Contents

Basics

- Gross output $\neq$ sold quantity: existing inventory can be sold; output can be added to inventory.

- **Gross Value Added** (GVA): value of gross output - value of intermediate goods. Useful to compare industries that vary in their intermediate input dependence. Defining GVA as V,

$$P^V V = PQ - P^M M \tag{1}$$

  where for multi-good firms the products can be interpreted as inner products. If $P^M M$ is not available, P^V can be equivalently quantified as expenditure in all other inputs, by definition.

- The *existence* of a GVA function inherently requires separability of intermediate goods from other inputs in the production function.

- *Total Factor Productivity* (TFP) is the variation in productivity not accounted for by variation in observed inputs. Explanations for this wedge include managerial ability, management practices, knowledge capital. If the researcher is unable to differentiate within inputs (e.g. labour by skill), the unaccounted-for heterogeneity will be conflated with TFP. If the functional form is wrong, the

misspecification error will be conflated with TFP. Just like the Solow residual, TFP is effectively a measure of our ignorance.

- For the purpose of the course, we treat firms as producing a **single good**. The implicit assumption is that sale prices serve as an appropriate weight to generate a unique 'aggregate output' quantity.

- Prices are often adjusted by deflating them with a sector-specific deflator, $P_{ijt}Q_{ijt}\frac{P_{j0}}{P_{jt}}$ – where j is firm's i industry. With the availability of panel data one can just include (sector-specific) time dummies and avoid deflating.

- **Capital**: usually constructed by cumulating investment and applying arbitrary depreciation rates. Both capital and **intemediate inputs** prices are deflated with sector-specific price indices.

- **Labour**: most commonly proxied by headcount or hours worked, irrespective of the quality of labour employed. The legitimacy of this approximation can be tested if both wage bill $\sum_{n\in L_{it}} W_{int}L_{int}$ and number of employees L_{it} is available. Assuming a Cobb-Douglas production function and noting $\sum_{n\in L_{it}} W_{int}L_{int} = \bar{W}_{it}L_{it}$, a Wald test for the estimated coefficient of the average wage in a logged Cobb-Douglas should suffice to determine whether the headcount is a good enough measure of labour input. If the coefficient on average wage is not significantly different from the coefficient on headcount, the wage bill might be a better measure of labour input; in any case, the inclusion of average wage provides an adjustment for labour quality under the assumption that market prices serve as sensible weights.

- If an input can be completely and costlessly adjusted in reaction to new information, it is called **perfectly flexible**. Intermediate goods and labour are often assumed to be flexible; capital is usually thought to have some adjustment cost: these include fixed costs and irreversibility, which enforce a region of action for low marginal return, and convex costs, which instead smooth adjustment over time.

The Problem with Flexible Inputs

$$Q_{it} = A_{it}K_{it}^{\beta_K}L_{it}^{\beta_L}M_{it}^{\beta_M}$$

for a perfectly flexible input – say labour – output maximisation implies

$$\frac{\partial Q_{it}}{\partial L_{it}} = \beta_L \frac{Q_{it}}{L_{it}}$$

while the dual revenue maximisation implies (where the financial cost of adjustment costs is Ξ_{it})[1]

$$\max_{L_{it}} \Pi_{it} \equiv P_{it}Q_{it} - R_{it}K_{it} - W_{it}L_{it} - P_{it}^M M_{it} - \Xi_{it} \Leftrightarrow \frac{\partial \Pi_{it}}{\partial L_{it}} = P_{it}\frac{\partial Q_{it}}{\partial L_{it}} - W_{it} = 0 \Leftrightarrow \frac{\partial Q_{it}}{\partial L_{it}} = \frac{W_{it}}{P_{it}}$$

Putting the two together,

$$\frac{W_{it}}{P_{it}} = \beta_L \frac{Q_{it}}{L_{it}} \Leftrightarrow L_{it}^{1-\beta_L} = \beta_L \frac{P_{it}}{W_{it}}A_{it}K_{it}^{\beta_K}M_{it}^{\beta_M} \tag{2}$$

---

[1]Expressing Ξ_{it} in financial terms allows not having to define a functional form for adjustment costs.

taking logs,

$$\ell_{it} = \frac{1}{1-\beta_L}\ln\beta_L + \frac{\beta_K}{1-\beta_L}k_{it} + \frac{\beta_M}{1-\beta_L}m_{it} + \frac{1}{1-\beta_L}(p_{it}-w_{it}) + \frac{1}{1-\beta_L}a_{it} \tag{3}$$

which shows that ℓ_{it} depends linearly on a_{it}, as higher Hicks-neutral technology increases the marginal return to any input.

If we assume that a_{it} is observed by the firm, and it is not controlled for by the econometrician, we have a standard endogeneity problem (*transmission bias*). This problem is less severe for capital, chosen at $t-1$, if there is no autocorrelation in a_{it}.

