}_i\right) \cdot \omega_i\right] \quad \text{since } \mathbb{E}\left[\omega_i\right] = 0 \\
\Rightarrow\ &\text{Cov}(\ell_i, \omega_i) = \mathbb{E}\left[\frac{\gamma_1 \cdot \omega_i}{1 - \alpha_L \gamma_1} \cdot \omega_i\right] \quad \text{since } \mathbb{E}\left[\ell_i\right] = \gamma_0 + \frac{\gamma_1\left(\alpha_0 + \alpha_L \gamma_0\right)}{1 - \alpha_L \gamma_1} \\
\Rightarrow\ &\text{Cov}(\ell_i, \omega_i) = \frac{\gamma_1 \cdot \text{Var}\left(\omega_i\right)}{1 - \alpha_L \gamma_1} \quad \text{since } \mathbb{E}\left[\omega_i\right] = 0
\end{aligned} \tag{3}$$

$$\begin{aligned}
&\text{Cov}(\ell_i, y_i) = \mathbb{E}\left[\left(\ell_i - \bar{\ell}_i\right) \cdot \left(y_i - \bar{y}_i\right)\right] \\
\Rightarrow\ &\text{Cov}(\ell_i, y_i) = \mathbb{E}\left[\gamma_1 \cdot \left(y_i - \bar{y}_i\right)^2\right] \quad \text{since } \left(\ell_i - \bar{\ell}_i\right) = \gamma_1 \cdot \left(y_i - \bar{y}_i\right) \\
\Rightarrow\ &\text{Cov}(\ell_i, y_i) = \gamma_1 \cdot \text{Var}\left(y_i\right) \quad \text{since } \mathbb{E}\left[\left(y_i - \bar{y}_i\right)^2\right] = \text{Var}(y_i) \\
\Rightarrow\ &\text{Cov}(\ell_i, y_i) = \frac{\gamma_1 \cdot \text{Var}\left(\omega_i\right)}{\left(1 - \alpha_L \gamma_1\right)^2}
\end{aligned} \tag{4}$$

Since variance of constants is zero and we derived y_i in previous part.

$$\text{Cov}(\ell_i, y_i) = \frac{\gamma_1}{(1 - \alpha_L \gamma_1)^2} \text{Var}(\omega_i).$$

Interpretation. *Any case done by students is fine give full marks for any case.*

Note that $\alpha_L > 0$. However, sign of γ_1 can differ and so would its magnitude.

Case 1 If $\gamma_1 < 0$ then $1 - \alpha_L \cdot \gamma_1 > 0$. Which gives us that

$$\text{Cov}(\ell_i, \omega_i) = \frac{\gamma_1 \cdot \text{Var}\left(\omega_i\right)}{1 - \alpha_L \gamma_1} < 0.$$

SInce $\hat{\alpha}^{OLS} = \alpha_L + \frac{S_{\ell\omega}}{S_{\ell\ell}} < \alpha_L$. Recall that $S_{\ell\omega} \propto \text{Cov}(\ell_i, \omega_i)$.

Case 2 If $\frac{1}{\alpha_L} > \gamma_1 > 0$ then $1 - \alpha_L \cdot \gamma_1 > 0$. Which gives us that

$$\text{Cov}(\ell_i, \omega_i) > 0.$$

SInce $\hat{\alpha}^{OLS} = \alpha_L + \frac{S_{\ell\omega}}{S_{\ell\ell}} > \alpha_L$. Recall that $S_{\ell\omega} \propto \text{Cov}(\ell_i, \omega_i)$.

Case 3 If $\gamma_1 > \frac{1}{\alpha_L}$ then $1 - \alpha_L \cdot \gamma_1 < 0$. Which gives us that

$$\text{Cov}(\ell_i, \omega_i) = \frac{\gamma_1 \cdot \text{Var}\left(\omega_i\right)}{1 - \alpha_L \gamma_1} < 0.$$

SInce $\hat{\alpha}^{OLS} = \alpha_L + \frac{S_{\ell\omega}}{S_{\ell\ell}} < \alpha_L$. Recall that $S_{\ell\omega} \propto \text{Cov}(\ell_i, \omega_i)$.

3.3 C. Noise Correlated with Inputs

Recall the primitives:

$$\ln P_i = \ln P_0 + \mu_i, \qquad y_i = \alpha_0 + \alpha_L \ell_i + \omega_i + \varepsilon_i, \qquad r_i = \ln P_i + y_i.$$

All residuals $\mu_i, \omega_i, \varepsilon_i$ have mean zero, unit variance, and are mutually independent.

C.1) Estimating equation for r_i on ℓ_i. Write the structural expression for revenue:

$$r_i = \ln P_i + y_i = (\ln P_0 + \mu_i) + \big(\alpha_0 + \alpha_L \ell_i + \omega_i + \varepsilon_i\big)$$
$$= \underbrace{\ln P_0 + \alpha_0}_{\beta_0} + \underbrace{\alpha_L}_{\beta_1} \ell_i + \underbrace{\mu_i + \omega_i + \varepsilon_i}_{e_i}.$$

Therefore the regression

$$r_i = \beta_0 + \beta_1 \ell_i + e_i$$

has

$$\boxed{\beta_0 = \ln P_0 + \alpha_0, \qquad \beta_1 = \alpha_L, \qquad e_i = \mu_i + \omega_i + \varepsilon_i.}$$

Since $\mu_i, \omega_i, \varepsilon_i$ are independent and unit variance,

$$\boxed{\mathrm{Var}(e_i) = \mathrm{Var}(\mu_i) + \mathrm{Var}(\omega_i) + \mathrm{Var}(\varepsilon_i) = 1 + 1 + 1 = 3.}$$

C.2) OLS under orthogonality: $\hat{\beta}_1^{\mathbf{OLS}} = \alpha_L$. The OLS probability limit satisfies

$$\begin{aligned}
\hat{\beta}_1^{OLS} &= \beta_1 + \frac{\mathrm{Cov}(\ell_i, e_i)}{\mathrm{Var}(\ell_i)} \\
&= \alpha_L + \frac{\mathrm{Cov}(\ell_i, e_i)}{\mathrm{Var}(\ell_i)} \\
&= \alpha_L + \frac{\mathrm{Cov}(\ell_i, \mu_i + \omega_i + \varepsilon_i)}{\mathrm{Var}(\ell_i)} \\
&= \alpha_L + \frac{\mathrm{Cov}(\ell_i, \mu_i) + \mathrm{Cov}(\ell_i, \omega_i) + \mathrm{Cov}(\ell_i, \varepsilon_i)}{\mathrm{Var}(\ell_i)} \\
&= \alpha_L + \frac{\mathbb{E}\left[\mu_i\left(\ell_i - \bar{\ell}_i\right)\right] + \mathbb{E}\left[\omega_i\left(\ell_i - \bar{\ell}_i\right)\right] + \mathbb{E}\left[\varepsilon_i\left(\ell_i - \bar{\ell}_i\right)\right]}{\mathrm{Var}(\ell_i)}
\end{aligned}$$

The orthogonality conditions, i.e.

$$\mathbb{E}[\mu_i(\ell_i - \bar{\ell})] = \mathbb{E}[\omega_i(\ell_i - \bar{\ell})] = \mathbb{E}[\varepsilon_i(\ell_i - \bar{\ell})] = 0,$$

hold, then $\hat{\beta}_1^{OLS} = \alpha_L$ and

$$\boxed{\hat{\beta}_1^{OLS} = \alpha_L,}$$

i.e. OLS consistently recovers the true structural coefficient.

C.3) Now let $\ell_i = \gamma_0 + \gamma_\mu \mu_i + u_i$ with $\gamma_\mu < 0$ and $u_i \perp (\mu_i, \omega_i, \varepsilon_i)$, $\mathrm{Var}(u_i) = 1$. We compute $(\mathrm{Cov}(\ell_i, e_i) = \mathbb{E}\big[e_i(\ell_i - \bar{\ell})\big])$

$$\begin{aligned}
\mathbb{E}\big[e_i(\ell_i - \bar{\ell})\big] &= \mathbb{E}\big[(\mu_i + \omega_i + \varepsilon_i) \cdot (\gamma_\mu \mu_i + u_i)\big] \\
&= \gamma_\mu\, \mathbb{E}[\mu_i^2] + \underbrace{\mathbb{E}[\omega_i(\gamma_\mu \mu_i + u_i)]}_{=0} + \underbrace{\mathbb{E}[\varepsilon_i(\gamma_\mu \mu_i + u_i)]}_{=0} \\
&= \gamma_\mu \cdot 1 = \gamma_\mu.
\end{aligned}$$

So

$$\boxed{\mathbb{E}\big[e_i(\ell_i - \bar{\ell})\big] = \gamma_\mu.}$$

C.4) Show $\hat{\beta}_1 < \alpha_L$ and intuition. From the OLS formula:

$$\hat{\beta}_1^{OLS} = \alpha_L + \frac{\text{Cov}(\ell_i, e_i)}{\text{Var}(\ell_i)} = \alpha_L + \frac{\gamma_\mu}{\text{Var}(\ell_i)}.$$

Because $\gamma_\mu < 0$ by assumption, the second term is negative, therefore

$$\boxed{\hat{\beta}_1^{OLS} < \alpha_L.}$$

Intuition: measurement noise in price (μ_i) is systematically correlated with input ℓ_i (through a negative coefficient γ_μ). Since μ_i also enters the revenue residual e_i, ℓ_i is negatively correlated with the residual; this covariance biases the slope downwards. Economically, when firms reduce measured labour when price noise is high (negative γ_μ), OLS underestimates the true labour effect on output.

C.5) Now let $\ell_i = \gamma_0 + \gamma_\mu \mu_i + \gamma_\omega \omega_i + u_i$ with $u_i \perp$ everything. Compute

$$\begin{aligned}\mathbb{E}\big[e_i(\ell_i - \bar{\ell})\big] &= \mathbb{E}\big[(\mu_i + \omega_i + \varepsilon_i) \cdot (\gamma_\mu \mu_i + \gamma_\omega \omega_i + u_i)\big] \\ &= \gamma_\mu \mathbb{E}[\mu_i^2] + \gamma_\omega \mathbb{E}[\omega_i^2] + 0 \\ &= \gamma_\mu + \gamma_\omega,\end{aligned}$$

because cross-terms vanish by independence and variances are unit. Hence

$$\boxed{\mathbb{E}\big[e_i(\ell_i - \bar{\ell})\big] = \gamma_\mu + \gamma_\omega.}$$

C.6) Condition for downward bias ($\hat{\alpha}_L^{\mathbf{OLS}} < \alpha_L$). From the OLS formula:

$$\hat{\beta}_1^{OLS} = \alpha_L + \frac{\gamma_\mu + \gamma_\omega}{\text{Var}(\ell_i)}.$$

Therefore

$$\hat{\beta}_1^{OLS} < \alpha_L \quad \Longleftrightarrow \quad \gamma_\omega < -\gamma_\mu.$$

Given $\gamma_\mu < 0$, this inequality means that the labor tfp complimentarity γ_ω must be smaller than the magnitude of the systematic relationship of labor and price measurement noise ($|\gamma_\mu| = -\gamma_\mu$ because $\gamma_\mu < 0$).

4 Question 4:

4.1 A. Three-Stage Least Squares vs. IV Method

Recall the system:

$$\begin{aligned} y_i &= \alpha + \beta x_i + \varepsilon_i, && (10)\\ x_i &= \gamma_0 + \gamma_1 z_i + e_i, && (11)\\ z_i &= \delta_0 + \delta_1 w_i + u_i, && (12)\\ \varepsilon_i &= \theta_0 e_i + \mu_{1i}, && (13)\\ e_i &= \rho_0 u_i + \mu_{2i}, && (14)\end{aligned}$$

with $w_i \perp\!\!\!\perp \mu_{1i}, \mu_{2i}, u_i$. All variables have zero mean unless otherwise stated.

1) Is x_i exogenous? Why? Exogeneity of x_i for the structural equation (10) requires $\mathrm{Cov}(x_i, \varepsilon_i) = 0$. Using (11) and (13):

$$\varepsilon_i = \theta_0 e_i + \mu_{1i}, \qquad x_i = \gamma_0 + \gamma_1 z_i + e_i.$$

Thus

$$\mathrm{Cov}(x_i, \varepsilon_i) = \theta_0 \,\mathrm{Cov}(x_i, e_i) + \mathrm{Cov}(x_i, \mu_{1i}).$$

Because μ_{1i} is independent of w_i, u_i (and hence of x_i) we have $\mathrm{Cov}(x_i, \mu_{1i}) = 0$. But

$$\mathrm{Cov}(x_i, e_i) = \mathrm{Cov}(\gamma_1 z_i + e_i, e_i) = \mathrm{Var}(e_i) + \gamma_1 \rho_0 Var(u_i) \neq 0,$$

Substitute z_i and e_i after expanding the covariance term.

2) Is z_i a valid instrument for x_i? Under what condition? An instrument z_i must satisfy (i) relevance $\mathrm{Cov}(z_i, x_i) \neq 0$ and (ii) exogeneity $\mathrm{Cov}(z_i, \varepsilon_i) = 0$.

Exogeneity:

$$\mathrm{Cov}(z_i, \varepsilon_i) = \mathrm{Cov}(z_i, \theta_0 e_i + \mu_{1i}) = \theta_0 \,\mathrm{Cov}(z_i, e_i) + \mathrm{Cov}(z_i, \mu_{1i}).$$

By assumptions μ_{1i} is independent of z_i, so the second term is zero. Using $e_i = \rho_0 u_i + \mu_{2i}$ and $z_i = \delta_0 + \delta_1 w_i + u_i$, we get

$$\mathrm{Cov}(z_i, e_i) = \rho_0 \,\mathrm{Cov}(z_i, u_i) = \rho_0 \,\mathrm{Var}(u_i).$$

Hence

$$\boxed{\mathrm{Cov}(z_i, \varepsilon_i) = \theta_0 \rho_0 \,\mathrm{Var}(u_i).} \neq 0$$

So z_i is endogenous.

3) Is w_i a valid instrument for z_i? Check relevance and exogeneity.

