Baumol and Bowen Cost Effects in Research Universities:
Econometric Appendix

A.1 Model Specification and Estimation

For Total Cost (*tc*) we specify the log-linear regression

$$\ln\left(tc_{it}\right) = \beta' x_{it} + \alpha_i + u_{it} \tag{1}$$

where α_i are individual effects, with $i = 1,\ldots,n$ and $t = 1,\ldots,T_i$. For the purpose of "deconstructing costs," and the results in Tables 3 and 4 of the paper, the parameters β are estimated using fixed effects. Denote the parameter estimates $\hat{\beta}$ and the robust cluster-corrected covariance matrix estimate $\hat{V}$.

A.2 The Partial Differentials

The log-linear specification implies the expectation

$$E\left(tc_{it}\right) = \exp\left(\beta' x_{it} + \alpha_i\right) E\left[\exp\left(u_{it}\right)\right] = \exp\left(\beta' x_{it}\right)\exp\left(\alpha_i\right) E\left[\exp\left(u_{it}\right)\right] \tag{2}$$

The predicted value should incorporate an estimate of $E\left[\exp\left(u_{it}\right)\right]$. We use the sample average of the fixed effects residuals $N^{-1}\sum_i \sum_t \exp\left(\hat{u}_{it}\right)$[4], where N is the total number of estimation sample observations. Other variants tried included the usual correction factor for the log-normal model $\exp\left(0.5\sigma^2\right)$, and also a group mean $\bar{\hat{u}}_{i\bullet} = T_i^{-1}\sum_t \exp\left(\hat{u}_{it}\right)$. Each of these corrections is very small and there were no meaningful differences among them in our calculations. Thus the predicted tc_{it} is

$$\widehat{tc}_{it} = \exp\left(\hat{\beta}' x_{it}\right)\exp\left(\hat{\alpha}_i\right)\left\{N^{-1}\sum_i \sum_t \exp\left(\hat{u}_{it}\right)\right\} \tag{3}$$

The total differential of $E\left(tc_{it}\right)$ is

---

[4] Cameron and Trivedi (2010, 108) .

$$dE\left(tc_{it}\right) = \exp\left(\beta' x_{it}\right)\exp\left(\alpha_i\right) E\left[\exp\left(u_{it}\right)\right]\left(\beta' dx_{it}\right) \tag{4}$$

We wish to compare outcomes in 1987 to those in 2005, and outcomes in 2008 to those in 2011. For each university *i*, we calculate

$$d(tc_i) = w_i \exp\left(\beta' x_{i,base}\right)\exp\left(\alpha_i\right) E\left[\exp\left(u_{it}\right)\right]\left(\beta' dx_i\right) \tag{5}$$

where $x_{i,base}$ are regressor values in the base year 1987 or 2008, and $dx_i = x_{i,2011} - x_{i,2008}$ or $dx_i = x_{i,2005} - x_{i,1987}$.[5] The differential is weighted by base year FTE student enrollment. Define

$$w_i = ftestu_{i,base} \Big/ \left(\sum_{i=1}^{n} ftestu_{i,base}\right) \tag{6}$$

Then

$$\begin{aligned} d(tc) &= \sum_{i=1}^{n} w_i \exp\left(\beta' x_{i,base}\right)\exp\left(\alpha_i\right) E\left[\exp\left(u_{it}\right)\right]\left(\beta' dx_i\right) \\ &= \sum_{i=1}^{n} c_i \exp\left(\beta' x_{i,base}\right)\left(\beta' dx_i\right) \end{aligned} \tag{7}$$

where $c_i = w_i \exp\left(\alpha_i\right) E\left[\exp\left(u_{it}\right)\right]$. The estimator of $d(tc)$ is

$$\widehat{d(tc)} = \sum_{i=1}^{n} \hat{c}_i \exp\left(\hat{\beta}' x_{i,base}\right)\left(\hat{\beta}' dx_i\right) \tag{8}$$

where $\hat{c}_i = w_i \exp\left(\hat{\alpha}_i\right)\left\{N^{-1}\sum_i \sum_t \exp\left(\hat{u}_{it}\right)\right\}$. Since $\widehat{d(tc)} = g\left(\hat{\beta}\right)$ is a nonlinear function of the estimator $\hat{\beta}$ inference uses the delta method[6]. The asymptotic distribution of the estimator in (8) is

$$g\left(\hat{\beta}\right) \overset{a}{\sim} N\left[g(\beta), JVJ'\right] \tag{9}$$

---

[5] Our estimation panel is unbalanced. However we use a common sample to compute the comparison values.