Panel Data Estimators

Panel data estimators can help in solving the endogeneity problem if one is willing to assume that the a_{it} term can be additively decomposed as, e.g., $a_{it} = a_i + \tau_t + \zeta_{it}$. In this case both Within Groups and First Difference estimators allow to eliminate a_i and account for τ_t by either inclusion of time dummies (WG) or a constant (FD). Consistency of either estimator depends on ζ_{it}: if inputs are predetermined with respect to ζ_{it}, FD is preferable; under strict exogeneity, WD can be used.[2]

Instrumental Variables

Given (3), an identifying condition can be formed as $\mathbb{E}[(p_{it}-w_{it})a_{it}] = 0$ and similarly for all inputs. This identifying condition rests on TFP not being related to output price and labour wages: for one, TFP cannot be correlated with market power. In any case, firm-level variation in input and output prices is only very rarely recorded in available data.

Identification with Flexible Inputs

It is rather evident that linear panel data estimators outlined rely on rather strong assumptions, and IV relies on rarely available data. More involved approaches have been developed to identify coefficient with flexible inputs, which vary depending on available data.

Quantity Data

Suppose a log Cobb-Douglas form with measurement error

$$\tilde{q}_{it} = \beta_K k_{it} + \beta_M m_{it} + \beta_L \ell_{it} + a_{it} + \epsilon_{it} \tag{4}$$

Approaches for quantity data can be divided in (i) **proxy methods** and (ii) **IV methods**. The former assume that a_{it} is a strictly monotonic function with a subset of inputs as *only* arguments. The latter impose structure on the dynamics of lnTFP and exploit identifying conditions thus created.

There are restrictions on the functional forms for an admissible proxy. Suppose for example that proxy s_{it} is a linear function of a_{it}, e.g. $a_{it} = \theta s_{it}$. The optimality conditions still imply (3): substituting θs_{it} in there, one can show that ℓ_{it} is a linear function of included inputs, s_{it}, and prices. Since we

---

[2]Given the persistence of firm input and output, these estimators tend to difference away much of the difference in levels that exists in the data, potentially compromising precise estimation.

already condition on inputs and s_{it} in (4), the only variation that ensures ℓ_{it} is not perfectly collinear with other variables is variation in prices. More generally, if the proxy is linear, **any flexible input can be *only* identified by idiosyncratic variation in its price and/or output price**. But the optimality conditions were derived under assumption of price-taking behaviour in the output market: hence, the sole identifying variation for flexible input parameters in this model must lie in independent differences in input prices. Given (3) and no variation in input price, any variable related to ℓ_{it} must be related to a_{it}, making it an invalid instrument.

Progress is possible by imposing further structure; specifically, under the assumption that the process for $(p_{it} - w_{it})$ is *strictly* more persistent than the process for a_{it}, one can define – by recursive substitution – identifying conditions on price lags unrelated to TFP lags (á la Blundell-Bond).

Expenditure Data

Suppose (4) applies, but we now have sales figures $P_{it}Q_{it}$ and expenditure in labour $P_{it}^L L_{it}$.

$$p_{it}q_{it} = p_{it} + \beta_K k_{it} + \beta_M m_{it} + \beta_L \ell_{it} + a_{it} \Leftrightarrow p_{it}q_{it} = (p_{it} - \beta_L p_{it}^L) + \beta_K k_{it} + \beta_M m_{it} + \beta_L(\ell_{it} + p_{it}^L) + a_{it} \quad (5)$$

To identify β_L we need to control for $(p_{it} - \beta_L p_{it}^L)$: the most likely candidate is time fixed effects, which however require no cross-sectional variation in prices. However, this assumption of no cross-sectional variation **contradicts** the identifying variation necessary to avoid perfect collinearity given (3). Hence, there seems to be no viable way of identifying β_L from expenditure without imposing further assumptions.[3]

Cost Share Identification

$$\frac{W_{it}}{P_{it}} = \beta_L \frac{Q_{it}}{L_{it}} \Leftrightarrow \frac{W_{it}/P_{it}}{Q_{it}/L_{it}} = \beta_L \quad (6)$$

Under perfect competition and Cobb Douglas functional form, **the above must hold exactly for each firm**.[4] This procedure will most likely return different β_Ls: this can be ascribed to misspecification, market power, the contemporaneous existence of multiple production functions, and measurement error. β_L can then be taken as the average of estimated cost shares; equivalently, if one is willing to assume that variation arises as a result of mean-zero measurement error, regression of log cost shares on a constant will return an estimate of $\ln \beta_L$.

Proxy Methods I: Olley and Pakes (1996)

The authors consider identification of $\{\beta_L, \beta_K\}$ in the following **value added** specification:

$$\tilde{v}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + a_{it} + \varepsilon_{it} \quad (7)$$

The **timing assumption** is that capital is chosen at $t-1$, after observing a_{it-1} and $t-1$ prices; furthermore, a_{it} is assumed to follow a **first-order Markov process** with respect to filtration Ω_{t-1}: $\mathbb{E}(a_{it}|\Omega_{t-1}) = g(a_{it-1})$ and hence $a_{it} = g(a_{it-1}) + \xi_{it}$, $\mathbb{E}(\xi_{it}|g(a_{it-1})) = 0$.