Relevance: From (12),

$$\mathrm{Cov}(w_i, z_i) = \mathrm{Cov}(w_i, \delta_1 w_i + u_i) = \delta_1 \,\mathrm{Var}(w_i) + \mathrm{Cov}(w_i, u_i).$$

But the assumption $w_i \perp\!\!\!\perp u_i$ implies $\mathrm{Cov}(w_i, u_i) = 0$, so $\mathrm{Cov}(w_i, z_i) = \delta_1 \,\mathrm{Var}(w_i)$. Thus relevant if $\delta_1 \neq 0$.

Exogeneity: Need $\mathrm{Cov}(w_i, \varepsilon_i) = 0$. Using (13) and (14):

$$\mathrm{Cov}(w_i, \varepsilon_i) = \theta_0 \,\mathrm{Cov}(w_i, e_i) + \mathrm{Cov}(w_i, \mu_{1i}).$$

By assumption $w_i \perp\!\!\!\perp \mu_{1i}$ and $w_i \perp\!\!\!\perp e_i$ (since $e_i = \rho_0 u_i + \mu_{2i}$ and $w_i \perp\!\!\!\perp u_i, \mu_{2i}$). Therefore both covariances are zero and $\mathrm{Cov}(w_i, \varepsilon_i) = 0$.

Answer: Yes — w_i is a valid instrument for z_i provided $\delta_1 \neq 0$ (relevance), and exogeneity holds by the maintained independence assumptions.

4) Can we say w_i is a valid instrument for x_i? Yes, provided w_i is (i) relevant for x_i and (ii) exogenous for the structural equation.

Relevance: Since $x_i = \gamma_0 + \gamma_1 z_i + e_i$ and $z_i = \delta_0 + \delta_1 w_i + u_i$, we can write (in population)

$$\mathrm{Cov}(w_i, x_i) = \gamma_1 \,\mathrm{Cov}(w_i, z_i) + \mathrm{Cov}(w_i, e_i).$$

Because $w_i \perp\!\!\!\perp e_i$, $\mathrm{Cov}(w_i, e_i) = 0$, and $\mathrm{Cov}(w_i, z_i) = \delta_1 \,\mathrm{Var}(w_i)$. Thus $\mathrm{Cov}(w_i, x_i) = \gamma_1 \delta_1 \,\mathrm{Var}(w_i)$, which is nonzero if $\gamma_1 \delta_1 \neq 0$.

Exogeneity: We need $\mathrm{Cov}(w_i, \varepsilon_i) = 0$, which was already shown in part (3) to hold by assumptions.

5) Consistency and bias of $\hat{\beta}_{IV} = \dfrac{S_{zy}}{S_{zx}}$. Let $\text{Cov}(z_i, y_i) = \frac{1}{n}\sum(z_i - \bar{z})(y_i - \bar{y}) = \frac{S_{zy}}{n}$ and similarly $\text{Cov}(z_i, x_i)$. Then,

$$\hat{\beta}_{IV} = \frac{\text{Cov}(z_i, y_i)}{\text{Cov}(z_i, x_i)}.$$

Compute the numerator:

$$\text{Cov}(z_i, y_i) = \text{Cov}\left(z_i, \alpha + \beta x_i + \varepsilon_i\right) = \beta\,\text{Cov}(z_i, x_i) + \text{Cov}(z_i, \varepsilon_i).$$

Therefore

$$\hat{\beta}_{IV} = \beta + \frac{\text{Cov}(z_i, \varepsilon_i)}{\text{Cov}(z_i, x_i)}.$$

We already found $\text{Cov}(z_i, \varepsilon_i) = \theta_0\rho_0\,\text{Var}(u_i)$. Hence the *bias* equals

$$\boxed{\hat{\beta}_{IV} - \beta = \frac{\theta_0\rho_0\,\text{Var}(u_i)}{\text{Cov}(z_i, x_i)}.}$$

Conclusion: $\hat{\beta}_{IV}$ is consistent iff $\text{Cov}(z_i, \varepsilon_i) = 0$, i.e. iff $\theta_0\rho_0 = 0$ (under our structural assumptions). Otherwise it is biased by the expression above.

6) Consistency and bias of $\hat{\beta}^w_{IV} = \dfrac{S_{wy}}{S_{wx}}$. Analogously,

$$\hat{\beta}^{IV}_w = \beta + \frac{\text{Cov}(w_i, \varepsilon_i)}{\text{Cov}(w_i, x_i)}.$$

Under the maintained independence assumptions, $\text{Cov}(w_i, \varepsilon_i) = 0$ (see part 3). Thus

$$\boxed{\hat{\beta}^{IV}_w = \beta = \beta}$$

provided $\text{Cov}(w_i, x_i) \neq 0$ (i.e. w_i is relevant for x_i, which requires $\gamma_1\delta_1 \neq 0$). The estimator is consistent and asymptotically unbiased.

Answer: $\hat{\beta}^w_{IV}$ is consistent (zero asymptotic bias) under the maintained assumptions and relevance $\gamma_1\delta_1 \neq 0$.

7) Three-Stage Least Squares (3SLS) procedure and consistency of $\hat{\beta}_{3SLS}$. Procedure summary:

- Stage 1: estimate $z_i = \delta_0 + \delta_1 w_i + u_i$ by OLS to get $\hat{z}_i = \hat{\delta}_0 + \hat{\delta}_1 w_i$. By LLN, $\hat{\delta}_1 \to \delta_1$ (consistent).
- Stage 2: estimate $x_i = \gamma_0 + \gamma_1\hat{z}_i + v_i$ (i.e., regress x on $\hat{z}$) and form $\hat{x}_i$ (the fitted values). Note $\hat{z}_i$ is a function of w_i only.
- Stage 3: estimate $y_i = \alpha + \beta\hat{x}_i + \nu_i$ by OLS; the estimate of β is $\hat{\beta}_{3SLS}$.

Sketch of consistency proof: Because $\hat{z}_i$ satisfies

$$\hat{z}_i = \hat{\delta}_0 + \hat{\delta}_1 w_i$$

Since $Cov(u_i, w_i) = 0$, we have $\hat{\delta}_i = \delta_i$.

Note that since

$$\begin{aligned}\hat{\gamma}_1^{2SLS} &= \frac{Cov(\hat{z}_i, x_i)}{Var(\hat{z}_i)} = \frac{Cov(\delta_0 + \delta_1 w_i, \gamma_0 + \gamma_1(\delta_0 + \delta_1 w_i + u_i) + e_i)}{Var(\delta_0 + \delta_1 w_i)} \\ &= \frac{Cov(\delta_1 w_i, \gamma_1\delta_1 w_i + \gamma_1 u_i + v_i)}{\delta_1^2 Var(w_i)} = \frac{\delta_1^2\gamma_1 Var(w_i)}{\delta_1 Var(w_i)} = \gamma_1\end{aligned} \tag{5}$$

Now, we effectivly have the following holding:

$$\hat{x}_i = \gamma_0 + \gamma_1 \hat{z}_i = \gamma_0 + \gamma_1 \delta_0 + \gamma_1 \delta_1 w_i = c_0 + \gamma_1 \delta_1 w_i \tag{6}$$

c_0 collects all intercept terms.

$$\begin{aligned}\hat{\beta}^{3SLS} &= \frac{Cov(\hat{x}_i, y_i)}{Var(\hat{x}_i)} = \frac{Cov(c_0 + \gamma_1\delta_1 w_i, \alpha + \beta(\gamma_0 + \gamma_1 z_i + e_i) + \epsilon_i)}{Var(c_0 + \gamma_1 \delta_1 w_i)} \\ &= \frac{Cov(\gamma_1\delta_1 w_i, \beta\gamma_1 z_i)}{\delta_1^2\gamma_1^2 Var(w_i)} = \frac{Cov(\gamma_1\delta_1 w_i, \beta\gamma_1(\delta_0 + \delta_1 w_i + u_i))}{\delta_1^2\gamma_1^2 Var(w_i)} = \frac{\delta_1^2\gamma_1^2\beta Var(w_i)}{\delta_1^2\gamma_1^2 Var(w_i)} = \beta\end{aligned} \tag{7}$$

4.2 Part B: Control Function

Consider the linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \qquad X_i = \delta_0 + \delta_1 Z_i + e_i,$$

with $\text{Cov}(Z_i, \varepsilon_i) = 0$ and $\text{Cov}(Z_i, e_i) = 0$.

We examine whether the first-stage residual

$$c_i := e_i = X_i - \delta_0 - \delta_1 Z_i$$

is a valid *control function* for the structural equation.

Condition [Control]. The control variable c_i must satisfy

$$\varepsilon_i = \gamma c_i + u_i,$$

where u_i is independent of both x_{1i} (here X_i) and c_i.

Define γ and u_i as followed:

$$\gamma = \frac{\text{Cov}(\varepsilon_i, e_i)}{\text{Var}(e_i)}, \qquad u_i = \varepsilon_i - \gamma e_i.$$

First we show that $\gamma \neq 0$, note it is sufficient to show $Cov(\epsilon_i, e_i) \neq 0$

$$\begin{aligned}Cov(\epsilon_i, e_i) &= Cov(\epsilon_i, X_i - \delta_0 - \delta_1 Z_i) \\ &= Cov(\epsilon_i, X_i) - \delta_1 Cov(\epsilon_i, Z_i) \\ &= Cov(\epsilon_i, X_i) \neq 0\end{aligned} \tag{8}$$

Now we have $\epsilon_i = \gamma e_i + u_i$, we show that $Cov(X_i, u_i) = 0$

$$\begin{aligned}Cov(X_i, u_i) &= Cov(X_i, \varepsilon_i - \gamma e_i) \\ &= Cov(X_i, \varepsilon_i) - \gamma Cov(X_i, e_i) \\ &= Cov(\delta_0 + \delta_1 Z_i + e_i, \varepsilon_i) - \gamma Cov(\delta_0 + \delta_1 Z_i + e_i, e_i) \\ &= Cov(e_i, \varepsilon_i) - \gamma Cov(e_i, e_i) \\ &= Cov(e_i, \varepsilon_i) - \frac{\text{Cov}(\varepsilon_i, e_i)}{\text{Var}(e_i)} Var(e_i) = 0\end{aligned} \tag{9}$$

There are not other regressors except for the constant term. Since $Var(e_i) > 0$, e_i is not linear transformation of the constant term. (Note: $Var(e_i) = 0$ only for constants and linear transformation or any deterministic transformation of a constant is a constant.) Therefore both conditions for a control function are satisfied.

4.3 C: Tricky Regressions

1) Is W_i an instrument for X_i?

No, W_i is not a valid instrument for X_i. Evaluate $\text{Cov}(W_i, \epsilon_i)$

$$\begin{aligned}\text{Cov}(W_i, \epsilon_i) &= \text{Cov}(W_i, \gamma W_i + u_i) \\ &= \text{Cov}(W_i, \gamma W_i) + \text{Cov}(W_i, u_i) \\ &= \gamma \cdot \text{Var}(W_i) + \text{Cov}(W_i, u_i) \neq 0\end{aligned}$$

Since W_i is correlated with the error term ϵ_i, it **violates the exclusion restriction** and cannot be used as a valid instrument for X_i.

2) Is W_i a control function for X_i?

No W_i is not a control function because $Cov(W_i, u_i) \neq 0$ even though $Cov(X_i, u_i) = 0$, we still have an endogeneity problem. Check lecture-3 for the conditions needed.

3) Combined Regression and Endogenous Variables

$$\begin{aligned}Y_i &= \beta_0 + \beta_1 X_i + \epsilon_i \\ Y_i &= \beta_0 + \beta_1 X_i + \gamma W_i + u_i\end{aligned}$$

This is the single combined regression. The new error term is u_i. To determine which variables are endogenous in this new equation, we must check their correlation with the error term u_i.

- **For X_i:** The problem states that $\text{Cov}(X_i, u_i) = 0$. Therefore, X_i is **exogenous** in this regression.
- **For W_i:** The problem states that $\text{Cov}(W_i, u_i) \neq 0$. Therefore, W_i is **endogenous** in this regression.

4) Strategy for Consistent Estimation

Find an IV or control for W_i. Then if IV is found one can use IV estimation on 2SLS. If control function is found then an OLS is sufficient.

5 Question 5: Spanish Dairy Farms

This question replicates the dynamic production–function exercise using panel data on Spanish dairy farms. Output (`milk`) is produced with inputs: number of cows (proxy for captial hence denoted as k from here on), feed, labour, and land. The purpose is to compare static OLS and fixed-effects estimates with dynamic panel GMM estimators that correct for endogeneity of inputs and persistence in productivity.

Wald-test statistic and over-identifying restriction tables are provided right after the code. The table for all regressions and linearhypothesis are provided in the end.

5.1 (i) Data Preparation

We first log-transform the variables, create lags, and difference them to build the panel dataset needed for dynamic estimation.

```
# ============================================================
# Q5: Production Function Estimation using Spanish Dairy Farms
# ============================================================
library(dplyr)
library(fixest)
library(car)
library(sandwich)
library(plm)
library(tidyverse)
```

```r

spanish_dairy_farms <- read_csv("spanish_dairy_farms.csv")

# ----------------------------------------------------------------
# Data preparation
# ----------------------------------------------------------------
df <- spanish_dairy_farms %>%
  filter(if_all(c(milk, cows, feed, labor, land), ~ . > 0)) %>%
  mutate(
    y = log(milk),
    k = log(cows),       # proxy for capital
    f = log(feed),       # feed input
    l = log(labor),
    n = log(land)
  ) %>%
  group_by(farm) %>%
  arrange(year, .by_group = TRUE) %>%
  mutate(
    y_l1 = lag(y, 1),
    k_l1 = lag(k, 1),
    f_l1 = lag(f, 1),
    l_l1 = lag(l, 1),
    n_l1 = lag(n, 1),
    dy = y - lag(y, 1),
    dk = k - lag(k, 1),
    df = f - lag(f, 1),
    dl = l - lag(l, 1),
    dn = n - lag(n, 1)
  ) %>%
  ungroup()
```

Comment: This creates a balanced panel suitable for both fixed-effects and GMM estimation. The logarithmic transformation allows interpretation of coefficients as elasticities.