[6] William Greene (2012, Theorem D.22, 1086).

where $J = \partial g(\beta)/\partial\beta'$, so that the estimator of the asymptotic variance of $\widehat{d(tc)}$ is $\hat{V}_{d(tc)} = \hat{J}\hat{V}\hat{J}'$, with $\hat{J} = \partial g(\beta)/\partial\beta'\big|_{\beta=\hat{\beta}}$ and $\hat{V}$ is a robust cluster corrected covariance matrix of $\hat{\beta}$.[7]

Given the form of the differential in (8) the Jacobian is

$$\begin{aligned} J &= \sum_{i=1}^{n} c_i \left[\exp\left(\beta' x_{i,base}\right)\left(\beta' dx_i\right) x'_{i,base} + \exp\left(\beta' x_{i,base}\right) dx'_i\right] \\ &= \sum_{i=1}^{n} c_i \exp\left(\beta' x_{i,base}\right)\left[\left(\beta' dx_i\right) x'_{i,base} + dx'_i\right] \end{aligned} \tag{10}$$

A.3 Deconstructions

Rather than the total differential we consider partial differentials using subsets of the regressor differential dx_i by setting some of its elements to zero. Specifically, a partial differential for the incremental effects would involve subsets of the independent variables as described in Section 5 of the paper; Output effects, cost savings, etc. To compare the theories of Baumol (*bau*) and Bowen (*bow*) we compute differential estimates for each. The Baumol components are salary and benefits, so

$$\begin{aligned} d(tc)_{tot}^{bau} &= d(tc)_{sal}^{bau} + d(tc)_{ben}^{bau} \\ &= \sum_{i=1}^{n} c_i \exp\left(\beta' x_{i,base}\right)\left[\beta'\left(dx_{i,sal}^{bau} + dx_{i,ben}^{bau}\right)\right] \\ &= \sum_{i=1}^{n} c_i \exp\left(\beta' x_{i,base}\right)\left(\beta' dx_{i,tot}^{bau}\right) \end{aligned} \tag{11}$$

The Bowen components of cost are productivity, salary, benefits, revenue and governance, so

$$\begin{aligned} d(tc)_{tot}^{bow} &= d(tc)_{prod}^{bow} + d(tc)_{sal}^{bow} + d(tc)_{ben}^{bow} + d(tc)_{rev}^{bow} + d(tc)_{gov}^{bow} \\ &= \sum_{i=1}^{n} c_i \exp\left(\beta' x_{i,base}\right)\left(\beta' dx_i^{bow}\right) \end{aligned} \tag{12}$$

where

[7] Coefficient estimation was carried out using Stata 13.0. Subsequent calculations were carried out in SAS 9.3/IML.

$$dx_i^{bow} = dx_{i,prod}^{bow} + dx_{i,sal}^{bow} + dx_{i,ben}^{bow} + dx_{i,rev}^{bow} + dx_{i,gov}^{bow} \tag{13}$$

We would like to test the null and alternative hypotheses

$$\begin{aligned} H_0 &: d(tc)_{tot}^{bow} - h \cdot d(tc)_{tot}^{bau} = h(\beta) \le 0 \\ H_1 &: d(tc)_{tot}^{bow} - h \cdot d(tc)_{tot}^{bau} = h(\beta) > 0 \end{aligned} \tag{14}$$

The test statistic is

$$t = \hat{h}(\beta) \Big/ se\left[\hat{h}(\beta)\right] \tag{15}$$

The numerator is

$$\begin{aligned} \hat{h}(\beta) &= \sum\nolimits_{i=1}^{n} c_i \exp\left(\hat{\beta}' x_{i,base}\right)\left(\hat{\beta}' dx_i^{bow}\right) - h \cdot \left[\sum\nolimits_{i=1}^{n} c_i \exp\left(\hat{\beta}' x_{i,base}\right)\left(\hat{\beta}' dx_i^{bau}\right)\right] \\ &= \sum\nolimits_{i=1}^{n} c_i \exp\left(\hat{\beta}' x_{i,base}\right)\left[\hat{\beta}'\left(dx_i^{bow} - h \cdot dx_i^{bau}\right)\right] \\ &= \sum\nolimits_{i=1}^{n} c_i \exp\left(\hat{\beta}' x_{i,base}\right)\hat{\beta}' dx_i^{h} \end{aligned} \tag{16}$$

where $dx_i^h = dx_i^{bow} - h \cdot dx_i^{bau}$. The denominator of the *t*-statistic uses a variance calculation based on the delta method. Note that the form of the differential in (8) and (16) is the same, and thus the Jacobian matrix (10) is of the same form in both cases.

A.4 References

Cameron, A. Colin and Trivedi, Pravin K. (2010) *Microeconometrics Using Stata, Revised Edition*, College Station, TX: Stata Press.

Greene, William E. (2012) *Econometric Analysis, Seventh Edition*, Upper Saddle River, NJ: Prentice-Hall.

Judge, George G., Hill, R. Carter, Griffiths, William E., Lütkepohl, Helmut and Lee, Tsoung-Chao (1988) *Introduction to the Theory and Practice of Econometrics, Second Edition*, New York: John Wiley and Sons, Inc.

Wooldridge, Jeffrey M. (2010) Econometric Analysis of Cross Section and Panel Data, Second Edition, Cambridge, MA: The MIT Press.