---

[3]Steve mentions deflated prices, but they don't really make a difference – this can be noticed by taking logs of deflated prices. Progress can be made by differencing the data, but this requires an assumption on the dynamic process of a_{it}.

[4]In fact, the result applies for more general production functions – e.g. translog – conditional on appropriate assumptions on competition environment and flexibility of the input.

Identifying Assumption. *$k_{it+1} = i_t(k_{it}, a_{it})$, where $i_t(\cdot)$ is strictly monotonic in a_{it} and **only** a function of its arguments.*

The strict monotonicity assumption imposes the absence of, for example, fixed adjustment costs and non-convex costs. The assumption that $i_t(\cdot)$ does not depend on any other variable instead requires no persistent autocorrelation in wages across firms, as such persistence would make investment a function of labour (problematic for first-stage identification of β_L).

OP First Stage

Given strict monotonicity of $i_t(\cdot)$ in a_{it}, invert to obtain $a_{it} = h_t(k_{it+1}, k_{it})$ and substitute in the production function:

$$\tilde{v}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + h_t(k_{it+1}, k_{it}) + \varepsilon_{it} \equiv \beta_L \ell_{it} + \phi_t(k_{it+1}, k_{it}) + \varepsilon_{it} \tag{8}$$

the assumption of no persistence in the real wage for each firm is *fundamental* as it provides the identifying variation necessary to consistently estimate β_L, the only parameter that can be obtained from the first stage: if there is persistence in the wage process, real wages enter $\phi_t(\cdot)$ via k_{t+1} and β_L is not separately identified. In practice, $\phi_t(\cdot)$ can be either approximated by a polynomial or treated non-parametrically; the choice will matter as to the procedure used for estimation, as in the latter case semi-parametric estimators á la Robinson (1988) will be necessary.

Validity of the proxy can be tested by testing significance of lagged inputs added to the first stage. If the proxy is valid, e.g. ℓ_{it-1} should have no explanatory power conditional on the inclusion of the proxy.

OP Second Stage

Conditional on having obtained estimates $\hat{\beta}_L$ and $\hat{\phi}_t(\cdot)$, under the assumption that $a_{it} = g(a_{it-1}) + \xi_{it}$ one has

$$\begin{aligned}
\tilde{v}_{it} - \hat{\beta}_L \ell_{it} &= \beta_K k_{it} + g(\hat{h}_{t-1}(k_{it}, k_{it-1})) + \xi_{it} + \varepsilon_{it} \\
\tilde{v}_{it} - \hat{\beta}_L \ell_{it} &= \beta_K k_{it} + g(\hat{\phi}_{t-1}(k_{it}, k_{it-1}) - \beta_K k_{it-1}) + \xi_{it} + \varepsilon_{it}
\end{aligned} \tag{9}$$

which might not be simple to estimate but at least – given the strong assumptions imposed to use the proxy – does not contain endogeneity. Equation (9) can, for example, be estimated by nonlinear least squares if $g(a_{it-1}) \equiv \rho a_{it-1}$.

It is worth noting that, when a_{it} is AR(1) and $k_{it+1} = i_t(a_{it}, k_{it}) = \theta a_{it} + \vartheta k_{it}$, one has

$$\tilde{v}_{it} - \hat{\beta}_L \ell_{it} = \beta_K k_{it} + \rho(\frac{k_{it}}{\theta} - \frac{\vartheta k_{it-1}}{\theta} - \beta_K k_{it-1}) + \xi_{it} + \varepsilon_{it}$$

$$\tilde{v}_{it} - \hat{\beta}_L \ell_{it} = (\beta_K + \frac{\rho}{\theta})k_{it} + -\rho(\frac{\vartheta}{\theta} + \beta_K)k_{it-1} + \xi_{it} + \varepsilon_{it}$$

from which it is evident that β_K is not identified. In particular, for linear $\phi_t(\cdot)$ one would need e.g. asymmetric convex adjustment costs, which would satisfy proxy requirements and add a source of variation.

Given $\hat{\phi}_t(\cdot)$ is a regressor generated from the first stage, inference will be non-standard – the easiest way to go about this is bootstrapping standard errors.

Alternatively, as noted in Wooldridge (2009), identifying conditions from both stages of the problem can be stacked and the entire system solved by non-linear GMM (note the different timing assumptions due to the first-order Markov assumption):

$$\text{First Stage: } \mathbb{E}(\varepsilon_{it}|\ell_{is} \ \forall \ s \leq t, k_{is} \ \forall \ s \leq t+1) = 0$$
$$\text{Second Stage: } \mathbb{E}(\varepsilon_{it} + \xi_{it}|\ell_{is} \ \forall \ s \leq t-1, k_{is} \ \forall \ s \leq t) = 0$$

(10)

the assumptions can be generally be expressed as $\mathbb{E}[Z'u(\theta)] = 0$. Estimation by non-linear GMM overrides the generated regressor problem, as standard errors can be obtained via the delta method; however, the **assumptions underlying the identifying conditions are the same as OP** – i.e. the issues with the methodology itself carry over.