5.2 (ii) OLS with Time Dummies

```r
ols_td <- feols(y ~ k + f + l + n | year,
                data = df, cluster = ~farm)

linearHypothesis(ols_td, "k + f + l + n = 1",
                 vcov = vcovCL(ols_td))   # Test CRS
```

Interpretation: OLS provides baseline elasticities of output with respect to each input, controlling for year-specific shocks. The CRS test examines whether returns to scale equal one.

5.3 (iii) Two-Way Fixed Effects (farm + year)

```r
fe_tw <- feols(y ~ k + f + l + n | farm + year,
               data = df, cluster = ~farm)

linearHypothesis(fe_tw, "k + f + l + n = 1",
                 vcov = vcovCL(fe_tw))

fe_tw2 <- feols(y ~ k + f + l + n + I(farm)| year,
               data = df)

wald(fe_tw2, keep = "farm")
```

Interpretation: Two-way fixed effects control for both farm-specific productivity and time effects such as market or weather shocks.

5.4 (iv) FE-Cochrane–Orcutt

Table 4: Wald Test for Nullity of I(farm)

Statistic	Value
F-statistic	6.3462
p-value	0.0119
Numerator DoF	1
Denominator DoF	1,471

Note: H0 is I(farm) = 0. VCOV is clustered by year.

```r
co <- feols(y ~ y_l1 + k + k_l1 + f + f_l1 +
             l + l_l1 + n + n_l1 | farm + year,
             data = df, cluster = ~farm)

linearHypothesis(co, "k + f + l + n = 1",
                 vcov = vcovCL(co))
```

```r
coeff_vec<-coef(co)
restriction_df<-data.frame(parameter=c("k","l","f","n","joint"),
                           formula_=c("y_l1*k+k_lag",
                                      "y_l1*l+l_lag",
                                      "y_l1*f+f_lag",
                                      "y_l1*n+n_lag",
                                      "sum of squares of above"),
                           test_statistic=c(coeff_vec[["y_l1"]]*coeff_vec[["k"]]+coeff_vec[[
                               "k_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["l"]]+coeff_vec[[
                                                "l_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["f"]]+coeff_vec[[
                                                "f_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["n"]]+coeff_vec[[
                                                "n_l1"]],
                                            0))
restriction_df$test_statistic[5]<-sum(restriction_df$test_statistic^2)
library(kableExtra)
library(knitr)
restriction_df %>%
  # Use mutate to rename columns for the final table header
  rename(
    "Parameter" = parameter,
    "Test-statistic formula" = formula_,
    "Test Statistic value" = test_statistic
  ) %>%
  # The kable() function converts the dataframe to LaTeX
  kable(
    format = "latex",     # Specify the output format
    booktabs = TRUE,      # Use booktabs for professional-looking lines
    caption = "Test Statistics for Over-identfying Restrictions",
    label = "common-factor-restrictions",
    digits = 5,           # Set rounding for numeric columns
    align = "lcr"         # Align columns: left, center, right
  ) %>%
  # kable_styling adds further customization
  kable_styling(
    latex_options = "striped",
    position = "center"
  ) %>%
  # Add a rule to separate the joint test statistic
  row_spec(5, hline_after = TRUE)
```

Listing 13: Test Statistics for Over-identfying Restrictions

Table 5: Test Statistics for Over-identfying Restrictions

Parameter	Test-statistic formula	Test Statistic value
k	y_l1*k+k_lag	-0.02391
l	y_l1*l+l_lag	-0.05792
f	y_l1*f+f_lag	0.05183
n	y_l1*n+n_lag	0.02385
joint	sum of squares of above	0.00718

5.5 (v) Arellano–Bond (GMM) Using Second Lags as Instruments

```
pdf <- pdata.frame(df, index = c("farm", "year"))

ab1 <- pgmm(
  y ~ k + f + l + n |
    lag(y, 2:2) + lag(k, 2:2) + lag(f, 2:2) +
    lag(l, 2:2) + lag(n, 2:2),
  data = pdf,
  effect = "individual",
  model = "onestep",
  transformation = "d"
)
summary(ab1)

crs_test_ab1 <- linearHypothesis(
  ab1, hypothesis.matrix = "k + f + l + n = 1", test = "Chisq")
```

Interpretation: The Arellano–Bond estimator removes fixed effects and uses lagged levels as instruments for differenced endogenous regressors, addressing endogeneity bias.

5.6 (vi) AB+CO

```
ab2 <- pgmm(
  y ~ lag(y,1) + k + lag(k,1) + f + lag(f,1) +
      l + lag(l,1) + n + lag(n,1) |
    lag(y,2:3) + lag(k,2:3) + lag(f,2:3) +
    lag(l,2:3) + lag(n,2:3),
  data = pdf,
  effect = "individual",
  model = "onestep",
  transformation = "d"
)
summary(ab2)

crs_test_ab2 <- linearHypothesis(
  ab2, hypothesis.matrix = "k + f + l + n = 1", test = "Chisq")
```

```
coeff_vec<-coef(ab2)
restriction_df<-data.frame(parameter=c("k","l","e","m","joint"),
                           formula_=c("y_l1*k+k_lag",
                                      "y_l1*l+l_lag",
                                      "y_l1*f+f_lag",
                                      "y_l1*n+n_lag",
                                      "sum of squares of above"),
                           test_statistic=c(coeff_vec[["lag(y, 1)"]]*coeff_vec[["k"]]+coeff_
                               vec[["lag(k, 1)"]],
                                            coeff_vec[["lag(y, 1)"]]*coeff_vec[["l"]]+coeff_
                                                vec[["lag(l, 1)"]],
                                            coeff_vec[["lag(y, 1)"]]*coeff_vec[["f"]]+coeff_
                                                vec[["lag(f, 1)"]],
```

Table 6: Test Statistics for Over-identifying restrictions

Parameter	Test-statistic formula	Test Statistic value
k	y_l1*k+k_lag	-0.02549
l	y_l1*l+l_lag	0.02137
f	y_l1*f+f_lag	0.12631
n	y_l1*n+n_lag	0.02936
joint	sum of squares of above	0.01792

```
                                                coeff_vec[["lag(y,␣1)"]]*coeff_vec[["n"]]+coeff_
                                                    vec[["lag(n,␣1)"]],
                                                0))
restriction_df$test_statistic[5]<-sum(restriction_df$test_statistic^2)
library(kableExtra)
library(knitr)
restriction_df %>%
  # Use mutate to rename columns for the final table header
  rename(
    "Parameter" = parameter,
    "Test-statistic␣formula" = formula_,
    "Test␣Statistic␣value" = test_statistic
  ) %>%
  # The kable() function converts the dataframe to LaTeX
  kable(
    format = "latex",     # Specify the output format
    booktabs = TRUE,      # Use booktabs for professional-looking lines
    caption = "}Test␣Statistics␣for␣Over-identifying␣restrictions",
    label = "common-factor-restrictions",
    digits = 5,           # Set rounding for numeric columns
    align = "lcr"         # Align columns: left, center, right
  ) %>%
  # kable_styling adds further customization
  kable_styling(
    latex_options = "striped",
    position = "center"
  ) %>%
  # Add a rule to separate the joint test statistic
  row_spec(5, hline_after = TRUE)
```

Listing 14: Testing Over-identifying restrictions due to CO

5.7 (vii) Exporting Regression Tables

```
etable(ols_td, fe_tw, co,
       tex = TRUE, digits = 3)
```

Comment: The exported LaTeX table summarises all static and dynamic specifications for easy comparison of estimated input elasticities and fit statistics.

5.8 (viii) Preferred Model and Concluding Remarks

The AB+CO model is generally preferred, as it controls for unobserved heterogeneity and input endogeneity while maintaining valid instruments. The marking for this question will depend on how *justifiable and economically coherent* the student's reasoning and preferred-model choice are, not on matching exact numerical coefficients.

Table 7: Q5: Production Function Estimates for Spanish Dairy Farms

Dependent Variable:	*log(milk output)*				
Model:	(1) OLS (Year FE)	(2) Two-way FE	(3) Dynamic FE	(4) AB	(5) AB+CO
Variables					
k	0.601***	0.638***	0.597***	0.544***	0.205
	(0.042)	(0.034)	(0.036)	(0.122)	(0.220)
f	0.445***	0.308***	0.271***	0.474***	0.314**
	(0.024)	(0.019)	(0.021)	(0.053)	(0.144)
l	0.026	0.035	0.043**	0.081	0.093
	(0.023)	(0.024)	(0.021)	(0.098)	(0.165)
n	0.024	0.041**	0.035*	0.125*	0.051
	(0.021)	(0.017)	(0.018)	(0.074)	(0.133)
y_l1			0.117***		0.465***
			(0.033)		(0.152)
k_l1			-0.094**		-0.121
			(0.038)		(0.138)
f_l1			0.020		-0.020
			(0.017)		(0.067)
l_l1			-0.063**		-0.022
			(0.028)		(0.164)
n_l1			0.020		0.006
			(0.013)		(0.077)
Fixed-effects					
Year	Yes	Yes	Yes	Yes	Yes
Farm	No	Yes	Yes	Yes	Yes
Fit statistics					
Observations	1,482	1,482	1,235	988	988
R^2	0.953	0.988	0.990	—	—
Within R^2	0.951	0.710	0.681	—	—
Sargan Test (p)	—	—	—	0.011	0.152
AR(1) Test (p)	—	—	—	1.4×10^{-10}	2.0×10^{-6}
AR(2) Test (p)	—	—	—	0.758	0.098
Wald Test (p)	—	—	—	< 0.001	< 0.001

Clustered (farm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 8: Tests for Constant Returns to Scale ($H_0 : \beta_k + \beta_f + \beta_l + \beta_n = 1$)

Model Specification	**Chi-Squared Statistic**	**P-value**
OLS (Pooled)	7.7611	0.0053 **
Fixed Effects (Two-Way)	0.4183	0.5178
Cochrane-Orcutt (Dynamic FE)	2.7843	0.0952 .
Arellano-Bond (Level Eq.)	4.4987	0.0339 *
Arellano-Bond (Diff Eq.)	1.4683	0.2256

*Signif. codes: '**' p¡0.01; '*' p¡0.05; '.' p¡0.1*

6 Question 6- India KLEMS

6.1 Overview

We estimate a Cobb–Douglas production function using the **India KLEMS** dataset. The dependent variable is real gross output (GO_{it}), and the explanatory variables are the logs of capital (K_{it}), labour (L_{it}), energy (E_{it}), materials (M_{it}), and services (S_{it}) inputs:

$$y_{it} = \alpha_K k_{it} + \alpha_L l_{it} + \alpha_E e_{it} + \alpha_M m_{it} + \alpha_S s_{it} + u_{it}.$$

We test for constant returns to scale (CRS): $\alpha_K + \alpha_L + \alpha_E + \alpha_M + \alpha_S = 1$, and compare several estimation strategies: pooled OLS, two-way fixed effects, dynamic (Cochrane–Orcutt) models, and Arellano–Bond–type GMM estimators.

6.2 (1) Data Preparation

We begin by importing the dataset, taking logarithms, and constructing lagged and differenced variables used in dynamic specifications.

```r
library(dplyr)
library(fixest)
library(car)
library(sandwich)
library(plm)
library(tidyverse)

klems_wide <- read_csv(
  file = "C:/Users/anubh/Dropbox/UG Emp IO/My Slides/R_markdownfiles/PSET Datasets/
      INDIAKLEMS09012024.csv"
)

df <- klems_wide %>%
  filter(if_all(c(GO_r, K_r, EMP, II_E_r, II_M_r, II_S_r), ~ . > 0)) %>%
  mutate(
    y = log(GO_r),
    k = log(K_r),
    l = log(EMP),
    e = log(II_E_r),
    m = log(II_M_r),
    s = log(II_S_r)
  ) %>%
  group_by(IndustryCode) %>%
  arrange(Year, .by_group = TRUE) %>%
  mutate(
    y_l1 = lag(y, 1),
    k_l1 = lag(k, 1),
    l_l1 = lag(l, 1),
    e_l1 = lag(e, 1),
    m_l1 = lag(m, 1),
    s_l1 = lag(s, 1),
    dy = y - lag(y, 1),
    dk = k - lag(k, 1),
    dl = l - lag(l, 1),
    de = e - lag(e, 1),
    dm = m - lag(m, 1),
    ds = s - lag(s, 1),
    dy_lag1 = lag(dy, 1),
    y_l2 = lag(y, 2),
    k_l2 = lag(k, 2),
    l_l2 = lag(l, 2),
    e_l2 = lag(e, 2),
    m_l2 = lag(m, 2),
    s_l2 = lag(s, 2)
  ) %>%
  ungroup()
```

Listing 15: Data preparation and transformation

6.3 (2) OLS with Time Dummies

We estimate a baseline OLS regression with year fixed effects to capture aggregate shocks. The CRS hypothesis $\alpha_K + \alpha_L + \alpha_E + \alpha_M + \alpha_S = 1$ is tested using a Wald test.

```
ols_td <- feols(y ~ k + l + e + m + s | Year,
                data = df, cluster = ~IndustryCode)

linearHypothesis(ols_td, "k + l + e + m + s = 1",
                 vcov = vcovCL(ols_td))  # CRS test
```

Listing 16: OLS with time dummies and CRS test

6.4 (3) Two–Way Fixed Effects

We next control for unobserved heterogeneity across both industries and years.

```
fe_tw <- feols(y ~ k + l + e + m + s | IndustryCode + Year,
               data = df, cluster = ~IndustryCode)

linearHypothesis(fe_tw, "k + l + e + m + s = 1",
                 vcov = vcovCL(fe_tw))
```

Listing 17: Two–way fixed–effects estimation

Two–way fixed effects account for technological differences across industries and macroeconomic shocks over time.

