Production function regressions, returns to scale, and externalities

Craig Burnside

The World Bank, Washington, DC 20433, USA

(Received February 1995; final version received March 1996)

Abstract

A number of recent papers have used simple linear regressions in an attempt to identify market structure, the extent of returns to scale, and possible external effects in U.S. manufacturing industries. The results obtained from these regressions have important implications for several branches of modern macroeconomics. As a result, the macro literature frequently cites specific numerical evidence from Caballero and Lyons (1992) and Hall (1990), which suggests that there are quantitatively significant increasing returns to scale, or external effects in U.S. manufacturing. In contrast, it is the argument of this paper that this evidence is not convincing. The most robust evidence suggests that the typical U.S. manufacturing industry displays constant returns with no external effects. On the other hand, there is significant heterogeneity across industries.

Key words: Cyclical utilization; External effects; Increasing returns; Instrumental variables

JEL classification: C33; D24; E32

1. Introduction

A number of recent papers have used simple linear regressions in an attempt to identify market structure, the extent of returns to scale, and possible external effects in U.S. manufacturing industries. These include, to name a few,

I thank Susanto Basu and Ricardo Caballero, who were very helpful in providing me with their data and with information on the procedures used in their papers. The data provided to me by Caballero are from Robert Hall, while the data provided to me by Basu are from Dale Jorgenson as described in Jorgenson, Gollop, and Fraumeni (1987) and Jorgenson (1990). Thanks are, therefore, also due to the originators of these data. In addition, I thank Susanto Basu, Jeff Campbell, Lawrence Christiano, Martin Eichenbaum, Sergio Rebelo, and an anonymous referee for discussions and comments

Bartelsman, Caballero, and Lyons (1994), Basu (1994), Basu and Fernald (1994, 1995a), Burnside, Eichenbaum, and Rebelo (1995a), Caballero and Lyons (1989, 1992), and Hall (1988, 1990). The results obtained from these regressions have important implications for several burgeoning literatures in macroeconomics. For example, following Romer (1986), there has been rising interest in models of growth with increasing returns to scale. Baxter and King (1992), Beaudry and Devereux (1993, 1994), Farmer and Guo (1994), Gali (1994), Hornstein (1993), and Kiyotaki (1988), to name just a few, have developed business cycle models which rely either on monopolistic competition, increasing returns, external effects, or some combination of these. Many theoretical models with multiple equilibria or indeterminacies – for example, Benhabib and Farmer (1994), Boldrin and Rustichini (1994), Gali (1992), and Weil (1989) – depend on the same features to obtain the multiplicity or indeterminacy. In many of these models, parameter values must lie in specific regions in order to obtain multiple or indeterminate equilibria. In other cases, the extent to which the predictions of the model are different from those of a constant returns–perfect competition model is tied directly to the parameterization of the model. As a result, many of these papers cite specific numerical evidence from Caballero and Lyons (1992) and Hall (1990), which suggests that there are quantitatively significant increasing returns to scale, or external effects in U.S. manufacturing. In contrast, it is the argument of this paper that this evidence is not convincing. In particular, here it will be argued that the most robust evidence suggests that the typical U.S. manufacturing industry displays constant returns with no external effects.

The evidence in Caballero and Lyons (1989, 1992) and Hall (1988, 1990) regarding returns to scale is based on instrumental variables regressions of industry-level value-added growth on industry level capital and labor input growth. Basu and Fernald (1994) have argued that this evidence is flawed due to misspecification under nonconstant returns or imperfect competition. Burnside, Eichenbaum, and Rebelo (1995a) argue that their results are due to a failure to allow for cyclical variation in the utilization of capital. Both sets of critics suggest that returns are close to being constant.

Evidence regarding external effects provided by Caballero and Lyons (1992) adds aggregate manufacturing value-added growth to these regressions. Basu and Fernald (1995a) have argued that Caballero and Lyons' (1992) evidence in favor of external effects is flawed, again due to misspecification under nonconstant returns or imperfect competition. Unfortunately, once they correct for these difficulties, their evidence is mixed. Gross output regressions suggest that there are no externalities, while modified doubly-deflated value-added regressions suggest that there are large external effects.

The argument of this paper is different. Here, it is argued that previous regressions may have been misleading for one of three reasons. First, most regression-based evidence is obtained by imposing cross-industry equality restrictions on parameters. These restrictions are useful, in that they lead to a single-parameter estimate which summarizes the entire manufacturing sector. Unfortunately, in almost every case, these restrictions are strongly rejected when tested. In some cases, restricted estimates tend to be upwardly biased relative to various summary statistics for unrestricted estimates. Furthermore, imposing the restriction eliminates the fundamental heterogeneity across industries.