Proxy Methods II: Levinsohn and Petrin (2003)

The authors consider identification of $\{\beta_L, \beta_K, \beta_M\}$ in the following **quantity** specification:

$$\tilde{q}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + \beta_M m_{it} + a_{it} + \varepsilon_{it}$$

(11)

Both ℓ_{it} and m_{it} are assumed to be flexible inputs.

Identifying Assumption. *$m_{it} = \mu_t(a_{it}, k_{it})$, where $\mu_t(\cdot)$ is strictly monotonic in a_{it} and **only** a function of the arguments.*

Just as in OP, we invert $\mu_t(\cdot)$ to obtain $a_{it} = \chi_t(m_{it}, k_{it})$ and substitute in:

$$\tilde{q}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + \beta_M m_{it} + \chi_t(m_{it}, k_{it}) + \varepsilon_{it} \equiv \beta_L \ell_{it} + \phi_t(m_{it}, k_{it}) + \varepsilon_{it}$$

(12)

where $\phi_t(m_{it}, k_{it}) \equiv \beta_K k_{it} + \beta_M m_{it} + \chi_t(m_{it}, k_{it})$. For the proxy assumptions to be respected, we require no *cross-sectional* variation in intermediate input prices and wages.

Note the difference with OP: previously, we were dealing with the relationship between k_{it+1} and the price of labour at t, hence we only required no **persistence** to ensure future capital choices are independent of current prices. Now instead we need the choice of flexible input m_{it} to be independent of **contemporaneous** prices for prices not to enter the process that determines m_{it}, the price of which then must not vary cross-sectionally.

This change is not harmless: we know from (3) that, in absence of cross-sectional price variation and conditioning on other input choices, **there is no residual variation in the flexible input** to identify its coefficient. Hence there is a **fundamental tension** between the variation required for identification of β_L in the first stage and the assumptions imposed to identify β_M, β_K in the second stage.

This can be shown to be the case by deriving (3) for m_{it} and ℓ_{it}: substituting ℓ_{it} with its optimality condition then returns m_{it} as a linear function of prices and k_{it}. The latter is conditioned upon, and the former are assumed not to vary in the cross-section – hence leaving no informative residual variation. The LP estimator simply **does not identify production function parameters**.

Blundell Bond (2000): Two-Stage Methods Are Dead ...

Assuming an AR(1) process[5] (**stronger assumption** than the simple first-order Markov process of OP/LP) for a_{it} in (7), we have

$$\tilde{v}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + \rho a_{it-1} + \xi_{it} + \varepsilon_{it} \tag{13}$$

which can be transformed to rid the equation of a_{it-1} and estimated with the dynamic panel data GMM estimator of Blundell Bond (2000). As in OP, assume ℓ_{it} is flexible (and hence correlated with p_{it} and a_{it}) and k_{it} is chosen at $t-1$ (and hence correlated with a_{it-1}).

Equation (13) can be **quasi-differenced** as follows:

$$\tilde{v}_{it} - \rho \tilde{v}_{it-1} = \beta_L(\ell_{it} - \rho \ell_{it-1}) + \beta_K(k_{it} - \rho k_{it-1}) + \xi_{it} + (\varepsilon_{it} - \rho \varepsilon_{it-1})$$

under standard assumptions on the measurement error process $\{\varepsilon_i\}_{t \in T}$ and on innovations to $\{a_i\}_{t \in T}$, the above equation is not prone to transmission bias. Note that $\{k_{is}, s \leq t\}$ are uncorrelated with the composite error term, as are $\{\ell_{is}, s \leq t-1\}$ (ℓ_{it} is correlated with ξ_{it} via its dependence on a_{it}); furthermore, $\{\tilde{v}_{is}, s \leq t-2\}$ are also independent of the error term. We thus have a large number of identifying conditions that allow estimation of $\{\beta_K, \beta_L, \rho\}$ by non-linear GMM.

Identifying Assumption. *The dynamics of a_{it} must be parametrically defined to determine the set of instruments. Measurement error must not be serially correlated.*

Identification - Flexible Inputs

Quasi-differencing (3),

$$\ell_{it} - \rho \ell_{it-1} = \frac{1-\rho}{1-\beta_L} \ln \beta_L + \frac{\beta_K}{1-\beta_L}(k_{it} - \rho k_{it-1}) + \frac{1}{1-\beta_L}(p_{it} - w_{it} - \rho p_{it-1} + \rho w_{it-1}) + \frac{1}{1-\beta_L}\xi_{it}$$

hence, for an instrument to be informative about $\ell_{it} - \rho \ell_{it-1}$ after conditioning on capital, it must be correlated with $(p_{it} - w_{it} - \rho p_{it-1} + \rho w_{it-1})$. Lagged values of the labour quasi-difference will then be informative instruments iff **there is persistence in the real wage process** – i.e. the lagged quasi-difference of the real wage is correlated with its own past lags – **and real wages vary in the cross-section**.[6]

Note however that the informativeness condition above is not sufficient: from the cost share approach, we have

$$\tilde{v}_{it} = -\ln \beta_L + \ell_{it} + (w_{it} - p_{it}) + \varepsilon_{it} \tag{14}$$

now, if the real wage follows an AR(1) process $(w_{it} - p_{it}) = \pi(w_{it-1} - p_{it-1}) + \zeta_{it}$, quasi-differencing the above yields

$$\tilde{v}_{it} - \pi \tilde{v}_{it-1} = -(1-\pi)\ln \beta_L + (\ell_{it} - \pi \ell_{it-1}) + \zeta_{it} + (\varepsilon_{it} - \pi \varepsilon_{it-1})$$

which is nested in (13) and identified by the same orthogonality conditions. Hence, it is impossible

---

[5]The following applies for any autoregressive / moving average process with limited length.