Testing whether Industry level heterogeneity exists or not:

```
fe_tw2 <- feols(y ~ k + l + e + m + s + I(IndustryCode)|  Year,
               data = df)

wald(fe_tw2, keep = "IndustryCode")
```

Listing 18: Two–way fixed–effects estimation

Table 9: Wald Test for Joint Significance of Industry Fixed Effects

Test Detail	**Value**
Hypothesis (H_0)	Joint nullity of all IndustryCode coefficients
Test Statistic (F)	10,029.4
Numerator DoF	26
Denominator DoF	1,061
P-value	< 0.001 ***
VCOV Estimation	Clustered by Year

*Signif. codes: '***'* $p < 0.001$

6.5 (4) FE-CO

To address serial correlation and persistence in output, we add lagged output and lagged input terms:

```
co <- feols(y ~ y_l1 + k + k_l1 + l + l_l1 + e + e_l1 +
                m + m_l1 + s + s_l1 | IndustryCode + Year,
            data = df, cluster = ~IndustryCode)
coef(co)
```

```

linearHypothesis(co, "k␣+␣l␣+␣e␣+␣m␣+␣s␣=␣1",
                 vcov = vcovCL(co))
```

Listing 19: Dynamic specification with lagged variables

The code for test-statistic of the over-identifying restrictions is below. The table for the regressions is towards the end, I will display the table of test-statistic here to avoid confusion

```
coeff_vec<-coef(co)
restriction_df<-data.frame(parameter=c("k","l","e","m","s","joint"),
                           formula_=c("y_l1*k+k_lag",
                                      "y_l1*l+l_lag",
                                      "y_l1*e+e_lag",
                                      "y_l1*m+m_lag",
                                      "y_l1*s+s_lag",
                                      "sum␣of␣squares␣of␣above"),
                           test_statistic=c(coeff_vec[["y_l1"]]*coeff_vec[["k"]]+coeff_vec[[
                               "k_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["l"]]+coeff_vec[[
                                                "l_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["e"]]+coeff_vec[[
                                                "e_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["m"]]+coeff_vec[[
                                                "m_l1"]],
                                            coeff_vec[["y_l1"]]*coeff_vec[["s"]]+coeff_vec[[
                                                "s_l1"]],
                                            0))
restriction_df$test_statistic[6]<-sum(restriction_df$test_statistic^2)
library(kableExtra)
library(knitr)
restriction_df %>%
  # Use mutate to rename columns for the final table header
  rename(
    "Parameter" = parameter,
    "Test-statistic␣formula" = formula_,
    "Test␣Statistic␣value" = test_statistic
  ) %>%
  # The kable() function converts the dataframe to LaTeX
  kable(
    format = "latex",     # Specify the output format
    booktabs = TRUE,      # Use booktabs for professional-looking lines
    caption = "Test␣Statistics␣for␣Common␣Factor␣Restrictions",
    label = "common-factor-restrictions",
    digits = 5,           # Set rounding for numeric columns
    align = "lcr"         # Align columns: left, center, right
  ) %>%
  # kable_styling adds further customization
  kable_styling(
    latex_options = "striped",
    position = "center"
  ) %>%
  # Add a rule to separate the joint test statistic
  row_spec(5, hline_after = TRUE)
```

Listing 20: Testing the over-identifying restrictions

6.6 (5) Arellano–Bond (t–2 Lags as Instruments)

We estimate a dynamic panel model using GMM with second–lag instruments to correct for endogeneity between inputs and unobserved productivity.

```
pdf <- pdata.frame(df, index = c("IndustryCode", "Year"))

ab1 <- pgmm(
  y ~ k + l + e + m + s |
```

Table 10: Test Statistics for Common Factor Restrictions

Parameter	Test-statistic formula	Test Statistic value
k	y_l1*k+k_lag	0.00405
l	y_l1*l+l_lag	0.01618
e	y_l1*e+e_lag	0.00538
m	y_l1*m+m_lag	0.00850
s	y_l1*s+s_lag	-0.00579
joint	sum of squares of above	0.00041

```r
    lag(y, 2:2) + lag(k, 2:2) + lag(l, 2:2) +
    lag(e, 2:2) + lag(m, 2:2) + lag(s, 2:2),
  data = pdf,
  effect = "individual",
  model = "onestep",
  transformation = "d"
)
summary(ab1)

# Test for CRS
crs_test_ab1 <- linearHypothesis(
  ab1,
  hypothesis.matrix = "k␣+␣l␣+␣m␣+␣e␣+␣s␣=␣1",
  test = "Chisq"
)
crs_test_ab1
```

Listing 21: Pseudo–Arellano–Bond (one–step GMM)

6.7 (6) Arellano–Bond + CO

As a robustness check, we extend the GMM model to include deeper lags of inputs and output.

```r
ab2 <- pgmm(
  y ~ lag(y, 1) + k + lag(k,1) + l + lag(l,1) +
      e + lag(e,1) + m + lag(m,1) + s + lag(s,1) |
    lag(y, 2:3) + lag(k, 2:3) + lag(l, 2:3) +
    lag(e, 2:3) + lag(m, 2:3) + lag(s, 2:3),
  data = pdf,
  effect = "individual",
  model = "onestep",
  transformation = "d"
)
summary(ab2)
coef(ab2)

crs_test_ab2 <- linearHypothesis(
  ab2,
  hypothesis.matrix = "k␣+␣l␣+␣m␣+␣e␣+␣s␣=␣1",
  test = "Chisq"
)
crs_test_ab2
```

Listing 22: Arellano–Bond with deeper lag instruments

```r
coeff_vec<-coef(ab2)
restriction_df<-data.frame(parameter=c("k","l","e","m","s","joint"),
                           formula_=c("y_l1*k+k_lag",
                                      "y_l1*l+l_lag",
                                      "y_l1*e+e_lag",
```

Table 11: Test Statistics for Common Factor Restrictions

Parameter	Test-statistic formula	Test Statistic value
k	y_l1*k+k_lag	-0.00504
l	y_l1*l+l_lag	0.02143
e	y_l1*e+e_lag	-0.01886
m	y_l1*m+m_lag	0.03208
s	y_l1*s+s_lag	-0.00180
joint	sum of squares of above	0.00187

```
                                                   "y_l1*m+m_lag",
                                                   "y_l1*s+s_lag",
                                                   "sum of squares of above"),
                                  test_statistic=c(coeff_vec[["lag(y, 1)"]]*coeff_vec[["k"]]+coeff_vec[["lag(k, 1)"]],
                                                   coeff_vec[["lag(y, 1)"]]*coeff_vec[["l"]]+coeff_vec[["lag(l, 1)"]],
                                                   coeff_vec[["lag(y, 1)"]]*coeff_vec[["e"]]+coeff_vec[["lag(e, 1)"]],
                                                   coeff_vec[["lag(y, 1)"]]*coeff_vec[["m"]]+coeff_vec[["lag(m, 1)"]],
                                                   coeff_vec[["lag(y, 1)"]]*coeff_vec[["s"]]+coeff_vec[["lag(s, 1)"]],
                                                   0))
restriction_df$test_statistic[6]<-sum(restriction_df$test_statistic^2)
library(kableExtra)
library(knitr)
restriction_df %>%
  # Use mutate to rename columns for the final table header
  rename(
    "Parameter" = parameter,
    "Test-statistic formula" = formula_,
    "Test Statistic value" = test_statistic
  ) %>%
  # The kable() function converts the dataframe to LaTeX
  kable(
    format = "latex",     # Specify the output format
    booktabs = TRUE,      # Use booktabs for professional-looking lines
    caption = "Test Statistics for Common Factor Restrictions",
    label = "common-factor-restrictions",
    digits = 5,           # Set rounding for numeric columns
    align = "lcr"         # Align columns: left, center, right
  ) %>%
  # kable_styling adds further customization
  kable_styling(
    latex_options = "striped",
    position = "center"
  ) %>%
  # Add a rule to separate the joint test statistic
  row_spec(5, hline_after = TRUE)
```

Listing 23: Arellano–Bond with deeper lag instruments

6.8 (7) Export of Regression Results

```
etable(ols_td, fe_tw, co, tex = TRUE)
```

Listing 24: Export regression table to LaTeX

Table 12: Production Function Estimation: India KLEMS (Q6)

	(1) OLS (Time FE)	(2) Two-way FE	(3) Cochrane–Orcutt	(4) AB(1) GMM	(5) AB+CO GMM
Dependent variable:			y (log real output)		
k	0.2803***	0.0298	0.0232	0.0164	-0.0411
	(0.0628)	(0.0614)	(0.0375)	(0.1355)	(0.0606)
l	0.0995	0.3103**	0.0964*	0.4037	0.1471
	(0.0830)	(0.1246)	(0.0509)	(0.2657)	(0.0957)
e	-0.0332	0.1138**	0.0745**	-0.0241	0.0375
	(0.0675)	(0.0542)	(0.0291)	(0.0572)	(0.0234)
m	0.3148**	0.4592***	0.4368***	0.6737***	0.5133***
	(0.1224)	(0.0731)	(0.0737)	(0.1289)	(0.0943)
s	0.3520**	0.0759	0.1215***	0.1143	0.1220**
	(0.1301)	(0.0798)	(0.0393)	(0.0753)	(0.0457)
y_{t-1}			0.9422***		0.9212***
			(0.0102)		(0.0242)
k_{t-1}			-0.0178		0.0461
			(0.0349)		(0.0594)
l_{t-1}			-0.0746		-0.1158
			(0.0506)		(0.0950)
e_{t-1}			-0.0649*		-0.0517**
			(0.0324)		(0.0236)
m_{t-1}			-0.4031***		-0.4523***
			(0.0701)		(0.0934)
s_{t-1}			-0.1203***		-0.1157**
			(0.0369)		(0.0505)
Fixed Effects:	Year	Year + Industry	Year + Industry	Industry	Industry
Model:	OLS	FE	Dynamic FE	GMM (Diff)	GMM (Diff)
Observations	1,134	1,134	1,107	1,080	1,080
R^2 / Within R^2	0.870 / 0.754	0.983 / 0.731	0.998 / 0.974	–	–
Sargan Test (p-val)	–	–	–	1.00	1.00
AR(1) (p-val)	–	–	–	0.0197	5.07×10^{-5}
AR(2) (p-val)	–	–	–	0.0373	0.735
Wald Test (p-val)	–	–	–	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$

Clustered (IndustryCode) standard errors in parentheses
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 13: Tests for Constant Returns to Scale ($H_0 : \beta_k + \beta_e + \beta_l + \beta_m + \beta_s = 1$)

Model Specification	**Chi-Squared Statistic**	**P-value**
OLS (Pooled)	0.2291	0.6322
Fixed Effects (Two-Way)	0.1526	0.6960
Cochrane-Orcutt (Dynamic FE)	17.6510	<0.001 ***
Arellano-Bond (Level Eq.)	15.1520	<0.001 ***
Arellano-Bond (Diff Eq.)	0.3972	0.5286

*Signif. codes: '***' p¡0.001*

6.9 Question 6.6: Preferred Estimates

The preferred specification is the dynamic Arellano–Bond GMM model including lagged output. It effectively controls for endogeneity, unobserved heterogeneity, and serial correlation, while passing key specification tests (Sargan, Wald). Although other models show high R^2, they are likely biased. Marks will be awarded based on how well students justify their model choice using both econometric reasoning and empirical evidence.

7 Question 7

7.1 A: Arellano–Bond with Serial Correlation

Start from the panel production function with individual and time effects

$$y_{it} \;=\; \alpha_L \ell_{it} + \alpha_K k_{it} + \eta_i + \delta_t + u_{it}, \tag{10}$$

and assume the idiosyncratic shock follows an AR(1):

$$u_{it} \;=\; \rho\, u_{i,t-1} + a_{it}, \qquad a_{it} \perp\!\!\!\perp \{\ell_{is}, k_{is}\}_{s \le t-1}. \tag{11}$$

7.1.1 Quasi–difference and placing $\rho y_{i,t-1}$ on the RHS

Apply the usual Cochrane–Orcutt algebra but move the $\rho y_{i,t-1}$ term to the right-hand side:

$$\begin{aligned} y_{it} - \rho y_{i,t-1} &= \alpha_L(\ell_{it} - \rho \ell_{i,t-1}) + \alpha_K(k_{it} - \rho k_{i,t-1}) + (1-\rho)\eta_i + (\delta_t - \rho\delta_{t-1}) + a_{it} \\ \Rightarrow \quad y_{it} &= \rho\, y_{i,t-1} + \alpha_L \ell_{it} - \alpha_L \rho\, \ell_{i,t-1} + \alpha_K k_{it} - \alpha_K \rho\, k_{i,t-1} \\ &\quad + (1-\rho)\eta_i + (\delta_t - \rho\delta_{t-1}) + a_{it}. \end{aligned}$$

Define the new parameters and re-label the composite fixed (and time) effects:

$$\begin{aligned} &\beta_1 := \rho, \\ &\beta_2 := \alpha_L, \qquad \beta_3 := -\alpha_L \rho, \\ &\beta_4 := \alpha_K, \qquad \beta_5 := -\alpha_K \rho, \\ &\mu_i := (1-\rho)\eta_i, \quad \tau_t := \delta_t - \rho\delta_{t-1}. \end{aligned}$$

With these definitions the quasi–differenced regression written with RHS lags as regressors becomes

$$\boxed{y_{it} \;=\; \beta_1\, y_{i,t-1} \;+\; \beta_2\, \ell_{it} + \beta_3\, \ell_{i,t-1} \;+\; \beta_4\, k_{it} + \beta_5\, k_{i,t-1} \;+\; \mu_i + \tau_t + a_{it}.} \tag{12}$$

Equation (12) is the qdPF representation.

7.1.2 Differencing to remove individual effects (Arellano–Bond step)

Take first differences of (12) to eliminate μ_i:

$$\begin{aligned}\Delta y_{it} &= y_{it} - y_{i,t-1} \\ &= \beta_1(y_{i,t-1} - y_{i,t-2}) + \beta_2 \Delta \ell_{it} + \beta_3(\ell_{i,t-1} - \ell_{i,t-2}) \\ &\quad + \beta_4 \Delta k_{it} + \beta_5(k_{i,t-1} - k_{i,t-2}) + (\tau_t - \tau_{t-1}) + (a_{it} - a_{i,t-1}).\end{aligned}$$

Equivalently,

$$\boxed{\Delta y_{it} \; = \; \beta_1 \, \Delta y_{i,t-1} \; + \; \beta_2 \, \Delta \ell_{it} + \beta_3 \, \Delta \ell_{i,t-1} \; + \; \beta_4 \, \Delta k_{it} + \beta_5 \, \Delta k_{i,t-1} \; + \; \Delta \tau_t + \Delta a_{it}.} \tag{13}$$

Note: the form above uses $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ and similarly for the other lagged differences.