Second, comparisons among the different sets of regression evidence are made more difficult by differences in the choice of sample and instruments. Point estimates are surprisingly sensitive to these choices yet they have received little attention from the literature. External effects regressions are particularly sensitive to the instrument set. This is argued to be the result of both the instruments and the measure of the external effect being highly correlated with aggregate business cycle dynamics.

Third, it is argued, as in Burnside, Eichenbaum, and Rebelo (1995a), that correcting for cyclical variation in the utilization of capital has a significant impact on point estimates of returns to scale and externality parameters. In particular, simple returns to scale regressions suggest that returns are close to constant. Furthermore, once external effects are allowed for, they appear to be either nonexistent or small. In no case are aggregate returns found to be increasing.

The next section outlines the theoretical bases of the regressions found in the literature. Section 3, using the data used by Hall (1988, 1990) and Caballero and Lyons (1989, 1992) as well as the data used by Basu and Fernald (1994, 1995a), contrasts restricted and unrestricted results obtained using three-stage least squares (3SLS) estimators. It also presents statistics which summarize the unrestricted results. Section 4 assesses the robustness of the results to different sample periods, instrument sets, and levels of aggregation. Section 5 summarizes the paper.

2. Production function regressions

This section describes a series of alternative specifications of the production function which will be considered in the empirical work. We begin with a general specification and work toward more specific special cases. In the most general case, assume that at the industry level, gross output is given by

$$Y_{it} = G(K_{it}, L_{it}, E_{it}, M_{it}, A_{it}), \tag{2.1}$$

where Y_{it} is the gross output of industry i at time t, K_{it} is capital services used by industry i at time t, L_{it} is labor services used by industry i at time t, E_{it} is energy used by industry i at time t, M_{it} is materials used by industry i at time t, and A_{it} is the level of technology in industry i at time t. At this point we make no specific assumptions about the function G, although the reader may assume that all functions are differentiable unless it is stated otherwise.

2.1. Caballero and Lyons' specification

Caballero and Lyons (1992) assume that the production function for gross output can be written as a function of value-added, energy, and materials,

$$Y_{it} = G(V_{it}, E_{it}, M_{it}),$$

where value-added at time t is

$$V_{it} = F(K_{it}, L_{it}, A_{it}). \tag{2.2}$$

The function F is assumed to be homogeneous of degree γ^F in (K, L) and of degree 1 in A. Furthermore, G is homogeneous of degree 1 in (V, E, M).

Caballero and Lyons (1992) also implicitly assume that it is appropriate to use a measure of value-added obtained by subtracting real materials inputs from real gross output (doubly-deflated value-added). As discussed by Basu and Fernald (1995a) this requires that G displays either (i) Leontief aggregation or (ii) Hicks aggregation.

Given these assumptions, when we totally differentiate (2.2) we obtain

$$\mathrm{d}V_{it} = F_{Kit}\mathrm{d}K_{it} + F_{Lit}\mathrm{d}L_{it} + (V_{it}/A_{it})\,\mathrm{d}A_{it}.$$

Divide through by V_{it} to obtain

$$\frac{\mathrm{d}V_{it}}{V_{it}} = \frac{F_{Kit}K_{it}}{V_{it}}\frac{\mathrm{d}K_{it}}{K_{it}} + \frac{F_{Lit}L_{it}}{V_{it}}\frac{\mathrm{d}L_{it}}{L_{it}} + \frac{\mathrm{d}A_{it}}{A_{it}}.$$

Let $\Delta x_{it} = \mathrm{d}X_{it}/X_{it}$ for any variable X_{it}. In the empirical work, Δx_{it} is associated with $\Delta \ln(X_{it})$. Using this notation we have

$$\Delta v_{it} = \frac{F_{Kit}K_{it}}{V_{it}}\Delta k_{it} + \frac{F_{Lit}L_{it}}{V_{it}}\Delta l_{it} + \Delta a_{it}. \tag{2.3}$$