[6]Variation in the cross-section is necessary so that the quasi-differenced real wage cannot be almost completely explained by the included constant.

to determine whether non-linear GMM returns the set of coefficients associated with one or the other equation. To distinguish the two we require the **real wage to follow a dynamic process that strictly nests the dynamic process of a_{it}**.

In particular, this implies that the sets of lags orthogonal to ξ_{it} and ζ_{it} are not exactly the same. Assume for example that the real wage follows an ARMA(1,1) process

$$w_{it} - p_{it} = \pi(w_{it-1} - p_{it-1}) + \zeta_{it} + \omega\zeta_{it-1}$$

and hence

$$\tilde{v}_{it} - \pi\tilde{v}_{it-1} = -(1-\pi)\ln\beta_L + (\ell_{it} - \pi\ell_{it-1}) + \underbrace{\zeta_{it} + \omega\zeta_{it-1} + (\varepsilon_{it} - \pi\varepsilon_{it-1})}_{\equiv u_{it}}$$

then the identifying condition $\mathbb{E}(\ell_{it-1}u_{it}) = 0$ is not valid for specification (14) resulting from the cost share approach, but it is still valid for specification (13) resulting from quasi-differencing the original production function: in fact, $\omega\mathbb{E}[\ell_{it-1}\zeta_{it-1}] \neq 0$, but $\mathbb{E}[\ell_{it-1}\xi_{it}] = 0$.[7]

Identification - Predetermined Inputs

A similar problem to OP Stage 2 arises for the identification of predetermined input coefficients: in particular, for a linear investment function, the set of moments does not identify a unique set of parameters from the reduced form. As a result, we require the investment function **not to be a linear function of a_{it} and k_{it} only**. Such requirement is for example fulfilled if the investment equation is linear but real wages have non-negligible persistence; alternatively, non-convex capital adjustment costs (fixed costs, irreversibility, asymmetric adjustment costs) suffice to cause the relationship to be non-linear.

Hypothesis Testing

Given BB is estimated with standard non-linear GMM, we can test whether a subset of these assumptions are appropriate via **Hansen's test of overidentifying restrictions** – conditional on assuming a number of conditions sufficient to make the system identified hold.

Furthermore, the validity of the dynamic specification can be tested: for example, for $a_{it} = \rho a_{it-1} + \xi_{it}$ the quasi-differenced model can be expressed as

$$\tilde{v}_{it} = \underbrace{\rho}_{\sigma_1}\tilde{v}_{it-1} + \underbrace{\beta_L}_{\sigma_2}\ell_{it} + \underbrace{(-\beta_L\rho)}_{\sigma_3}\ell_{it-1} + \underbrace{\beta_K}_{\sigma_4}k_{it} + \underbrace{(-\beta_K\rho)}_{\sigma_5}k_{it-1} + \xi_{it} + (\varepsilon_{it} - \rho\varepsilon_{it-1})$$

implying the following testable nonlinear restrictions: $\sigma_3 = -\sigma_1\sigma_2$ and $\sigma_5 = -\sigma_1\sigma_4$.

Generalisations and Limitations

- Unsurprisingly, as long as measurement error is not endogenous,[8] the Blundell-Bond estimator can accommodate **arbitrary sources of measurement error from multiple inputs**, as identifying conditions can be derived for correctly-measured inputs. OP and LP estimators are unable to account for measurement error in a similar manner.

---

[7]This clearly assumes strict exogeneiry of ℓ with respect to measurement error ε, as assumed pretty much throughout.
[8]If measurement error is endogenous, identifying conditions will need to change accordingly.

- **Timing assumptions** are not fundamental to the BB estimator, as long as sufficient moment conditions can be found given the data to obtain an identified system. OP and LP rest on timing assumptions with little room for modification.

- The BB estimator **can accommodate fixed effects in** a_{it}, as the specification of the a_{it} process is only specified to aptly transform the data. OP/LP cannot allow for FE due to the assumed Markovian structure.[9]

- BB allows to stack identifying conditions from orthogonality between u_{it} and **both instruments in levels and in differences** (*system GMM*).

- BB **does not generalise easily to processes for** a_{it} **that are linear in non-linear lags**, unless we assume no measurement error (i.e. no ε_{it}). This is because the nonlinear lags would, by substitution, contain products of inputs and their additive measurement error, and very strong assumptions would be needed on the form of the measurement error to find suitable instruments. Nonlinearity due to production functions other than Cobb-Douglas (e.g. translog) is also problematic, but BB can be extended to such a case more readily than to nonlinear lags.