7.2 Instrument Variables

- In (13) the potentially endogenous regressors are $\Delta \ell_{it}$, $\Delta \ell_{i,t-1}$, $\Delta\, k_{it}$, $k_{i,t-1}$ and $y_{i,t-1}$ (since a_{it} correlates with t variables and a_{it-1} correlates with $t-1$ variables). Note $t-2$ is not correlated with a_{it} nor a_{it-1} neither are $t-3$,$t-4$,.... However, all those lags correlate with inputs by LD and KD applied iteratively. **Any two lags beyond (and including) $t-2$ is correct.**
- Valid moment conditions / instruments under standard assumptions:

$$\begin{aligned}&\text{use } \{y_{i,t-2}, y_{i,t-3}, \dots\} \text{ as instruments for } \Delta y_{i,t-1}, \Delta \ell_{it}, \Delta \ell_{i,t-1}, \Delta k_{it}, \Delta k_{i,t-1}, \\ &\text{use } \{\ell_{i,t-2}, \ell_{i,t-3}, \dots\} \text{ as instruments for } \Delta y_{i,t-1}, \Delta \ell_{it}, \Delta \ell_{i,t-1}, \Delta k_{it}, \Delta k_{i,t-1}, \\ &\text{use } \{k_{i,t-2}, k_{i,t-3}, \dots\} \text{ as instruments for } \Delta y_{i,t-1}, \Delta \ell_{it}, \Delta \ell_{i,t-1}, \Delta k_{it}, \Delta k_{i,t-1}.\end{aligned}$$

7.2.1 Estimation Using `pgmm()` with $(t-2)$ and $(t-3)$ Instruments

We now implement the Arellano–Bond estimator using the `pgmm()` function from the `plm` package. All variable creation and transformation steps are carried out using the `tidyverse` syntax for cleaner, reproducible code.

Step 1: Load the required libraries and dataset. We begin by loading the necessary packages and importing the dataset `INDIAKLEMS09012024.csv`. This dataset contains annual information by industry and year.

```r
library(tidyverse)
library(plm)

# Load dataset
ind <- read_csv("INDIAKLEMS09012024.csv")
```

Listing 25: Load libraries and dataset.

Step 2: Create and transform variables using `mutate()` and sort observations. We retain only the relevant variables — value added (`VA_r`), employment (`EMP`), and capital (`K_r`) — convert them to logarithmic form, and explicitly sort the dataset by industry and year using `arrange()`. This ensures proper temporal ordering when computing lags within each panel.

```r
ind <- ind %>%
  select(IndustryCode, Year, VA_r, EMP, K_r) %>%
  arrange(IndustryCode, Year) %>%          # ensure correct chronological order
  mutate(
    y = log(VA_r),    # Output (Value Added)
    l = log(EMP),     # Labor
    k = log(K_r)      # Capital
```

```
)
```

Listing 26: Variable creation, transformation, and sorting.

Step 3: Declare the panel structure. We declare the dataset as a panel object using `pdata.frame`, specifying the cross–sectional and time identifiers.

```
pdata <- pdata.frame(ind, index = c("IndustryCode", "Year"))
```

Listing 27: Declare panel structure.

Step 4: Specify and estimate the Arellano–Bond model. We estimate the re–parameterized Cochrane–Orcutt form:

$$y_{it} = \beta_1 y_{i,t-1} + \beta_2 \ell_{it} + \beta_3 \ell_{i,t-1} + \beta_4 k_{it} + \beta_5 k_{i,t-1} + \mu_i + \tau_t + a_{it}.$$

Endogenous regressors include the lagged dependent variable and current inputs, while the second and third lags $(t-2, t-3)$ serve as instruments. The estimation uses the one-step Arellano–Bond GMM with first–difference transformation.

```
ab_model <- pgmm(
  formula = y ~ lag(y, 1) + l + lag(l, 1) + k + lag(k, 1) |
              lag(y, 2:3) + lag(l, 2:3) + lag(k, 2:3),
  data = pdata,
  effect = "individual",
  model = "onestep",
  transformation = "d"
)
```

Listing 28: Estimate Arellano–Bond model with $(t-2)$ and $(t-3)$ instruments.

Table 14: Step 5: Summarize estimation results and diagnostics.

Variable	Estimate	Std. Error	z-value	p-value
$\text{lag}(y, 1)$	0.9173	0.0183	50.13	$< 2 \times 10^{-16}$ ***
ℓ_{it}	0.0007	0.3377	0.00	0.9984
ℓ_{it-1}	0.0408	0.3035	0.13	0.8930
k_{it}	0.1558	0.0958	1.63	0.1040
k_{it-1}	-0.1052	0.1060	-0.99	0.3212
Model Diagnostics				
Sargan test (over-id)	$\chi^2(232) = 27.00, \quad p = 1.000$			
AR(1) test	$z = -3.34, \quad p = 0.0008$			
AR(2) test	$z = -0.42, \quad p = 0.675$			
Wald test (coefficients)	$\chi^2(5) = 59127.56, \quad p < 2.22 \times 10^{-16}$			

Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

7.3 Q7 (B): Levinsohn–Petrin Control Function Approach

7.3.1 Part B(1): First–Stage Specification

Assume a parametric control function of the form

$$\phi_t(k_{it}, m_{it}, r_t) = c_K\, k_{it} + c_M\, m_{it} + c_{KM}\, k_{it} m_{it} + \delta_t.$$

Substituting this into the production function

$$y_{it} = \alpha_L \ell_{it} + \alpha_K k_{it} + \alpha_M m_{it} + \phi_t(k_{it}, m_{it}, r_t) + e_{it},$$

we obtain

$$y_{it} = \alpha_L \ell_{it} + (\alpha_K + c_K)k_{it} + (\alpha_M + c_M)m_{it} + c_{KM}(k_{it}m_{it}) + \delta_t + e_{it}.$$

Hence, the empirical first–stage equation is

$$y_{it} = \alpha_L \ell_{it} + \underbrace{(\alpha_K + c_K)k_{it} + (\alpha_M + c_M)m_{it} + c_{KM}k_{it}m_{it} + \delta_t}_{\varphi_{it}} + e_{it},$$

where $\varphi_{it} = (\alpha_K + c_K)k_{it} + (\alpha_M + c_M)m_{it} + c_{KM}k_{it}m_{it} + \delta_t$

—

7.3.2 Part B(2): Second Stage and Identification

Assume productivity evolves as

$$\omega_{it} = h(\omega_{i,t-1}) + \xi_{it}, \qquad h(\omega) = \pi_0 + \pi_1\omega + \pi_2\omega^2.$$

Replacing unobserved $\omega_{i,t-1}$ by its fitted proxy $\widehat{\varphi}_{i,t-1}$ from the first stage gives

$$\begin{aligned}\hat{\varphi}_{it} - \alpha_K k_{it} - \alpha_M m_{it} &= \pi_0 + \pi_1 \cdot (\hat{\varphi}_{it-1} - \alpha_K k_{it-1} - \alpha_M m_{it-1}) + \pi_2 \cdot (\hat{\varphi}_{it-1} - \alpha_K k_{it-1} - \alpha_M m_{it-1})^2 + \xi_{it} \\ \Rightarrow \hat{\varphi}_{it} &= \pi_0 + \alpha_K k_{it} + \alpha_M m_{it} + \gamma_{\hat{\varphi},1}\hat{\varphi}_{it-1} + \gamma_{k,1}k_{it-1} + \gamma_{m,1}m_{it-1} + \gamma_{\hat{\varphi},2}\hat{\varphi}^2_{it-1} \\ &\quad + \gamma_{k,2}k^2_{it-1} + \gamma_{m,2}m^2_{it-1} + \gamma_{m\hat{\varphi}} \cdot \hat{\varphi}_{it-1} \cdot m_{it-1} + \gamma_{k\hat{\varphi}} \cdot \hat{\varphi}_{it-1} \cdot k_{it-1} + \gamma_{mk} \cdot k_{it-1} \cdot m_{it-1} + \xi_{it}\end{aligned}$$

—

7.3.3 Part B(3): Empirical Implementation Using `fixest`

We now estimate the Levinsohn–Petrin control function model using the NBER Hospice dataset. For each specification $(\deg(\varphi_t), \deg(h)) = (2, 3), (3, 5), (4, 7)$, we follow a two–stage approach.

Step 1: Load libraries and dataset. We use `tidyverse` for data manipulation and `fixest` for both stages of the estimation.

```
library(tidyverse)
library(readr)
library(fixest)

root <- "C:/Users/anubh/Dropbox/UG Emp IO/My Slides/R_markdownfiles/PSET Datasets/
    HOSPICE_NBER"
df_op <- read_csv(paste0(root, "/hospc_nber.csv"))
```

Step 2: Define polynomial degrees. We will estimate three combinations of polynomial degrees for φ_t (first stage) and $h(\cdot)$ (second stage):

$$(\deg(\varphi_t), \deg(h)) = (2, 3), (3, 5), (4, 7)$$

```
deg_vec1 <- c(2, 3, 4)    # degrees for varphi_t
deg_vec2 <- c(3, 5, 7)    # degrees for h(.)
m1 <- list()              # store first-stage regressions
m2 <- list()              # store second-stage regressions
```

Step 3: Stage 1 (within the loop). For each polynomial degree: 1. take logs of all inputs and outputs; 2. estimate the first–stage model

$$\log y_{it} = \alpha_L \ell_{it} + \varphi_t(k_{it}, m_{it}) + e_{it},$$

where φ_t is a degree–deg(φ_t) polynomial in (k, m) interacted with year; 3. compute $\widehat{\varphi}_{it}$ as the fitted value minus the labor contribution $\alpha_L \ell_{it}$.

```
# Example of Stage 1 (inside the loop)
m1[[i]] <- feols(
  ly ~ l + I(year):poly(k, m, degree = deg, raw = TRUE),
  data = dat
)
dat$phi_hat <- fitted(m1[[i]]) - coef(m1[[i]])[["l"]] * dat$l
```

Step 4: Stage 2 (within the same loop). After obtaining $\widehat{\varphi}_{it}$, we construct lags:

$$\widehat{\varphi}_{i,t-1}, \quad m_{i,t-1}, \quad m_{i,t-2}, \quad k_{i,t-1}.$$

Then we estimate the second stage:

$$\widehat{\varphi}_{it} = \alpha_K k_{it} + \alpha_M m_{it} + h(\widehat{\varphi}_{i,t-1}, k_{i,t-1}, m_{i,t-1}) + \xi_{it},$$

where m_{it} is instrumented by $m_{i,t-2}$.

```
# Example of Stage 2 (inside the loop)
m2[[i]] <- feols(
  fml = phi_hat ~ k + poly(lag_phi_hat, k_lag, m_lag,
                          degree = deg, raw = TRUE) |
        m ~ m_lag2,
  data = dat2
)
```

Step 5: The complete loop (Stages 1 and 2 together). Now that both stages are clear, below is the full loop combining them. This runs all three polynomial combinations sequentially.

```
for (i in 1:3) {
  deg <- deg_vec1[i]

  # Prepare and clean data
  dat <- df_op %>%
    mutate(ly = log(y1),
           l  = log(labor_input),
           m  = log(equipment_input),
           k  = log(capital_state)) %>%
    na.omit() %>%
    mutate(across(c(ly, l, k, m), as.numeric)) %>%
    na.omit()

  # --- Stage 1: Estimate varphi interacted with year ---
  print(paste0("First-stage␣regression␣degree␣=␣", deg))
  m1[[i]] <- feols(
    ly ~ l + I(year):poly(k, m, degree = deg, raw = TRUE),
    data = dat
  )
  dat$phi_hat <- fitted(m1[[i]]) - coef(m1[[i]])[["l"]] * dat$l

  # --- Build lagged variables for Stage 2 ---
```

```
  dat2 <- dat %>%
    group_by(prvdr_num) %>%
    arrange(year, .by_group = TRUE) %>%
    mutate(lag_phi_hat = dplyr::lag(phi_hat, 1),
           m_lag = dplyr::lag(m, 1),
           k_lag = dplyr::lag(k, 1),
           m_lag2 = dplyr::lag(m, 2)) %>%
    ungroup() %>%
    na.omit()

  # --- Stage 2: IV regression (instrument m with m_lag2) ---
  deg <- deg_vec2[[i]]
  print(paste0("Second-stage regression degree = ", deg))
  m2[[i]] <- feols(
    fml = phi_hat ~ k + poly(lag_phi_hat, k_lag, m_lag,
                             degree = deg, raw = TRUE) |
          m ~ m_lag2,
    data = dat2
  )
}
```

Listing 29: Full two–stage loop for Levinsohn–Petrin estimation.

Step 6: Summarize and compare specifications. Finally, we report the main coefficients of interest (α_L, α_K, and α_M) from all three specifications:

```
etable(m1[[1]], m2[[1]], m1[[2]], m2[[2]], m1[[3]], m2[[3]],
       keep   = c("l", "m", "k"),
       drop   = c("\\(Intercept\\)", "^year::", "poly\\("),
       dict   = c("l"="$\\alpha_L$",
                  "k"="$\\alpha_K$",
                  "m"="$\\alpha_M$"),
       se.below    = TRUE,
       fitstat     = ~ n + r2 + ar2 + f,
       digits      = 3,
       signif.code = c("***"=0.01,"**"=0.05,"*"=0.1),
       tex = TRUE)
```

7.3.4 4) Which one to prefer

Observe the pattern in the R^2 values. Although increasing the polynomial degrees theoretically allows for a more flexible approximation of the functions $\varphi_t(\cdot)$ and $h(\cdot)$, it does not translate into improved statistical performance. As the degree rises, the number of parameters to be estimated increases rapidly while the number of observations remains fixed. This expansion leads to overfitting, inflated variance, and consequently very large standard errors.