With the further assumption that factor markets are perfectly competitive, capital and labor are each paid the marginal cost of producing a unit of value-added times their marginal value-added products; i.e. $P_{Jt} = MC_{it}^F F_{Jit}$ for $J = K, L$. This implies that total value-added cost is given by

$$\begin{aligned} C_{it}^F &= P_{Kt}K_{it} + P_{Lt}L_{it} \\ &= MC_{it}^F\,(F_{Kit}K_{it} + F_{Lit}L_{it}) \\ &= MC_{it}^F\,(\gamma^F V_{it}), \end{aligned}$$

and that each term

$$\frac{F_{Jit}J_{it}}{V_{it}} = \frac{P_{Jt}J_{it}}{MC_{it}^F V_{it}} = \gamma^F\frac{P_{Jt}J_{it}}{C_{it}^F} = \gamma^F c_{Jit}^F,$$

where c^F_{Jit} denotes factor J's share in value-added costs in industry i at time t. So we can rewrite (2.3) as

$$\begin{aligned}\Delta v_{it} &= \gamma^F \left(c^F_{Kit}\Delta k_{it} + c^F_{Lit}\Delta l_{it}\right) + \Delta a_{it} \\ &= \gamma^F \Delta x^F_{it} + \Delta a_{it}.\end{aligned} \tag{2.4}$$

Caballero and Lyons (1992) estimate (2.4) using 3SLS by assuming that the stock of capital is the appropriate measure of capital services and by imputing the cost of capital in each period to obtain its cost share.

2.2. Basu and Fernald's specification

Basu and Fernald (1994) assume that the function G is differentiable and homogeneous of degree γ^G in (K,L,E,M) and, without loss of generality, that it is homogenous of degree 1 in A_{it}. With these assumptions and competitive factor markets, using similar arguments as above, we obtain

$$\begin{aligned}\Delta y_{it} &= \gamma^G \left(c^G_{Kit}\Delta k_{it} + c^G_{Lit}\Delta l_{it} + c^G_{Eit}\Delta e_{it} + c^G_{Mit}\Delta m_{it}\right) + \Delta a_{it} \\ &= \gamma^G \Delta x^G_{it} + \Delta a_{it},\end{aligned} \tag{2.5}$$

where c^G_{Jit} now represents factor J's share in total gross output costs.

Basu and Fernald (1994) estimate (2.5) using 3SLS by assuming that the stock of capital is the appropriate measure of capital services and by imputing the cost of capital in each period. They use the simple average of the cost shares at $t-1$ and t in implementing the local approximation implicit in (2.5).

It is difficult to compare estimates of γ^G obtained by Basu and Fernald with those of γ^F obtained by Caballero and Lyons. These comparisons must be made industry-by-industry, depending on the shares of energy and materials in costs because

$$\gamma^F = \gamma^G \left(1 - c^G_M - c^G_E\right) / \left[1 - \gamma^G \left(c^G_M + c^G_E\right)\right].$$

As Basu and Fernald (1994) and Rotemberg and Woodford (1995) point out, the relevant parameter in a representative agent model is γ^F. Of course, γ^F and γ^G are equal when $\gamma^G = 1$.

2.3. Burnside, Eichenbaum, and Rebelo's specification

Burnside, Eichenbaum, and Rebelo (1995a) argue that capital services are inappropriately measured using the stock of capital. Instead, they argue that capital services should be measured by effective units of capital used in production. In one of their specifications they assume a strong complementarity between electricity use and capital hours in producing capital services. Here I interpret their specification as applying to energy use. In their most restricted specification the production function for gross output would be

$$Y_{it} = \min(a_w W_{it}, a_m M_{it}),$$

where a_m and a_v are positive constants,

$$W_{it} = A_{it} H(L_{it}, S_{it}),$$

and

$$S_{it} = \min(a_k K_{it} H_{it}, a_e E_{it}),$$

where H_{it} is the number of hours (or effective hours) over which each unit of capital is used in a given period of time. In this specification, W_{it} represents a value-added like aggregate of labor, L_{it}, and capital services, S_{it}. Capital services are a Leontief function of capital hours and energy use.[1] The function H is assumed to be homogeneous of degree γ^H in (L, S). Given these assumptions and competitive factor markets, we have

$$\begin{aligned} \Delta y_{it} &= \Delta w_{it} \\ &= \gamma^H \left(c^H_{Lit} \Delta l_{it} + c^H_{Sit} \Delta s_{it}\right) + \Delta a_{it} \\ &= \gamma^H \left(c^H_{Lit} \Delta l_{it} + c^H_{Sit} \Delta e_{it}\right) + \Delta a_{it} \\ &= \gamma^H \Delta x^H_{it} + \Delta a_{it}, \end{aligned} \tag{2.6}$$

where c^H_{Jit} is factor J's share in total gross output costs net of materials costs.

Below, (2.6) is estimated using 3SLS.

