1 Identification of a Cobb-Douglas Production Function

(a)

We have by definition that

$$y_t = f(k_t, l_t, m_t) + \omega_t + \epsilon_t$$

and assume, as in GNR, that k_t and l_t are dynamic but m_t is freely adjustable in each period, conditional on ω_t.

Following GNR, we can write

$$\begin{aligned}
\mathbb{E}[y_t \mid \Gamma_t] &= \mathbb{E}[f(k_t, l_t, m_t)] + \mathbb{E}[\omega_t \mid \Gamma_t] + \mathbb{E}[\epsilon_t \mid \Gamma_t] \\
&= \mathbb{E}[f(k_t, l_t, m_t)] + \mathbb{E}[\delta_0 + \delta_1 \omega_{t-1} + \eta_t \mid \Gamma_t] + \mathbb{E}[\epsilon_t \mid \Gamma_t] \\
&= \mathbb{E}[f(k_t, l_t, m_t)] + \delta_0 + \delta_1 \mathbb{E}[\omega_{t-1} \mid \Gamma_t] \\
&= \alpha_k k_t + \alpha_l l_t + \alpha_m \mathbb{E}[m_t \mid \Gamma_t] + \delta_0 + \delta_1 h(\phi(k_{t-1}, l_{t-1}, m_{t-1}) + d_{t-1} - f(k_{t-1}, l_{t-1}, m_{t-1})),
\end{aligned}$$

where we have rewritten ω_t using its definition and use the fact that η_t and ϵ_t are exogenous.

(b)

The firm's problem with respect to materials is

$$\max_{M_t} P_t \mathbb{E}[F(K_t, L_t, M_t) \exp(\omega_t + \epsilon_t) \mid \Gamma_t] - \rho_t M_t.$$

The first order condition is given by

$$P_t \frac{\partial}{\partial M_t} F(K_t, L_t, M_t) \exp(\omega_t) \mathcal{E} = \rho_t.$$

Rearranging, we have

$$\alpha_m K_t^{\alpha_k} L_t^{\alpha_l} M_t^{\alpha_m - 1} = \frac{\rho_t}{P_t \mathcal{E}} \exp(-\omega_t).$$

Taking logs,

$$\begin{aligned}
(\alpha_m - 1) m_t &= \log\left(\frac{\rho_t}{P_t \mathcal{E}}\right) - \omega_t - \alpha_k k_t - \alpha_l l_t - \log(\alpha_m) \\
&= d_t - \omega_t - \alpha_k k_t - \alpha_l l_t - \log(\alpha_m).
\end{aligned}$$

Recall from above that

$$\begin{aligned}
\mathbb{E}[y_t \mid \Gamma_t] &= \alpha_k k_t + \alpha_l l_t + \alpha_m \mathbb{E}[m_t \mid \Gamma_t] + \mathbb{E}[\omega_t \mid \Gamma_t] \\
&= \alpha_k k_t + \alpha_l l_t + \frac{\alpha_m}{\alpha_m - 1} \mathbb{E}[d_t - \omega_t - \alpha_k k_t - \alpha_l l_t - \log(\alpha_m) \mid \Gamma_t] + \mathbb{E}[\omega_t \mid \Gamma_t] \\
&= \left(1 - \frac{\alpha_m}{\alpha_m - 1}\right)(\alpha_k k_t + \alpha_l l_t + \mathbb{E}[\omega_t \mid \Gamma_t]) + \frac{\alpha_m}{\alpha_m - 1}(d_t - \log(\alpha_m)) \\
&= \frac{1}{1 - \alpha_m}(\alpha_k k_t + \alpha_l l_t + \mathbb{E}[\omega_t \mid \Gamma_t]) + \frac{\alpha_m}{\alpha_m - 1}(d_t - \log(\alpha_m)) \\
&= \frac{1}{1 - \alpha_m}(\alpha_k k_t + \alpha_l l_t + h(\phi_{t-1} + d_{t-1} - f(k_{t-1}, l_{t-1}, m_{t-1})) + \frac{\alpha_m}{\alpha_m - 1}(d_t - \log(\alpha_m))
\end{aligned}$$

(c)