- By requiring persistent variation in the real wage, BB **requires a measure of quantity of labour, not simply a measure of the wage bill**. Using the wage bill to quantify labour inherently assumes that there exist market-level compensations that provide appropriate weights to combine different labour inputs; in turn, this requires no systematic cross-sectional variation in wages to allow comparing wage bills among firms.[10]

Bond and Söderbom (2005): Identification with no Price Variation

β_K can be identified without resorting to variation in capital price: first, if the production function is Cobb-Douglas, labour is flexible, and firms are price-takers, the cost share approach returns consistent estimates of β_L for each firm; second, one can then net labour out from the log Cobb Douglas equation and, if say a_{it} is AR(1) and capital is predetermined, β_K can be directly estimated from

$$\tilde{v}_{it} - \rho\tilde{v}_{it-1} = \hat{\beta}_L(\ell_{it} - \rho\ell_{it-1}) + \beta_K(k_{it} - \rho k_{it-1}) + \xi_{it} + (\varepsilon_{it} - \rho\varepsilon_{it-1})$$

Bond and Söderbom (2005) explore whether asymmetric adjustment costs – expressed in financial terms to avoid having to specify a functional form for the costs – are sufficient to identify $\{\beta_L, \beta_K\}$.

Ackerberg, Caves and Frazer (2015): ...Long Live Two-Stage Methods!

$$\tilde{v}_{it} = \beta_L\ell_{it} + \beta_K k_{it} + a_{it} + \varepsilon_{it} \tag{15}$$

The ACF estimator proceeds as follows:

1. Use a LP proxy variable not with the aim of estimating any coefficient, but rather with the aim of **eliminating measurement error**

---

[9]OP/LP would probably reply that the Markovian assumption is however less restrictive than the full parameterisation required by BB.

[10]It is worth noting that the cost share approach does not assume *anything* about the wage process, as long as the production function is Cobb-Douglas and firms are price takers.

2. **Specify a dynamic process for a_{it} and estimate coefficients using BB GMM** on the data rid of measurement error.

Identifying Assumption. *For intermediate inputs m_{it}, $m_{it} = \mu_t(k_{it}, \ell_{it}, a_{it})$. m_{it} depends on these inputs **only** and is **strictly monotonic** in a_{it}.*

One can then invert $\mu_t(\cdot)$ to obtain $a_{it} = h_t(k_{it}, \ell_{it}, m_{it})$ and

$$\tilde{v}_{it} = \beta_L \ell_{it} + \beta_K k_{it} + h_t(k_{it}, \ell_{it}, m_{it}) + \varepsilon_{it} \equiv \phi_t(k_{it}, \ell_{it}, m_{it}) + \varepsilon_{it}$$

where $\phi_t(k_{it}, \ell_{it}, m_{it}) \equiv \beta_L \ell_{it} + \beta_K k_{it} + h_t(k_{it}, \ell_{it}, m_{it})$. A consistent estimate of $\phi(\cdot)$, given standard assumptions on measurement error, is a consistent estimate of unobserved error-free output.
Allowing $\phi_t(\cdot) = \hat{\phi}_t(\cdot) + \varsigma_{it}$, one then has

$$\hat{\phi}_t(\cdot) \equiv \widehat{(\tilde{v}_{it} - \varepsilon_{it})} = a_{it} + \beta_L \ell_{it} + \beta_K k_{it} - \varsigma_{it}$$
$$a_{it} = \hat{\phi}_t(\cdot) - \beta_L \ell_{it} - \beta_K k_{it} + \varsigma_{it}$$

which can be **estimated given a parametric specification for a_{it}**. For example, if a_{it} is AR(1) the equation above can be quasi-differenced to treat $\hat{\phi}_t(\cdot)$ as the dependent variable and identifying restrictions follow in the same way as in BB. Given the similarity with BB, affine requirements are necessary to ensure that the estimating equation does not have statistically equivalent representations. Sufficient persistence in the real wage process and/or asymmetric adjustment costs can provide the additional conditions necessary to isolate the estimated parameter set.
Just as in LP and OP, **inference is not standard** due to the generated regressor $\hat{\phi}_t(\cdot)$. Bootstrapping can provide valid confidence intervals.