2.4. Specifications with external effects

Each of the previous specifications can be modified to allow for an external effect in production. Authors differ as to how to model external effects. When working with two-digit manufacturing sector data, Caballero and Lyons (1989) use aggregate manufacturing input as an index for the external effect. Caballero and Lyons (1992) use aggregate manufacturing output as the index for the external effect. One reason for using the input-based measure rather than the output-based one is that spurious external effects can result in industries for which industry output is a large share of the total. For this reason, this paper consistently uses an aggregate input measure as the externality index.

Define $\Delta x^F_t = \sum_i \Delta x^F_{it}$, $\Delta x^G_t = \sum_i \Delta x^G_{it}$, and $\Delta x^H_t = \sum_i \Delta x^H_{it}$. The typical assumption is that the production function being estimated displays a constant

[1] In a related paper, Abbott, Griliches, and Hausman (1988) adjust for capital utilization by assuming that capital hours are related to hours per worker.

elasticity with respect to the externality index. As a result this paper considers the following specifications:

$$\Delta v_{it} = \gamma^F \Delta x_{it}^F + \eta^F \Delta x_t^F + \Delta a_{it}, \tag{2.7}$$

$$\Delta y_{it} = \gamma^G \Delta x_{it}^G + \eta^G \Delta x_t^G + \Delta a_{it}, \tag{2.8}$$

$$\Delta y_{it} = \gamma^H \Delta x_{it}^H + \eta^H \Delta x_t^H + \Delta a_{it}. \tag{2.9}$$

3. Estimates of returns to scale and external effects

This section presents estimates of Eqs. (2.4)–(2.9). These estimates are obtained using standard 3SLS methods. In each case, estimates obtained by imposing cross-industry restrictions on the parameters are compared to unrestricted estimates. The cross-industry restrictions are tested, and in most cases strongly rejected. As a result, alternative summary statistics for the unrestricted estimates are used to assess the robustness of inference to relaxation of the restriction. In some cases, returns to scale appear smaller when the restriction is relaxed.

3.1. Data

The regressions presented in this section are performed using two data sets. The first is Hall's (1988) data on a panel of 20 two-digit SIC code level manufacturing industries. The data are annual over the period 1953–1984. The data set includes information on value-added, labor input, and the capital stock. Given that Caballero and Lyons (1992) also use this data, it is used here to estimate Eqs. (2.4) and (2.7).

The second data set, which covers the same industries (with some definitional differences) and the same sample period, is described in detail in Jorgenson, Gollop, and Fraumeni (1987). It contains information on gross output, labor input, the stock of capital, energy use, and materials inputs. These data are used by Basu and Fernald (1994, 1995a) and are used here to estimate Eqs. (2.5), (2.6), (2.8), and (2.9).

To use 3SLS a set of instruments must be found for the right-hand side variables in these equations. The baseline set of instruments is the same as that found in Basu and Fernald (1995a). It consists of the growth rate of real military purchases of goods and services, the growth rate of the world price of oil, a dummy variable representing the political party of the president, and one lag of each of these variables. These instruments have been extensively discussed in the literature. Alternative instruments are discussed below.

3.2. Value-added regressions

This section presents estimates of Eqs. (2.4) and (2.7).

Table 1
Value-added returns to scale regressions

Industry	$\hat{\gamma}^F$	Industry	$\hat{\gamma}^F$
20	−0.13 (0.59)	30	0.97 (0.11)
21	−0.38 (0.63)	31	0.73 (0.32)
22	0.73 (0.28)	32	1.06 (0.11)
23	0.77 (0.16)	33	2.12* (0.17)
24	1.31 (0.25)	34	1.22 (0.34)
25	1.19 (0.26)	35	1.18 (0.16)
26	1.31 (0.38)	36	1.09 (0.13)
27	1.19 (0.26)	37	1.04 (0.25)
28	0.64 (0.43)	38	1.11 (0.24)
29	−1.34 (0.35)	39	0.09 (0.45)
Summary statistics			
$\hat{\gamma}^F_{\text{med}}$		1.02	
$\bar{\gamma}^F$		0.79 (0.09)	
$\tilde{\gamma}^F$		0.90 (0.11)	
$\bar{\sigma}_\gamma$		0.72 (0.09)	
$\tilde{\sigma}_\gamma$		0.67 (0.11)	
Restricted estimates			
$\hat{\gamma}^F$		1.12* (0.05)	

$\hat{\gamma}^F_{\text{med}}$, $\bar{\gamma}^F$, and $\tilde{\gamma}^F$ are the median, mean, and weighted averages of the industry level estimates of γ^F. $\bar{\sigma}_\gamma$ and $\tilde{\sigma}_\gamma$ are the measures of dispersion across the industry level estimates of γ^F as described in the text.