Hence, higher polynomial degrees are not necessarily desirable: they raise model complexity without improving precision. In practice, one should increase the polynomial degree only when the sample size is sufficiently large to support the additional parameters. It is also important to note that, in Stage 1, each polynomial term is interacted with year dummies, further expanding the number of estimated coefficients.

Finally, the estimated capital elasticity is negative. This likely reflects the inelastic demand structure faced by hospitals, as discussed in Question 3, where an inelastic demand curve can give rise to negative or near-zero estimates of input elasticities.

Table 15: Levinsohn–Petrin Estimation Results: Polynomial Specifications $(2,3)$, $(3,5)$, and $(4,7)$

	(2,3)		**(3,5)**		**(4,7)**	
Dependent variable:	ly	$\hat{\varphi}$	ly	$\hat{\varphi}$	ly	$\hat{\varphi}$
Input elasticities						
α_L	0.687***		0.685***		0.692***	
	(0.007)		(0.007)		(0.007)	
α_M		0.006		0.0005		-0.071
		(0.005)		(0.0010)		(2.51)
α_K		-0.029***		-0.026***		-0.699
		(0.002)		(0.009)		(25.9)
Fit statistics						
Observations	10,339	5,248	10,339	5,248	10,339	5,248
R^2	0.7517	0.9009	0.7518	−25.41	0.7530	−23,368,515.7
Adj. R^2	0.7515	0.9005	0.7516	−25.51	0.7527	−23,511,908.3
F–test	5,211.7	1,742.2	3,128.8	−258.15	2,098.3	−162.97

Notes: IID standard errors in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

8 Q8

8.1 A: ACF-model when labor is collinear with control function

8.1.1 8(A)(1): The First Stage under the ACF Extension

We begin with the standard log–linear production function:

$$y_{it} = \alpha_L \ell_{it} + \alpha_K k_{it} + \alpha_M m_{it} + \omega_{it} + \varepsilon_{it},$$

where ω_{it} denotes the unobserved firm–specific productivity shock and ε_{it} is an i.i.d. error term.

Following Ackerberg, Caves and Frazer (ACF), suppose that productivity is a linear function of current inputs and time effects:

$$\omega_{it} = \varphi_t(k_{it}, \ell_{it}, m_{it}) = c_K k_{it} + c_L \ell_{it} + c_M m_{it} + \delta_t.$$

Substituting the above expression for ω_{it} into the production function gives

$$\begin{aligned} y_{it} &= \alpha_L \ell_{it} + \alpha_K k_{it} + \alpha_M m_{it} + (c_K k_{it} + c_L \ell_{it} + c_M m_{it} + \delta_t) + \varepsilon_{it} \\ &= \underbrace{(\alpha_L + c_L)\ell_{it} + (\alpha_K + c_K)k_{it} + (\alpha_M + c_M)m_{it} + \delta_t}_{\varphi_t} + \varepsilon_{it}. \end{aligned}$$

8.1.2 Q8(A)(2): The Second Stage under the ACF Extension

Let $\widehat{\varphi}_{it}$ denote the predicted control–function values from the first stage. Note that $\omega_{it} = \widehat{\varphi}_{it} - \alpha_K k_{it} - \alpha_L l_{it} - \alpha_M m_{it})$. Under the ACF timing assumptions the productivity law of motion is

$$\omega_{it} = h(\omega_{i,t-1}) + \xi_{it},$$

where $h(\cdot)$ is a deterministic function and ξ_{it} is the innovation. Imposing $h \approx \pi_0 + \pi_1 \omega + \pi_2 \omega^2$

$$\hat{\varphi}_{it} - \alpha_L l_{it} - \alpha_K k_{it} - \alpha_M m_{it} = \pi_0 + \pi_1 \cdot (\hat{\varphi}_{it-1} - \alpha_L l_{i-1t} - \alpha_K k_{it-1} - \alpha_M m_{it-1}) + \pi_2 \cdot (\hat{\varphi}_{it-1} - \alpha_L l_{i-1t} - \alpha_K k_{it-1} - \alpha_M m_{it-1})^2 + \xi_{it}$$

$$\begin{aligned} \Rightarrow \hat{\varphi}_{it} &= \pi_0 + \alpha_L l_{it} + \alpha_K k_{it} + \alpha_M m_{it} + \gamma_{\hat{\varphi},1}\hat{\varphi}_{it-1} + \gamma_{l,1} l_{it-1} + \gamma_{k,1} k_{it-1} + \gamma_{m,1} m_{it-1} + \gamma_{\hat{\varphi},2}\hat{\varphi}^2_{it-1} \\ &+ \gamma_{l,2} l^2_{it-1} + \gamma_{k,2} k^2_{it-1} + \gamma_{m,2} m^2_{it-1} + \gamma_{m\hat{\varphi}}\hat{\varphi}_{it-1} t m_{it-1} + \gamma_{k\hat{\varphi}}\hat{\varphi}_{it-1} k_{it-1} + \\ &+ \gamma_{mk} k_{it-1} m_{it-1} + \gamma_{l\hat{\varphi}}\hat{\varphi}_{it-1} l_{it-1} + \gamma_{lk} k_{it-1} l_{it-1} + \gamma_{lm} m_{it-1} l_{it-1} + \xi_{it} \end{aligned}$$

8.1.3 Q8(A)(3): Implementation of the ACF Estimator

In this section, we implement the Ackerberg–Caves–Frazer (ACF) estimator for the production function using the `hospc_nber.csv` dataset. The procedure consists of two stages: the first stage approximates the unobserved productivity control function $\varphi_t(\cdot)$, and the second stage recovers the structural elasticities $(\alpha_L, \alpha_K, \alpha_M)$ using lagged instruments for endogenous inputs.

Step 1: Data preparation. We begin by loading the required packages, reading the data, and constructing the logarithmic transformations of output, labor, materials, and capital. We ensure numeric conversion and removal of missing values for clean estimation.

```
library(tidyverse)
library(readr)
library(fixest)

# Define directory
root <- "C:/Users/anubh/Dropbox/UG␣Emp␣IO/My␣Slides/R_markdownfiles/PSET␣Datasets/HOSPICE_NBER"

# Load data
df_op <- read_csv(paste0(root, "/hospc_nber.csv"))

# Create log variables
dat <- df_op %>%
  mutate(
    ly = log(y1),
    l  = log(labor_input),
    m  = log(equipment_input),
    k  = log(capital_state)
  ) %>%
  na.omit() %>%
  mutate(across(c(ly, l, m, k), as.numeric))
```

Step 2: Explanation of the first stage. In the first stage, we approximate the control function $\widehat{\psi}_{it} = \varphi_t(l_{it}, k_{it}, m_{it})$ using a flexible polynomial of the input variables interacted with year dummies. This stage captures the predictable component of productivity conditional on input choices. The fitted values $\widehat{\psi}_{it}$ serve as an estimate of the unobserved productivity term ω_{it}.

```
# --- FIRST STAGE CHUNK (within the loop) ---

print(paste0("First-stage␣regression␣degree␣=␣", deg))
m1[[i]] <- feols(
  ly ~ I(year):poly(l, k, m, degree = deg, raw = TRUE),
  data = dat
)

# Predicted productivity proxy
dat$phi_hat <- fitted(m1[[i]])
```

Step 3: Explanation of the second stage. The second stage relies on the productivity law of motion

$$\omega_{it} = h(\omega_{i,t-1}) + \xi_{it},$$

which is implemented by regressing the estimated productivity $\widehat{\varphi}_{it}$ on lagged $\widehat{\varphi}_{i,t-1}$ and inputs. Labor and materials are treated as endogenous and instrumented using their second lags. This stage identifies the structural elasticities $(\alpha_L, \alpha_K, \alpha_M)$ while controlling for the dynamic evolution of productivity.

```
# --- SECOND STAGE CHUNK (within the loop) ---

```

```
dat2 <- dat %>%
  group_by(prvdr_num) %>%
  arrange(year, .by_group = TRUE) %>%
  mutate(
    lag_phi_hat = dplyr::lag(phi_hat, 1),
    m_lag  = dplyr::lag(m, 1),
    k_lag  = dplyr::lag(k, 1),
    m_lag2 = dplyr::lag(m, 2),
    l_lag  = dplyr::lag(l, 1),
    l_lag2 = dplyr::lag(l, 2)
  ) %>%
  ungroup() %>%
  na.omit()

deg <- deg_vec2[[i]]
print(paste0("Second-stage regression degree = ", deg))

m2[[i]] <- feols(
  fml = phi_hat ~ k + poly(lag_phi_hat, l_lag, k_lag, m_lag,
                           degree = deg, raw = TRUE)
        | l + m ~ l_lag2 + m_lag2,
  data = dat2
)
```

Step 4: The complete loop. Below is the full `for`-loop that executes the above two stages for each pair of polynomial degrees. The first-stage polynomial degrees are $(2, 3, 4)$, and the corresponding second-stage degrees are $(3, 5, 7)$.

```
deg_vec1 <- c(2, 3, 4)
deg_vec2 <- c(3, 5, 7)
m1 <- list()
m2 <- list()

for (i in 1:3) {
  deg <- deg_vec1[i]

  # ----- First-stage regression -----
  print(paste0("First-stage␣regression␣degree␣=␣", deg))
  m1[[i]] <- feols(
    ly ~ I(year):poly(l, k, m, degree = deg, raw = TRUE),
    data = dat
  )
  dat$phi_hat <- fitted(m1[[i]])

  # ----- Second-stage regression -----
  dat2 <- dat %>%
    group_by(prvdr_num) %>%
    arrange(year, .by_group = TRUE) %>%
    mutate(
      lag_phi_hat = dplyr::lag(phi_hat, 1),
      m_lag  = dplyr::lag(m, 1),
      k_lag  = dplyr::lag(k, 1),
      m_lag2 = dplyr::lag(m, 2),
      l_lag  = dplyr::lag(l, 1),
      l_lag2 = dplyr::lag(l, 2)
    ) %>%
    ungroup() %>%
    na.omit()

  deg <- deg_vec2[[i]]
  print(paste0("Second-stage␣regression␣degree␣=␣", deg))

  m2[[i]] <- feols(
    fml = phi_hat ~ k + poly(lag_phi_hat, l_lag, k_lag, m_lag,
                             degree = deg, raw = TRUE)
            | l + m ~ l_lag2 + m_lag2,
    data = dat2
  )
}
```

Listing 30: Full two–stage loop for ACF Estimation.

Step 5: Displaying results. The results from all stages and specifications are summarized in a single `etable`. The table presents the estimated elasticities $(\alpha_L, \alpha_M, \alpha_K)$ across different polynomial specifications.

```
etable(
  m1[[1]], m2[[1]], m1[[2]], m2[[2]], m1[[3]], m2[[3]],
  headers = list("(2,3)" = 2, "(3,5)" = 2, "(4,7)" = 2),
  tex = TRUE,
  fitstat = ~ n + r2 + ar2 + f,
  keep = c("l", "m", "k"),
  drop = c("\\(Intercept\\)", "^year::", "poly\\("),
  dict = c("l" = "$\\alpha_L$", "m" = "$\\alpha_M$", "k" = "$\\alpha_K$"),
```

```
  depvar = FALSE,
  se.below = TRUE,
  signif.code = c("***" = 0.01, "**" = 0.05, "*" = 0.1),
  digits = 3,
  title = "ACF␣Estimation␣Results:␣Polynomial␣Specifications␣$(2,3)$,␣$(3,5)$,␣and␣$(4,7)$",
  label = "tab:ACF_results",
  notes = "IID␣standard␣errors␣in␣parentheses.␣Significance␣codes:␣***␣$p<0.01$,␣**␣$p<0.05$
     ,␣*␣$p<0.10$.",
  style.tex = style.tex("aer")
)
```

Summary. The ACF estimator accounts for the dynamic nature of productivity by including $\widehat{\psi}_{i,t-1}$ in the second stage. Labor and materials are treated as endogenous and instrumented using their lagged values. The resulting estimates of α_L, α_M, and α_K represent structural input elasticities that correct for simultaneity bias arising from firm-level productivity shocks.

Table 16: ACF Estimation Results: Polynomial Specifications $(2,3)$, $(3,5)$, and $(4,7)$

	(2,3)		**(3,5)**		**(4,7)**	
Dependent variable:	*ly*	$\hat{\varphi}$	*ly*	$\hat{\varphi}$	*ly*	$\hat{\varphi}$
Input elasticities						
α_L		0.693***		-0.823		-0.224
		(0.020)		(30.0)		(177.1)
α_M		0.005		-0.753		-3.63
		(0.021)		(30.8)		(3,167.9)
α_K		-0.028***		0.100		1.06
		(0.008)		(10.9)		(1,012.5)
Observations	10,339	5,248	10,339	5,248	10,339	5,248
R^2	0.75335	0.99017	0.76032	-35,893.9	0.76098	-120,288,392.2
Adjusted R^2	0.75313	0.99010	0.75988	-36,176.6	0.76019	-121,003,296.4
F-test	3,505.3	3,348.1	1,722.8	-126.97	964.85	-168.26

Notes: IID standard errors in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

8.1.4 4) Which one to prefer

Observe the pattern in the R^2 values. Although increasing the polynomial degrees theoretically allows for a more flexible approximation of the functions $\varphi_t(\cdot)$ and $h(\cdot)$, it does not translate into improved statistical performance. As the degree rises, the number of parameters to be estimated increases rapidly while the number of observations remains fixed. This expansion leads to overfitting, inflated variance, and consequently very large standard errors.

Hence, higher polynomial degrees are not necessarily desirable: they raise model complexity without improving precision. In practice, one should increase the polynomial degree only when the sample size is sufficiently large to support the additional parameters. It is also important to note that, in Stage 1, each polynomial term is interacted with year dummies, further expanding the number of estimated coefficients.

Finally, the estimated capital elasticity is negative. This likely reflects the inelastic demand structure faced by hospitals, as discussed in Question 3, where an inelastic demand curve can give rise to negative or near-zero estimates of input elasticities.