<div align="center">

ACF-BB Comparison

</div>

Taking the simple case in which $a_{it} = \rho_1 a_{it-1} + \rho_2 a_{it-2} + \xi_{it}$, we have

$$\phi_t(\cdot) - \beta_L \ell_{it} - \beta_K k_{it} = \rho_1(\phi_{t-1}(\cdot) - \beta_L \ell_{it-1} - \beta_K k_{it-1}) + \rho_2(\phi_{t-2}(\cdot) - \beta_L \ell_{it-2} - \beta_K k_{it-2}) + \xi_{it} \quad \text{(ACF)}$$

$$\tilde{v}_{it} - \beta_L \ell_{it} - \beta_K k_{it} = \rho_1(\tilde{v}_{it-1} - \beta_L \ell_{it-1} - \beta_K k_{it-1}) + \rho_2(\tilde{v}_{it-2} - \beta_L \ell_{it-2} - \beta_K k_{it-2}) + \xi_{it} \quad \text{(BB)}$$

First, the two estimators form identifying conditions from different sets of moments: while both estimators use conditions on the orthogonality of $\{\ell_{is}, s \leq t-1\}$ and $\{k_{is}, s \leq t\}$, they differ in that BB uses $\{\tilde{v}_{is}, s \leq t-2\}$ while ACF uses $\{\phi_s(\cdot), s \leq t-1\}$ thanks to the assumed functional form of the proxy. Secondly, BB requires measurement error to be uncorrelated, whereas ACF does not need to impose any structure on the measurement error as it is eliminated in the first stage.[11]
Finally, while ACF – provided validity of the proxy – extends to the case of nonlinear technological change more straightforwardly, BB (as mentioned above) does not.

A further point to note is that **the ACF estimator should only be used to estimate value added production functions**: to estimate a quantity production function, one would have to assume that $\mu_t(k_{it}, \ell_{it}, a_{it})$ is an optimality condition – as $\mu_t(\cdot)$ does not depend on intermediate prices–, ruling out cross-sectional variation in input prices and thus ruling out any identifying variation for the second stage – given (3) and given other inputs are conditioned upon.

---

[11]Effectively, ACF replaces the assumption on measurement error with the assumptions on the proxy.

Testing the Validity of the ACF Proxy

Under the assumption of correct production function specification, the validity of proxy $h_t(k_{it}, \ell_{it}, m_{it})$ can be tested by **including lags of the inputs in the first stage equation**: if their associated coefficients result significantly different from zero, this can be interpreted as evidence of the proxy conditions not holding, as lags should be uninformative about a_{it} conditional on the proxy fully controlling for it.

ACF under Leontief Production Function

Leontief production function $Q_{it} = \min\{\theta_V V_{it}, \theta_M M_{it}\}$ – where V is value added and M are intermediate goods –; cost minimisation implies $Q_{it} = \theta_V V_{it} = \theta_M M_{it}$.

In logarithms, $m_{it} = v_{it} + (\ln \theta_V - \ln \theta_M)$; hence the Leontief already assumes $m_{it} = \mu(K_{it}, L_{it}, A_{it})$. In particular, $\tilde{v}_{it} = (\ln \theta_V - \ln \theta_M) - m_{it} + \varepsilon_{it}$ and the proxy function takes a particularly simple functional form in this case. A regression of $\tilde{v}_{it}$ on a constant and m_{it} should return a coefficient of 1: if this hypothesis is rejected, it could either be that the production function is not Leontief or m_{it} is not a valid proxy. Furthermore, if the production function is assumed to be Leontief there is **no need for a two-step procedure**: including m_{it} and a constant in the second stage specification suffices under maintained assumptions. Inference is furthermore standard as no regressor is generated in the first stage.

Finally, note that the Leontief is a particular case in which analysing quantity or value added does not matter, as $Q_{it} = \theta_V V_{it} \Leftrightarrow q_{it} = \ln \theta_V + v_{it}$.

Coda: Gross Output Production Functions with Flexible Inputs

The cost share approach to estimating β for flexible inputs holds both for Leontief and Cobb-Douglas production functions (assuming firms are price takers):

$$\frac{P_{it}^M M_{it}}{P_{it} Q_{it}} = \beta_M \qquad \frac{P_{it}^L L_{it}}{P_{it} Q_{it}} = \beta_L$$

which together with an assumption on the returns to scale to the industry yield an estimate of the β vector that does not require estimating the production function. If one is uncomfortable at the idea of not running even a single `reg y x, robust` in Stata, for any flexible input j $\hat{\beta}_J j_{it}$ can be subtracted from $\tilde{v}_{it}$ and, after appropriate differencing that depends on assumed structure of a_{it}, orthogonality conditions can be formed following BB.

The identifying variation depends on available data:

- **quantity data**: the parameter is identified with sufficient persistence of the real price of the flexible input and/or with asymmetric adjustment costs;

- **expenditure data**: the parameter is identified under the assumption of asymmetric adjustment costs – cross-sectional price variation is inconsistent with the idea that market prices provide an appropriate weight to aggregate sub-classes of the input.

As noted above, ACF does not apply directly to quantity production function estimation: however, ACF's algorithm can be modified by

1. estimating β_M, β_L using the cost share approach;

2. residualising $\tilde{q}_{it}$ to $\tilde{q}_{it} - \hat{\beta}_M m_{it} - \hat{\beta}_L \ell_{it}$;

3. estimating β_K (assuming K is predetermined and M, L are flexible) using BB given the assumption on a_{it}.

this is because, although in ACF with quantities $\mu_t(k_{it}, \ell_{it}, a_{it})$ must be interpreted as an optimality condition and hence no input price variation is permitted, such **price variation is here not necessary** – coefficients on flexible inputs are estimated via the cost share approach, making the 'optimality condition' interpretation not problematic.