* Significantly greater than 1 at the 5% level.

Table 1 presents estimates of γ_i^F based on Eq. (2.4) which range from a low of -1.34 in industry 29 (petroleum refining) to a high of 2.12 in industry 33 (primary metals). The last row of Table 1 presents the estimate obtained when we impose the restriction $\gamma_i^F = \gamma^F$ for all i. This estimate is 1.12 with a standard error of 0.05. By imposing the restriction we find significantly increasing returns. However, returning to the unrestricted estimates we find only one industry (33, primary metals) in which γ_i^F is individually significantly greater than 1. Furthermore, the median of the estimated γ_i^F's is $\hat{\gamma}_{\text{med}}^F = 1.02$. Define

$$\bar{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\hat{\gamma}_i, \qquad \tilde{\gamma} = \sum_{i=1}^{N} s_i\hat{\gamma}_i,$$

where s_i is the average share of industry i in total manufacturing value-added over the sample period. The mean is $\bar{\gamma}^F = 0.79$, and the weighted average (based on value-added) is $\tilde{\gamma}^F = 0.90$. These results are more suggestive of constant or decreasing returns.

Which set of results is more credible? Imposing the restriction lends a lot of precision to the estimates, but the restriction is strongly rejected when tested. The $\chi^2(19)$ statistic for $\mathrm{H}_0\colon \gamma_i^F = \gamma^F$ for all i is 184.9 with a p-value of 0.000. This suggests that imposing the restriction could be highly misleading. Burnside (1994) shows that the restricted estimate of γ^F is approximately a weighted average of the unrestricted estimates. When there is no covariance across equations, the weight given to $\hat{\gamma}_i^F$ is proportional to the inverse of its variance. This is exactly the right thing to do (in order to improve precision) when the restriction is valid. Even when the restriction is not valid, it is an interesting summary statistic which can be justified by a desire to weight industry level estimates according to their precision. The median, mean, and weighted average are provided here as alternatives, the latter providing an alternative that weights the estimates according to economic significance.

Analogous dispersion measures can be defined as

$$\bar{\sigma}_{\gamma}^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{\gamma}_i - \bar{\gamma})^2, \qquad \tilde{\sigma}_{\gamma}^2 = \sum_{i=1}^{N} s_i(\hat{\gamma}_i - \tilde{\gamma})^2.$$

One interpretation of the weighted dispersion measure is that it would be the population variance of the returns to scale parameter measured across firms if all firms produced the same amount of value-added but industries had different numbers of firms. The dispersion measures are also given in Table 1.

The dispersion measures highlight another problem: that imposing H_0 leads to a false sense of precision. The restricted estimate of 1.12 has an estimated standard error of 0.05. This leads one to believe that there is significant evidence of

Table 2
Value-added external effects regressions

Industry	$\hat{\gamma}^F$	$\hat{\eta}^F$	Industry	$\hat{\gamma}^F$	$\hat{\eta}^F$
20	−1.42 (0.93)	0.49 (0.31)	30	0.28 (0.29)	1.17† (0.48)
21	0.87 (0.74)	0.50 (0.37)	31	0.79 (0.50)	0.01 (0.48)
22	2.74* (0.98)	−2.18 (1.27)	32	1.56* (0.24)	−0.35 (0.36)
23	0.91 (0.38)	−0.10 (0.42)	33	2.80* (0.40)	−1.22 (0.73)
24	1.00 (0.36)	0.49 (0.65)	34	1.68 (1.33)	−0.48 (1.96)
25	1.61* (0.37)	−0.87 (0.63)	35	0.97 (0.27)	0.33 (0.50)
26	2.03 (1.11)	−0.41 (0.84)	36	−0.32 (0.28)	2.08† (0.46)
27	−0.23 (0.41)	0.66† (0.23)	37	0.40 (0.39)	1.20† (0.72)
28	−2.21 (0.87)	1.67† (0.57)	38	0.41 (0.35)	0.76† (0.40)
29	−1.21 (0.37)	0.17 (0.38)	39	−0.66 (0.92)	0.92 (1.02)
Summary statistics					
$\hat{\gamma}^F_{\text{med}}$		0.83	$\hat{\eta}^F_{\text{med}}$		0.41
$\bar{\gamma}^F$		0.60 (0.14)	$\bar{\eta}^F$		0.24 (0.17)
$\tilde{\gamma}^F$		0.44 (0.16)	$\tilde{\eta}^F$		0.42† (0.21)
$\bar{\sigma}_\gamma$		1.30 (0.21)	$\bar{\sigma}_\eta$		0.97 (0.22)
$\tilde{\sigma}_\gamma$		1.41 (0.24)	$\tilde{\sigma}_\eta$		1.01 (0.22)
Restricted estimates					
$\hat{\gamma}^F$		0.94 (0.06)	$\hat{\eta}_F$		0.33† (0.09)