8.2 Question 8, Part B

8.2.1 B.1) First-Stage Regression

The Olley-Pakes procedure begins with the production function:

$$y_{it} = \alpha_L \ell_{it} + \alpha_K k_{it} + \omega_{it} + e_{it}$$

The key assumption is that the unobserved productivity, ω_{it}, can be expressed as an invertible function of the firm's state variables (capital k_{it}) and investment (i_{it}), which is denoted as $\omega_{it} = \phi_t(k_{it}, i_{it}, r_{it})$.

We can substitute this into the production function:

$$y_{it} = \alpha_L \ell_{it} + \alpha_K k_{it} + \phi_t(k_{it}, i_{it}, r_{it}) + e_{it}$$

The goal of the first stage is to estimate the coefficient of the flexible input, α_L, while controlling for the influence of capital and productivity non-parametrically. We can group the terms involving capital and the productivity proxy:

$$y_{it} = \alpha_L \ell_{it} + \varphi_t(k_{it}, i_{it}) + e_{it}$$

where $\varphi_t(k_{it}, i_{it}) = \alpha_K k_{it} + \omega_{it} = \alpha_K k_{it} + \phi_t(k_{it}, i_{it}, r_{it})$.

The problem provides a specific parametric form for ϕ_t:

$$\phi_t = c_1 k_{it} + c_2 k_{it}^2 + d \cdot i_{it} + \delta_t$$

Substituting this into the expression for φ_t gives:

$$\varphi_t(k_{it}, i_{it}) = \alpha_K k_{it} + c_1 k_{it} + c_2 k_{it}^2 + d \cdot i_{it} + \delta_t$$

This leads to the following linear regression for the first stage:

$$y_{it} = \alpha_L \ell_{it} + (\alpha_K + c_1) k_{it} + c_2 k_{it}^2 + d \cdot i_{it} + \delta_t + e_{it}$$

From this regression, **the elasticity of labor, α_L, is identified**. The coefficients on capital and investment are combinations of structural parameters, so the capital elasticity, α_K, is not identified at this stage.

B.2) Quasi-Differenced Second-Stage Regression

The second stage uses the law of motion for productivity, which is assumed to be a first-order Markov process:

$$\omega_{it} = \mathbb{E}[\omega_{it} | \omega_{it-1}] + \xi_{it} = h(\omega_{it-1}) + \xi_{it}$$

The innovation term, ξ_{it}, is assumed to follow an AR(1) process:

$$\xi_{it} = \rho \xi_{it-1} + a_{it}$$

To handle the serial correlation in ξ_{it}, we apply the Cochrane-Orcutt transformation. We start with the equation for ω_{it}:

$$\omega_{it} = h(\omega_{it-1}) + \xi_{it}$$

Lagging this equation by one period gives:

$$\omega_{it-1} = h(\omega_{it-2}) + \xi_{it-1}$$

Multiply the lagged equation by ρ and subtract it from the original equation:

$$\omega_{it} - \rho \omega_{it-1} = (h(\omega_{it-1}) + \xi_{it}) - \rho(h(\omega_{it-2}) + \xi_{it-1})$$
$$\omega_{it} - \rho \omega_{it-1} = h(\omega_{it-1}) - \rho h(\omega_{it-2}) + (\xi_{it} - \rho \xi_{it-1})$$

Since $a_{it} = \xi_{it} - \rho \xi_{it-1}$, the equation becomes:

$$\omega_{it} = \rho \omega_{it-1} + h(\omega_{it-1}) - \rho h(\omega_{it-2}) + a_{it}$$

This is the quasi-differenced regression. To make it operational, we replace ω_{it} with an expression derived from the first stage: $\omega_{it} = \hat{\varphi}_{it} - \alpha_K k_{it}$, where $\hat{\varphi}_{it} = y_{it} - \hat{\alpha}_L \ell_{it}$. This yields the final expression:

$$(\hat{\varphi}_{it} - \alpha_K k_{it}) - \rho(\hat{\varphi}_{it-1} - \alpha_K k_{it-1}) = h(\omega_{it-1}) - \rho h(\omega_{it-2}) + a_{it}$$

where $\omega_{t-s} = \hat{\varphi}_{i,t-s} - \alpha_K k_{i,t-s}$ for $s = 1, 2$.

B.3) Second-Stage Regression with Parametric $h(\omega)$

For the final step, we substitute the given parametric form for the TFP transition function, $h(\omega) = \pi_0 + \pi_1\omega + \pi_2\omega^2$, into the quasi-differenced equation derived in B.2.

The terms involving $h(\cdot)$ become:

$$h(\omega_{it-1}) = \pi_0 + \pi_1\omega_{it-1} + \pi_2\omega_{it-1}^2$$

$$\rho h(\omega_{it-2}) = \rho(\pi_0 + \pi_1\omega_{it-2} + \pi_2\omega_{it-2}^2)$$

Substituting these into the equation from B.2 gives the complete second-stage regression:

$$(\hat{\varphi}_{it} - \alpha_K k_{it}) - \rho(\hat{\varphi}_{it-1} - \alpha_K k_{it-1}) = \left[\pi_0 + \pi_1\omega_{it-1} + \pi_2\omega_{it-1}^2\right] - \rho\left[\pi_0 + \pi_1\omega_{it-2} + \pi_2\omega_{it-2}^2\right] + a_{it}$$

Remembering that ω is a function of the unknown parameter α_K, we define:

$$\omega_{it-1}(\alpha_K) = \hat{\varphi}_{it-1} - \alpha_K k_{it-1}$$

$$\omega_{it-2}(\alpha_K) = \hat{\varphi}_{it-2} - \alpha_K k_{it-2}$$

This gives the following

$$\hat{\varphi}_{it} = \alpha_K k_{it} + \rho(\hat{\varphi}_{it-1} - \alpha_K k_{it-1}) + \pi_0(1-\rho) + \pi_1(\hat{\varphi}_{it-1} - \alpha_K k_{it-1}) - \rho\pi_1(\hat{\varphi}_{it-2} - \alpha_K k_{it-2}) \\ + \pi_2(\hat{\varphi}_{it-1} - \alpha_K k_{it-1})^2 - \rho\pi_2(\hat{\varphi}_{it-2} - \alpha_K k_{it-2})^2 + a_{it}$$

This equation can be re-parameterized as a linear regression and recover α_K.

Solution: Question 8, Part B

B.4) R Code Implementation

```
library(tidyverse)
library(readr)
library(fixest)

# Your folder with the downloaded NBER CSVs
root <- "C:/Users/anubh/Dropbox/UG Emp IO/My Slides/R_markdownfiles/PSET Datasets/"

df_op<-read_csv(paste0(root,"/RFSD.csv"))

deg_vec1=c(2,3,4)
deg_vec2=c(3,5,7)
m1<-list()
m2<-list()
for (i in 1:3) {
  deg<-deg_vec1[i]

  # prep + lags
  dat <- df_op %>%
    mutate(ly=log(Revenue),
           l=log(labour),
           k=log(Capital),
           i=log(investment)) %>%
    na.omit()

  dat<-dat[abs(dat$k)<Inf,]
  dat<-dat[abs(dat$l)<Inf,]
  dat<-dat[abs(dat$i)<Inf,]
  dat<-dat[abs(dat$ly)<Inf,]

  dat <- dat %>%
    mutate(across(c(ly, l, k, i), as.numeric)) %>%
```

```
    na.omit()

  # --- STAGE 1 (inline poly) ---
  print(paste0("First-stage␣regression␣degree␣=␣",deg,"\n"))
  m1[[i]] <- feols(
    ly ~ l + I(year):poly(k, i, degree = deg, raw = TRUE),
    data = dat
  )

  dat$phi_hat <- fitted(m1[[i]]) - coef(m1[[i]])[["l"]]

  dat2<- dat %>%
    group_by(firm_id) %>%
    arrange(year, .by_group = TRUE) %>%
    mutate(lag_phi_hat =dplyr::lag(phi_hat,1),
           lag_phi_hat2 =dplyr::lag(phi_hat,2),
           k_lag = dplyr::lag(k,1),
           k_lag2 = dplyr::lag(k,2)) %>%
    ungroup() %>%
    na.omit()

  deg<-deg_vec2[[i]]
  print(paste0("Second-stage␣regression␣degree␣=␣",deg,"\n"))
  m2[[i]]<-feols(fml= phi_hat ~ k +
                   poly(lag_phi_hat,k_lag,degree = deg,raw=TRUE) +
                   poly(lag_phi_hat2,k_lag2,degree = deg,raw=TRUE),
                 data=dat2)
}

# --- Generate LaTeX table from results ---
library(fixest)
etable(
  m1[[1]], m2[[1]], m1[[2]], m2[[2]], m1[[3]], m2[[3]],
  headers = list(
    "(2,3)" = 2,
    "(3,5)" = 2,
    "(4,7)" = 2
  ),
  tex = TRUE,
  fitstat = ~ n + r2 + ar2 + f,
  keep = c("l", "m", "k"),
  drop = c("\\(Intercept\\)", "^year::", "poly\\("),
  dict = c("l" = "$\\alpha_L$",
           "m" = "$\\alpha_M$",
           "k" = "$\\alpha_K$"),
  depvar = F,
  se.below = TRUE,
  signif.code = c("***" = 0.01, "**" = 0.05, "*" = 0.1),
  digits = 3,
  title = "ACF␣Estimation␣Results:␣Polynomial␣Specifications␣$(2,3)$,␣$(3,5)$,␣and␣$(4,7)$",
  label = "tab:LP_results",
  notes = "IID␣standard␣errors␣in␣parentheses.␣Significance␣codes:␣***␣$p<0.01$,␣**␣$p<0.05$
      ,␣*␣$p<0.10$."
)
```

Listing 31: R Code for Olley-Pakes Estimation

Table 17: OP-CO Estimation Results: Polynomial Specifications $(2,3)$, $(3,5)$, and $(4,7)$

	(2,3)		(3,5)		(4,7)	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
α_L	0.783***		0.782***		0.782***	
	(0.002)		(0.002)		(0.002)	
α_K		0.158***		0.162***		0.158***
		(0.003)		(0.003)		(0.003)
Fit statistics						
Observations	207,931	37,221	207,931	37,221	207,931	37,221
R^2	0.59648	0.94753	0.59724	0.94798	0.59771	0.94706
Adjusted R^2	0.59646	0.94750	0.59722	0.94792	0.59768	0.94696
F-test	51,224.3	35,354.1	30,831.5	16,523.5	20,593.9	9,360.2

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

9 Q9

9.1 Question 9A

9.1.1 (1) Largest Elasticities

- **ISIC 311:** Materials (0.500) has the highest elasticity, followed by Capital (0.240).
- **ISIC 381:** Materials (0.420) is the highest, followed by Capital (0.290) and Skilled Labour (0.188).

Interpretation: Both industries are materials-intensive. ISIC 381 is relatively more skill- and capital-intensive than ISIC 311.

9.1.2 (2) Returns to Scale (RTS)

$$\text{RTS} = \sum_j \alpha_j$$

$$\begin{aligned}\text{ISIC 311: } & 0.139 + 0.051 + 0.085 + 0.023 + 0.500 + 0.240 = 1.038, \\ \text{ISIC 381: } & 0.172 + 0.188 + 0.081 + 0.020 + 0.420 + 0.290 = 1.171.\end{aligned}$$

$$t = \frac{\widehat{RTS} - 1}{SE(RTS)}$$

$$\begin{aligned}t_{311} &= \frac{1.038 - 1}{0.059} = 0.64 \Rightarrow \text{Not significant (CRS).} \\ t_{381} &= \frac{1.171 - 1}{0.075} = 2.28 \Rightarrow \text{Significant (IRS).}\end{aligned}$$

Conclusion: ISIC 311 exhibits constant returns to scale, whereas ISIC 381 shows statistically significant increasing returns.

9.1.3 3) Complementarity via Cross-Partial for the Three Pairs

For Cobb–Douglas,

$$\frac{\partial^2 Y}{\partial X_i \partial X_j} = \alpha_i \alpha_j \frac{Y}{X_i X_j}.$$

Using $Y_{311} = 164.74$, $Y_{381} = 256.56$ and $L_{un} = 100$, $L_{sk} = 50$, $K = 100$, we compute the cross-partials for the three requested pairs:

Pair of inputs/$\frac{\partial^2 Y}{\partial X_i \partial X_j}$	$\alpha_{un}\alpha_{sk}\frac{Y_{311}}{L_{un}L_{sk}}$	$\alpha_{un}\alpha_K\frac{Y_{311}}{L_{un}K}$	$\alpha_{sk}\alpha_K\frac{Y_{311}}{L_{sk}K}$
ISIC 311	$0.139 \cdot 0.051 \cdot \frac{164.74}{100 \cdot 50} = 0.0002336$	$0.139 \cdot 0.240 \cdot \frac{164.74}{100 \cdot 100} = 0.0005496$	$0.051 \cdot 0.240 \cdot \frac{164.74}{50 \cdot 100} = 0.0004033$
ISIC 381	$0.172 \cdot 0.188 \cdot \frac{256.56}{100 \cdot 50} = 0.0016592$	$0.172 \cdot 0.290 \cdot \frac{256.56}{100 \cdot 100} = 0.0012797$	$0.188 \cdot 0.290 \cdot \frac{256.56}{50 \cdot 100} = 0.0027975$

All cross-partials are positive. Comparing magnitudes shows that ISIC 381 has substantially larger cross-partials for each pair, hence stronger complementarity between the three considered input pairs.

9.1.4 (4) Given prices and inputs: which inputs are least/most optimally employed?

Value marginal product: For Cobb–Douglas,

$$VMP_j \;=\; P \cdot MP_j \;=\; P \cdot \alpha_j \frac{Y}{x_j}.$$

The higher the value of $|VMP_j - w_j|$ the worse is its FOC holding, therefore it is less optimally employed.