Markup Estimation: A Primer

$$\frac{P}{MC} \equiv \mu > 1 \tag{16}$$

Under monopolistic competition, $\pi_M \equiv p(q)q - c(q)$, hence

$$\frac{\partial \pi_M}{\partial q} = 0 \Rightarrow p + \frac{\partial p(q)}{\partial q} q = \frac{\partial c(q)}{\partial q} \Leftrightarrow \frac{p(q) - c'(q)}{p(q)} = -\frac{\partial p(q)}{\partial q} \frac{q}{p} \equiv \frac{1}{|\eta|} \Leftrightarrow \mu = \frac{\eta}{\eta - 1} \tag{17}$$

Under pricing power, optimal price depends on output which in turn depends on TFP; hence the **difference between output quantity and revenue measures of output value is important**:

$$\text{Output Elasticity: } \frac{\partial \ln Q_{it}}{\partial \ln M_{it}} = \beta_M \qquad \text{Revenue Elasticity: } \frac{\partial \ln P_{it} Q_{it}}{\partial \ln M_{it}} = \gamma_M$$

We keep assuming that **firms are price takers in input markets** and also assume that input prices are common in the cross-section.

Klette and Griliches (1996)

$$q_{it} = \beta_0 + \beta_K k_{it} + \beta_L \ell_{it} + \beta_M m_{it} + a_{it} \tag{18}$$

with constant elasticity of demand and real prices defined as

$$Q_{it} = Q_t \Big(\frac{P_{it}}{P_t}\Big)^{-\eta} \qquad R_{it} = \frac{P_0}{P_t}(P_{it}Q_{it}) \tag{19}$$

taking logs and rearranging, we have

$$q_{it} - q_t = -\eta(p_{it} - p_t) \Leftrightarrow r_{it} = -\frac{1}{\eta}(q_{it} - q_t) + p_0 + q_{it} \Leftrightarrow r_{it} = \frac{\eta - 1}{\eta} q_{it} + \frac{1}{\eta} q_t + p_0 \tag{20}$$

and, for $\gamma_J \equiv \frac{\eta - 1}{\eta} \beta_J \equiv \frac{\beta_J}{\mu} \leq \beta_J$,

$$r_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_L \ell_{it} + \gamma_M m_{it} + \frac{1}{\eta} q_t + p_0 + \frac{\eta - 1}{\eta} a_{it} \tag{21}$$

which shows that **revenue elasticities will, in this case, underestimate output elasticities** from the Cobb Douglas. While one can in principle include the log of (sector) output to identify η and then estimate output elasticities, this precludes the inclusion of time fixed effects – which would not allow for identification of η.[12]

---

[12]Identification can be re-established via knowledge of the returns to scale – see this post.

De Loecker and Warzynski (2012)

The authors show that, if inputs are flexible and from cost minimisation,

$$\mu = \beta_{it}^M \underbrace{\frac{P_{it}Q_{it}}{P_{it}^M M_{it}}}_{1/\alpha_{it}^M} \tag{22}$$

which, from the result in Klette and Griliches that $\gamma_J = \frac{\beta_J}{\mu}$, implies

$$\frac{\gamma_M}{\alpha_{it}^M} = \frac{\beta_{it}^M/\mu}{\beta_{it}^M/\mu} = 1 \tag{23}$$

hence, under Cobb Douglas and constant elasticity of demand, the ratio between revenue elasticity and the cost share of the flexible input provides **no information** regarding the markup – if the parametric assumptions are correct, the ratio should equal 1 for any value of cost shares and elasticities.

DLW(2012) then consider the relationship between estimated markup $\hat\mu$ and exporting status in each sector, where markup is estimated as above:

$$\ln \hat\mu = \delta_0 + \delta_1 D_{it} + \varepsilon_{it} \text{ where } \hat\mu = \frac{\beta_{jt}^M}{\alpha_{it}^M} \tag{24}$$

for firm i in sector j. Given $\ln \hat\mu = \ln \beta_{jt}^M - \ln \alpha_{it}^M$, inclusion of a sector-specific constant will control of any (log-)additive error in estimated sector output elasticity. Hence, δ_1 is consistently estimated under $\mathbb{E}(D_{it}\varepsilon_{it}) = 0$; however, sector-specific means of $\ln \mu_{it}$ will not be consistent unless one has consistent estimates of β_{it}^M. For DLW's analysis, it can be noted that the inclusion of sector dummies makes estimation **independent** of estimated sector quantity elasticities β_{jt}; hence, the negative of the estimated cost share can simply be used as dependent variable:

$$\ln \beta_{jt}^M - \ln \alpha_{it}^M = \ln \hat\mu = \delta_0 + \phi^\mathsf{T} \mathbb{1}_{jt} + \delta_1 D_{it} + \varepsilon_{it} \Leftrightarrow -\ln \alpha_{it}^M = \delta_0 + \phi^\mathsf{T} \mathbb{1}_{jt} + \delta_1 D_{it} + \varepsilon_{it}$$