$\hat{\gamma}^F_{\text{med}}$, $\bar{\gamma}^F$ and $\tilde{\gamma}^F$ are the median, mean, and weighted averages of the industry level estimates of γ^F. $\hat{\eta}^F_{\text{med}}$, $\bar{\eta}^F$, and $\tilde{\eta}^F$ are the median, mean, and weighted averages of the industry level estimates of η^F. $\bar{\sigma}_\gamma$, $\tilde{\sigma}_\gamma$, $\bar{\sigma}_\eta$, and $\tilde{\sigma}_\eta$ are the measures of dispersion across the industry level estimates of γ^F and η^F as described in the text.

* Significantly greater than 1 at the 5% level. † Significantly greater than 0 at the 5% level.

increasing returns. However, H_0 is rejected. The median standard error of the unrestricted estimates is 0.26. The simple and weighted averages of the unrestricted estimates have standard errors of 0.09 and 0.11, respectively. Furthermore, the dispersion measures are $\bar{\sigma}_\gamma = 0.72$ and $\tilde{\sigma}_\gamma = 0.67$, which indicates that there is considerable dispersion across the 3SLS estimates of γ^F. Now one is led to conclude that the typical industry displays constant returns, but that returns to scale varies greatly across industries.

Table 2 provides estimates of (2.7). The restricted estimate of γ^F is 0.94 (and insignificantly different than 1) while the restricted estimate of η^F is 0.33 (and significantly greater than 0). The degree of overall returns to scale at the aggregate level is $\gamma^F/\left(1-\eta^F\right) = 1.40$. Summing up, the restricted results suggest that manufacturing is typified by close to constant internal returns with strong external effects and increasing returns in the aggregate.

Again, however, the cross-industry restrictions on the parameters are rejected strongly. The $\chi^2(38)$ statistic for the joint hypothesis that $\gamma_i^F = \gamma^F$ and $\eta_i^F = \eta^F$ for all i is 217.7 with a p-value of 0.000. The individual restrictions are also strongly rejected, with p-values of 0.000 and 0.002, respectively. When the restrictions are relaxed, we find quite different results. First, as shown in Table 2, the median, mean, and weighted average of the unrestricted estimates of γ^F are much smaller than the restricted estimate and imply sharply decreasing internal returns. On the other hand, the median, mean, and weighted average of the unrestricted estimates of η^F are in the neighbourhood of the restricted estimate. Taken together, though, the unrestricted results are puzzling: they imply roughly constant returns in the aggregate with sharply decreasing returns at the industry level. Importantly, the unrestricted estimates do not allow for the sharp inferences about the parameters obtained when the restriction was imposed. They are highly dispersed and estimated very imprecisely.

Results obtained using value-added data are also sensitive to the instrument set and the sample. This is discussed below.

3.3. Gross output regressions

This section presents results obtained using the second data set described above. The instruments used in the regressions are identical to the ones used in the value-added regressions.

Rather than present an entire set of results, Table 3 presents summary statistics for the unrestricted regressions and the restricted results for Eq. (2.5). The restricted estimate of γ^G is 0.96 with a standard error of 0.02. A similar finding is the basis of Basu and Fernald's (1994) conclusion that a failure to model gross output more generally is responsible for the finding of increasing returns with value-added data.

However, once again the restriction is rejected with a $\chi^2(19)$ statistic of 332.4 and a p-value of 0.000. The mean ($\bar{\gamma}^G = 0.88$) and weighted average ($\tilde{\gamma}^G = 0.87$)

Table 3
Gross output returns to scale regressions

Summary statistics for unrestricted estimates	
$\hat{\gamma}^G_{\text{med}}$	0.96
$\bar{\gamma}^G$	0.88 (0.03)
$\tilde{\gamma}^G$	0.87 (0.05)
$\bar{\sigma}_\gamma$	0.34 (0.04)
$\tilde{\sigma}_\gamma$	0.38 (0.06)
Restricted estimates	
$\hat{\gamma}^G$	0.96 (0.02)