Recall that

$$\mathbb{E}[y_t \mid k_t, l_t, m_t] \equiv \phi(k_t, l_t, m_t) + d_t.$$

By the FOC on materials,

$$\mathbb{E}[y_t \mid k_t, l_t, m_t] = \mathbb{E}[\omega_t + f(k_t, l_t, m_t) \mid k_t, l_t, m_t] = \mathbb{E}[d_t + m_t - log(\alpha_m) \mid k_t, l_t, m_t].$$

Thus,

$$\phi_t = m_t - \log(\alpha_m).$$

Plugging this into the previous expression,

$$\mathbb{E}[y_t \mid \Gamma_t] = \frac{1}{1-\alpha_m}\left[\alpha_k k_t + \alpha_l l_t + \delta_0 + \delta_1 h(m_{t-1} - \log(\alpha_m) + d_{t-1} - f(k_{t-1}, l_{t-1}, m_{t-1})\right] + \frac{\alpha_m}{\alpha_m - 1}(d_t - \log(\alpha_m)).$$

(d)

We assume that prices are fixed, so d_t doesn't change. By the logic of GNR, the only variables that exogenously shifts m_t is η_t, which is unobserved. Conditional on Γ_t, the expectation of m_t is a constant. Because we have $\frac{1}{1-\alpha_m}$ multiplying the other factor coefficients, these also cannot be identified.

(e)

We can use the firm's materials FOC to write

$$\mathbb{E}[Y_t/M_t \mid \Gamma_t]\alpha_m = \rho/P \implies \alpha_m = \frac{\rho}{P} \cdot \frac{1}{\mathbb{E}\left[\frac{Y_t}{M_t} \mid \Gamma_t\right]}.$$

So we can identify α_m, since the right-hand side is just data. From this, we can identify the other factor coefficients, e.g. by the procedure in (d). Finally, we can back out $\hat{\omega}_t$ for each observation, and run an autogregression of $\hat{\omega}_t$ on $\hat{\omega}_{t-1}$. This will identify δ_0 and δ_1.

2 Estimation of Production Functions

(a) Summary Statistics

I haven't used Julia in the past to produce latex tables so forgive the somewhat shabby formatting of the tables.

Variable	Observations	Mean	Median	Std Dev	p25	p75
Output	11393	13.7903	13.6358	1.82123	12.2636	15.2589
Investment	11393	8.77443	9.75418	4.18482	7.61776	11.7141
Capital	11393	11.9133	11.8011	2.09978	10.1414	13.6564
Hours	11393	4.90301	4.58619	1.39403	3.66766	6.11561
Intermediate C	11393	13.2585	13.1634	1.95939	11.7206	14.8021

Table 1: Summary Statistics - Unbalanced Panel

Investment has the highest variance in the data. There are no firms with zero labour or zero materials used. There seems to be significant entry and exit across industries and years.

year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1990	14	36	50	61	35	9	56	26	15	15	64	16	73	13	37	39	25	13
1991	18	60	68	102	53	10	92	43	23	22	97	23	112	29	54	71	46	20
1992	23	75	86	124	58	11	108	52	27	25	127	28	133	37	70	86	58	25
1993	24	74	84	112	63	12	99	56	25	28	136	30	126	36	74	84	65	25
1994	29	85	92	119	75	11	105	59	22	35	129	25	129	35	76	85	72	24
1995	32	86	92	123	80	9	106	59	24	40	128	25	131	32	68	86	77	24
1996	34	91	88	125	84	9	100	59	25	41	126	23	132	41	69	92	80	24
1997	35	91	88	163	93	14	118	70	30	44	135	19	139	50	86	103	95	29
1998	30	88	85	160	90	12	112	67	31	39	128	18	135	46	83	101	92	26
1999	29	71	71	137	76	9	91	58	27	27	107	15	121	32	72	81	84	22

Table 2: Industry Year Counts - Unbalanced Panel

year	Zero Investment	Zero Labour	Zero Materials
1990	74	0	0
1991	143	0	0
1992	189	0	0
1993	239	0	0
1994	215	0	0
1995	196	0	0
1996	179	0	0
1997	193	0	0
1998	137	0	0
1999	112	0	0

Table 3: Zero Counts - Unbalanced Panel

(b) Summary Statistics - Balanced Panel

The count of observations is 2,470 and lower as expected given entry and exit. Means of relevant variables are higher implying that there is selection into the balanced panel, i.e. larger or "better" firms are more likely to stay in the sample throughout.