Numerical VMPs (rounded to 3 decimal places):

	ISIC 311			ISIC 381		
Input	VMP	Price w	$\lvert VMP - w\rvert$	VMP	Price w	$\lvert VMP - w\rvert$
Unskilled (L_{ul})	0.687	5	4.313	1.324	5	3.676
Skilled (L_{sk})	0.504	25	24.496	2.894	25	22.106
Electricity (E)	2.100	2	0.100	3.117	2	1.117
Fuels (F)	1.137	8	6.863	1.539	8	6.461
Materials (M)	0.824	1	0.176	1.078	1	0.078
Capital (K)	1.186	10	8.814	2.232	10	7.768

Table 18: *
Cells in green denote the most optimally employed inputs (minimum $|VMP - w|$); cells in red denote the least optimally employed inputs (maximum $|VMP - w|$).
ISIC 311: Most optimal – Electricity (0.100); Least – Skilled labour (24.496).
ISIC 381: Most optimal – Materials (0.078); Least – Skilled labour (22.106).

Profits: Profit is computed as

$$\Pi = P \cdot Y - \sum_j w_j x_j.$$

With the given input quantities and prices we obtain (rounded to two decimals):

$$\Pi_{311} = 3 \cdot 164.74 - \big(5 \cdot 100 + 25 \cdot 50 + 2 \cdot 20 + 8 \cdot 10 + 1 \cdot 300 + 10 \cdot 100\big) = -2675.78,$$

$$\Pi_{381} = 3 \cdot 256.56 - \big(5 \cdot 100 + 25 \cdot 50 + 2 \cdot 20 + 8 \cdot 10 + 1 \cdot 300 + 10 \cdot 100\big) = -2400.32.$$

Both industries yield negative profits with the supplied input bundle and prices; ISIC 381 is the more profitable of the two (it has the less negative profit, −2400.32 vs −2675.78).

9.1.5 (5) MRTS between Capital and Skilled Labour

For Cobb–Douglas,

$$MRTS_{K,sk} = \frac{MP_{sk}}{MP_K} = \frac{\alpha_{sk}/L_{sk}}{\alpha_K/K} = \frac{\alpha_{sk}}{\alpha_K} \cdot \frac{K}{L_{sk}}.$$

Given $K = 100$ and $L_{sk} = 50$ ($K/L_{sk} = 2$):

$$MRTS_{311} = 2 \times \frac{0.051}{0.240} = 0.425,$$
$$MRTS_{381} = 2 \times \frac{0.188}{0.290} = 1.297.$$

Interpretation:

- ISIC 311 has the **lower** MRTS (0.425) compared to ISIC 381 (1.297).

9.1.6 (6) Underutilized or overutilized inputs

The only correct method is to compare the inputs to profit maximizing values. Since, we have derived the cost-minimizing level of inputs for a given output, and the corresponding output function. I will directly solve for the optimal level of output and then substitute that output level back to cost-minimizing input value to obtain profit maximizing level of inputs.
Recall from the cost–minimization problem that for a Cobb–Douglas production function

$$Y = A \prod_i X_i^{\alpha_i}, \qquad \sum_i \alpha_i = \alpha.$$

the cost function can be written as

$$C(Y) = \kappa Y^{1/\alpha},$$

where κ is a constant that depends on input prices and technology parameters.

$$\kappa \;=\; \alpha \cdot \left(\frac{\prod_j w_j^{\alpha_j}}{A \prod_j \alpha_j^{\alpha_j}} \right)^{1/\alpha}.$$

The profit function is therefore

$$\Pi = PY - C(Y) = PY - \kappa Y^{1/\alpha}.$$

Taking the derivative with respect to Y,

$$\frac{d\Pi}{dY} = P - \frac{\kappa}{\alpha} Y^{(1/\alpha)-1} = 0 \quad \Rightarrow \quad Y^* = \left(\frac{P\alpha}{\kappa} \right)^{\frac{\alpha}{1-\alpha}}.$$

The second derivative is given by:

$$\frac{d^2\Pi}{dY^2} = -\frac{\kappa(1-\alpha)}{\alpha^2} Y^{(1/\alpha-2)} > 0 \quad \text{since } \alpha > 1 \text{ for both industries}$$

The profit function is convex and increasing in Y, beyond Y^*. This means $Y \to \infty$ if the profits have to be maximized, to acheive that all inputs $\to \infty$. **Therefore, all inputs are underutilized.**

This is not an artefact of the production estimates being incorrect or wrong, but rather assuming prices (output+input) will not change as output and inputs of industries is increased. Because, production function estimation more often than not ignores the market structure (demand+nature of competition) these strange conclusions can be obtained.

Conclusion: Since $\alpha > 1$ for both ISIC 311 and ISIC 381, profit rises indefinitely with scale. All inputs are effectively *under-utilized* at the current bundle.

9.1.7 (7) Markup

Differentiate the cost function obtain marginal cost:

$$MC(Y) \;=\; \frac{dC}{dY} = \frac{\kappa}{\alpha} Y^{\frac{1}{\alpha}-1} = \frac{\kappa}{\alpha} Y^{\frac{1-\alpha}{\alpha}}.$$

Thus the MC curve is a power function of output with exponent $(1-\alpha)/\alpha$.

	ISIC 311	ISIC 381
Sum of elasticities α	1.038	1.171
Constant $\frac{\kappa}{\alpha}$ (defined above)	8.551	5.143
Observed output Y	164.74	256.56
Marginal cost $MC(Y) = \frac{\kappa}{\alpha} Y^{(1-S)/S}$	8.862	7.258
Price P	3	3
Markup $(P - MC)/MC$	−0.6615	−0.5866
Markup (percent)	−66.15%	−58.66%

9.2 Part B (Hospital production function)

The production specification is

$$Y_{it} = A(q_{it})\, K_{it}^{\beta_K}\, L_{it}^{\beta_\ell}, \qquad A(q_{it}) = \exp(a_0 + a_q q_{it} + \omega_{it}).$$

Table 2 (given) reports the following estimates (standard errors in parentheses):

	OLS	FE	OP
Quality effort, a_q	−0.0028 (0.0007)	−0.0018 (0.0004)	−0.0124 (0.0042)
Capital, β_K	0.4607 (0.0209)	0.1788 (0.0514)	0.5134 (0.0468)
Labor, β_ℓ	0.6723 (0.0149)	0.1855 (0.0119)	0.2453 (0.0319)

9.2.1 1) OLS → FE: covariance of inputs and TFP

OLS estimates of β_K and β_ℓ are substantially larger than the FE estimates (OLS: $\beta_K \approx 0.4607$, $\beta_\ell \approx 0.6723$; FE: $\beta_K \approx 0.1788$, $\beta_\ell \approx 0.1855$). The reason behind this change is explained by a positive correlation between inputs and hosptial+time fixed effects. There is systematic positive correlation inputs and hospital+time components of TFP. Opposite hols for quality choice

9.2.2 2) FE → OP

The OP estimates differ from FE: β_K rises under OP (0.5134) relative to FE (0.1788), while β_ℓ is moderate under OP (0.2453) and larger than FE's labor coefficient. This is indicative of negative correlation between inputs and transitory firm productivity shocks. Opposite holds for quality choice.

9.2.3 3) Trade-offs and the quality dimension

The estimated coefficient on quality effort a_q is negative in all estimators (OLS: −0.0028, FE: −0.0018, OP: −0.0124). A negative a_q implies that, holding observable inputs constant, higher measured quality effort is associated with lower measured output in these regressions. This indicates a trade-off: increasing measured quality effort is associated with lower measured output. This does show a moral dilemma, as providing best quality would amount to lower hospital annual output and lead to lower revenue. Depending upon the nature of the demand, it might be in the best interest of the hospital to engage in lower quality treatments for their patients.

9.2.4 4) Profit maximization: case $P = W_K = W_\ell = W_q = 1$, $a_0 = \omega_{it} = 0$

Let $P = 1$, $W_K = W_\ell = W_q = 1$, and assume $a_0 = 0$, $\omega_{it} = 0$. Consider the variable transformation $Q = \exp(Q)$, then $Q^{a_q} = \exp(a_q q)$. Then

$$Y = Q^{\alpha_q} K^{\beta_K} L^{\beta_\ell}, \qquad \Pi = P \cdot Y - (W_K K + W_\ell L + W_q Q) = Y - K - L - Q.$$

First-order conditions (interior) for (K, L, Q):

$$\frac{\partial \Pi}{\partial K} = \beta_K \frac{Y}{K} - 1 = 0 \quad \Rightarrow \quad K = \beta_K Y,$$

$$\frac{\partial \Pi}{\partial L} = \beta_\ell \frac{Y}{L} - 1 = 0 \quad \Rightarrow \quad L = \beta_\ell Y,$$

$$\frac{\partial \Pi}{\partial q} = a_q \frac{Y}{Q} - 1 = 0 \quad \Rightarrow \quad Q = a_q Y.$$

Remarks:

- Because the empirical a_q estimates are negative (all three estimators), the equation $Q = a_q Y$ has no interior solution (RHS is negative while LHS is strictly positive). Therefore the interior first-order condition for Q cannot be satisfied. With the constraint $q \geq 0$ the profit-maximizing quality choice is the boundary $q^* = 0$.
- With $q^* = 0$ we have $Q^* = 1$ and $Y = K^{\beta_K} L^{\beta_\ell}$. Using the FOCs for K and L we get

  $$K = \beta_K Y, \qquad L = \beta_\ell Y.$$

  Substitute into the production function:

  $$Y = (\beta_K Y)^{\beta_K} (\beta_\ell Y)^{\beta_\ell} = \beta_K^{\beta_K} \beta_\ell^{\beta_\ell} Y^{\beta_K + \beta_\ell}.$$

  Hence

  $$Y^{1-(\beta_K+\beta_\ell)} = \beta_K^{\beta_K} \beta_\ell^{\beta_\ell},$$

  and if $\beta_K + \beta_\ell < 1$ we obtain the finite interior scale:

  $$Y^* = \left(\beta_K^{\beta_K} \beta_\ell^{\beta_\ell}\right)^{\frac{1}{1-(\beta_K+\beta_\ell)}}.$$

  The corresponding optimal factors are

  $$K^* = \beta_K Y^*, \qquad L^* = \beta_\ell Y^*, \qquad q^* = 0.$$

- Since $\beta_K + \beta_\ell \leq 1$, the above solution is valid and global maxima given the constraint $q^* \geq 0$.

9.2.5 5) Now suppose $P = 1 + \exp(q) = 1 + Q$. Optimality conditions

With $P(q) = 1 + Q$ the profit function is

$$\Pi = \big(1 + Q\big) Y - K - L - Q, \qquad Y = Q^{a_q} K^{\beta_K} L^{\beta_\ell}.$$

First-order conditions:

$$\frac{\partial \Pi}{\partial K}: \quad \big(1+Q\big)\beta_K \frac{Y}{K} - 1 = 0 \quad \Rightarrow \quad K = \big(1+Q\big)\beta_K Y,$$

$$\frac{\partial \Pi}{\partial L}: \quad \big(1+Q\big)\beta_\ell \frac{Y}{L} - 1 = 0 \quad \Rightarrow \quad L = \big(1+Q\big)\beta_\ell Y,$$

$$\frac{\partial \Pi}{\partial Q}: \quad Y + \big(1+Q\big) a_q \frac{Y}{Q} - 1 = 0.$$

The q-first-order condition can be written as

$$Y\big[Q + a_q(1+Q)\big] = Q.$$

Assuming $\big[Q + a_q(1+Q)\big] > 0$

$$Y = \frac{Q}{Q + (1+Q)a_q}.$$

Let $\beta = \beta_\ell + \beta_K$. This gives us:

$$Y = Q^{a_q} K^{\beta_K} L^{\beta_\ell} \tag{14}$$

$$\Rightarrow Y = \beta_\ell^{\beta_\ell} \beta_K^{\beta_K} Q^{a_q} (1+Q)^{\beta} Y^{\beta} \tag{15}$$

$$\Rightarrow Y^{1-\beta} = \beta_\ell^{\beta_\ell} \beta_K^{\beta_K} Q^{a_q} (1+Q)^{\beta} \tag{16}$$

$$\Rightarrow \left[\frac{Q}{Q + (1+Q)a_q}\right]^{1-\beta} = \beta_\ell^{\beta_\ell} \beta_K^{\beta_K} Q^{a_q} (1+Q)^{\beta} \tag{17}$$

Full-marks upto the above equation. If the student gave the full set of equations that can be used to solve for optimal quantities and brainstormed on potential solutions, then provide full marks. Below, is how it should be done and go even further. In the exams, numerical solutionof optimization problems won't be expected, but in PSET-2 it will be expected.

Let Q^* be the solution of the above equation. It is solved numerically by minimizing the squared difference between the two sides. The code below defines the objective function, enforces $Q > 0$, and uses `optim()` with a newton method to find the minimizing Q^* using the OP estimates ($a_q = -0.0124$, $\beta_\ell = 0.2453$, $\beta_K = 0.5134$, $\beta = 0.7587$). The algorithm starts at $Q = 2$ and restricts the domain to $[10^{-4}, 50]$. The solution obtained is $Q^* \approx 2.52$ with an objective value near zero, confirming that the equality holds to high precision.

```r
# OP parameter estimates
a_q  <- -0.0124; beta_l <- 0.2453; beta_k <- 0.5134
beta <- beta_l + beta_k

# Define squared-error objective
objective <- function(Q) {
  if (Q <= 0) return(1e6)
  lhs <- (Q / (Q + (1 + Q) * a_q))^(1 - beta)
  rhs <- (beta_l^beta_l) * (beta_k^beta_k) *
         (Q^a_q) * ((1 + Q)^beta)
  (lhs - rhs)^2
}

# Minimize with bounded domain
result <- optim(par = 2, fn = objective, method = "L-BFGS-B",
                lower = 1e-4, upper = 50)

cat("Q*␣=", result$par, "Objective␣=", result$value, "\n")
```

The numerical minimization yields $Q^* = 1.5063$ with an objective value of approximately 2.15×10^{-16}, indicating that the equality holds to an extremely high degree of precision at this solution. You can check that $\left[Q^* + a_q(1 + Q^*)\right] > 0$. Find the rest of the quantities similarly. The important thing to note is that when price of hospital's service is dependent on quality, then hospitals will have an incentive to provide better quality, more than the minimum threshold.

10 Question 10

Note: Follow the same steps as in Question 9. The numbers will change, but many other things are the same. Part B is almost identical (even numerically). Dear TF, for the PSET, the exact numbers do not matter—just make sure they follow the correct logic.