$\hat{\gamma}^G_{\text{med}}$, $\bar{\gamma}^G$, and $\tilde{\gamma}^G$ are the median, mean, and weighted averages of the industry level estimates of γ^G. $\bar{\sigma}_\gamma$ and $\tilde{\sigma}_\gamma$ are the measures of dispersion across the industry level estimates of γ^G as described in the text.

of the unrestricted estimates of γ^G are lower than the restricted estimate, while the median $\hat{\gamma}^G_{\text{med}} = 0.96$ is larger. Certainly, Basu and Fernald's basic inference seems intact; there is little evidence here in favor of increasing returns. On the other hand, the unrestricted estimates are sufficiently dispersed that for 4 of the 20 industries, returns to scale are significantly increasing (in contrast to only 1 from the value-added regressions).

Again, there is some dispersion of results across the industries. The measures of spread are $\bar{\sigma}_\gamma = 0.34$ and $\tilde{\sigma}_\gamma = 0.37$, indicating considerably less variety across the estimates of γ^G than in the corresponding value-added regressions. The typical industry appears to have roughly constant or slightly decreasing returns to scale.[2]

Turning to Eq. (2.8), Table 4 presents estimates of γ^G and η^G. First, consider the restricted estimates. These estimates are somewhat similar to the ones found by Basu and Fernald (1995a), although they obtain a much smaller estimate of η^G.[3] The returns to scale parameter $\hat{\gamma}^G = 0.90$ with a standard error of 0.02, while $\hat{\eta}^G = 0.10$ with a standard error of 0.03. These estimates suggest that overall returns to scale $\gamma^G/(1-\eta^G)$ are almost exactly constant, while internal

[2] It is important to note that since the restricted estimate and the summary statistics of the unrestricted estimates of γ^G are all less than 1, the implied values of γ^F are marginally smaller. So the gross output specification is not automatically decreasing the degree of returns to scale as it does when the estimates are greater than 1.

[3] The differences stem from the fact that their sample includes 1985 and they keep the automobile component of industry 37 as a separate industry.

Table 4
Gross output external effects regressions

Summary statistics for unrestricted estimates			
$\hat{\gamma}^G_{\text{med}}$	0.89	$\hat{\eta}^G_{\text{med}}$	0.15
$\bar{\gamma}^G$	0.85 (0.05)	$\bar{\eta}^G$	0.08 (0.05)
$\tilde{\gamma}^G$	0.83 (0.07)	$\tilde{\eta}^G$	0.11† (0.05)
$\bar{\sigma}_\gamma$	0.49 (0.05)	$\bar{\sigma}_\eta$	0.47 (0.06)
$\tilde{\sigma}_\gamma$	0.46 (0.06)	$\tilde{\sigma}_\eta$	0.44 (0.05)
Restricted estimates			
$\hat{\gamma}^G$	0.90 (0.02)	$\hat{\eta}^G$	0.10† (0.03)

$\hat{\gamma}^G_{\text{med}}$, $\bar{\gamma}^G$, and $\tilde{\gamma}^G$ are the median, mean, and weighted averages of the industry level estimates of γ^G. $\hat{\eta}^G_{\text{med}}$, $\bar{\eta}^G$, and $\tilde{\eta}^G$ are the median, mean, and weighted averages of the industry level estimates of η^G. $\bar{\sigma}_\gamma$, $\tilde{\sigma}_\gamma$, $\bar{\sigma}_\eta$, and $\tilde{\sigma}_\eta$ are the measures of dispersion across the industry level estimates of γ^G and η^G as described in the text.

† Significantly greater than 0 at the 5% level.

returns to scale are decreasing. Using gross output reduces the magnitude of the external effect but does not eliminate it completely.

Once again, the cross-industry restrictions being imposed are overwhelmingly rejected by the data. The $\chi^2(38)$ statistic for the joint hypothesis that $\gamma_i^G = \gamma^G$ and $\eta_i^G = \eta^G$ for all i is 492.1 with a p-value of 0.000. The individual hypotheses are strongly rejected as well. However, in this case there appears to be little bias due to the imposition of the restriction. The mean, median, and weighted averages of the unrestricted coefficients are all close to the restricted estimates. The dispersion measures indicate that there is still considerable heterogeneity among the industries. In fact, there are four industries with significantly increasing internal returns and five industries with significant external effects. Interestingly, there is a strong negative correlation between $\hat{\gamma}_i^G$ and $\hat{\eta}_i^G$. Most industries with strong internal returns display negative external effects. To this extent the results are quite similar to the unrestricted findings with value-added data.