Variable	Observations	Mean	Median	Std Dev	25th Percentile	75th Percentile
Output	2470	14.0358	14.1047	1.6304	12.5887	15.3642
Investment	2470	9.44331	10.2731	3.75307	8.4784	11.9286
Capital	2470	12.321	12.2325	1.86071	10.8466	13.8514
Hours	2470	5.15916	5.07243	1.26825	4.00891	6.2508
Intermediate C	2470	13.4876	13.5886	1.75933	12.0704	14.8781

Table 4: Summary Statistics - Balanced Panel

year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1990	7	21	15	27	15	3	27	13	6	6	26	5	33	5	10	15	10	3
1991	6	21	14	26	15	3	27	14	6	6	26	5	33	5	10	16	11	3
1992	7	21	14	26	14	3	27	14	6	6	26	5	33	5	10	16	11	3
1993	7	21	14	26	14	3	27	14	6	6	26	5	33	5	10	16	11	3
1994	7	21	14	26	14	3	27	14	6	6	26	5	33	5	10	16	11	3
1995	7	21	14	27	14	3	28	13	6	6	26	5	32	5	10	16	11	3
1996	7	21	14	27	14	3	28	13	6	6	26	5	32	5	10	16	11	3
1997	6	21	14	28	14	3	28	13	6	6	26	5	32	5	10	16	11	3
1998	6	21	14	27	14	4	27	13	6	6	26	5	33	5	10	16	11	3
1999	6	21	14	27	13	4	27	13	7	6	26	5	33	5	10	16	11	3

Table 5: Industry Year Counts - Balanced Panel

year	Zero Investment	Zero Labour	Zero Materials
1990	19	0	0
1991	28	0	0
1992	29	0	0
1993	32	0	0
1994	29	0	0
1995	29	0	0
1996	27	0	0
1997	24	0	0
1998	16	0	0
1999	22	0	0

Table 6: Zero Counts - Balanced Panel

(c) Standard Regression Models - Balanced Panel

I execute the standard models in julia. I am going to skip printing output in tables for this one because I don't have much experience printing these out of julia. But the Jupyter notebook prints the output.

I choose industry no. 7. From these models, we can learn that the coefficient on materials is high consistently across models, implying a high elasticity of output with respect to materials. Can't interpret the coefficients too well, the sign on capital seems to change. Negative coefficient on capital does not make sense to me.

(d) Standard Regression Models - Unbalanced Panel

Using the balanced panel, the coefficient on materials is now lower. The coefficient on labour is larger now. Thus the balanced panel likely has firms with large material usage and smaller labour usage in otuput.

(e) Estimation in Stata

I did the parts with a Stata command in Stata. I didn't want to spend time figuring out how to do them in Julia. I attempted to code up OP myself but gave up midway for the sake of time.

	(1) AB	(2) BB	(3) OP	(4) LP
L.log of gross output	0.328***	0.279***		
	(0.078)	(0.069)		
log of capital	0.072**	0.023	0.108	0.158
	(0.035)	(0.027)	(0.068)	(0.103)
log of labour	0.157	0.026	0.892***	0.277***
	(0.109)	(0.081)	(0.050)	(0.031)
log of materials	0.662***	0.673***		
	(0.069)	(0.070)		
Constant	-1.215	0.649		
	(0.790)	(0.441)		
Observations	607	794	987	987

Standard errors in parentheses
* p¡0.10, ** p¡0.05, *** p¡0.01

Table 7: Production Function Estimates

(e) ACF

The coefficients I get are $\beta_k = 0.25$ and $\beta_l = 0.82$. I think covergence is failing and I can't successfully debug. But both grid-search and Nelder-Mead yield similar coefficients. And these estimates make sense in comparison to the estimates in ACF (2015). I bootstrap to get standard errors.

(e) GNR

I did first stage and gave up. Sorry!