Do these results suggest that increasing returns are present in some industries while external effects are present in others? Probably not. Out of the four industries that display significantly increasing internal returns, only one does so in the absence of a negative external effect. Out of the five industries that display significantly positive external effects, only one does so in the absence of decreasing internal returns. This suggests that in the unrestricted regressions the first-stage fitted values of the regressors are highly collinear.

The sensitivity of the gross output based results to the instrument set and the sample is discussed below.

3.4. Gross output regressions with the energy correction

Finally, this section considers estimates of Eqs. (2.6) and (2.9) which are based on Burnside, Eichenbaum, and Rebelo's (1995a) correction for capital utilization. Table 5 presents results for the returns to scale Eq. (2.6). The restricted estimate of γ^H is 0.94 with a standard error of 0.03. It is insignificantly different from 1 at the 5% level. Once again, though, the cross-industry restriction is rejected with a $\chi^2(19)$ statistic of 96.3 and a p-value of 0.000. Again, we see the common pattern that $\hat{\gamma}^H_{\text{med}} = 0.90$, $\bar{\gamma}^H = 0.80$, and $\tilde{\gamma}^H = 0.89$ are all smaller than the restricted estimate. However, the finding that returns are nonincreasing is robust. The dispersion of the unrestricted estimates of γ^H is similar to what was found for the gross output regressions, with $\bar{\sigma}_\gamma = 0.40$ and $\tilde{\sigma}_\gamma = 0.38$.

Table 6 presents results for external effects and internal returns to scale. The restricted estimates are $\hat{\gamma}^H = 0.94$ and $\hat{\eta}^H = 0.00$ with standard errors both equal to 0.04. So Burnside, Eichenbaum, and Rebelo's correction for capital utilization appears to eliminate the external effect completely while preserving roughly constant returns.

Once more, however, the restriction across industries is rejected with a $\chi^2(38)$ statistic of 187.9 and a p-value of 0.000. As in the gross output case, though, inference is not dramatically affected by the restriction. The unrestricted returns to

Table 5
Energy-corrected returns to scale regressions

Summary statistics for unrestricted estimates	
$\hat{\gamma}_{\text{med}}$	0.90
$\bar{\gamma}$	0.80 (0.07)
$\tilde{\gamma}$	0.89 (0.06)
$\bar{\sigma}_\gamma$	0.40 (0.07)
$\tilde{\sigma}_\gamma$	0.38 (0.06)
Restricted estimates	
$\hat{\gamma}^H$	0.94 (0.03)

$\hat{\gamma}^H_{\text{med}}$, $\bar{\gamma}^H$, and $\tilde{\gamma}^H$ are the median, mean, and weighted averages of the industry level estimates of γ^H. $\bar{\sigma}_\gamma$, and $\tilde{\sigma}_\gamma$ are the measures of dispersion across the industry level estimates of γ^H as described in the text.

Table 6
Energy-corrected external effects regressions

Summary statistics for unrestricted estimates			
$\hat{\gamma}^H_{\text{med}}$	0.88	$\hat{\eta}^H_{\text{med}}$	-0.03
$\bar{\gamma}^H$	1.00 (0.11)	$\bar{\eta}^H$	-0.07 (0.12)
$\tilde{\gamma}^H$	1.00 (0.09)	$\tilde{\eta}^H$	0.02 (0.11)
$\bar{\sigma}_\gamma$	0.51 (0.11)	$\bar{\sigma}_\eta$	0.59 (0.13)
$\tilde{\sigma}_\gamma$	0.42 (0.09)	$\tilde{\sigma}_\eta$	0.53 (0.11)
Restricted estimates			
$\hat{\gamma}^H$	0.94 (0.04)	$\hat{\eta}^H$	0.00 (0.04)

$\hat{\gamma}^H_{\text{med}}$, $\bar{\gamma}^H$, and $\tilde{\gamma}^H$ are the median, mean, and weighted averages of the industry level estimates of γ^H. $\hat{\eta}^H_{\text{med}}$, $\bar{\eta}^H$, and $\tilde{\eta}^H$ are the median, mean, and weighted averages of the industry level estimates of η^H. $\bar{\sigma}_\gamma$, $\tilde{\sigma}_\gamma$, $\bar{\sigma}_\eta$, and $\tilde{\sigma}_\eta$ are the measures of dispersion across the industry level estimates of γ^H and η^H as described in the text.

scale parameter has median $\hat{\gamma}^H_{\text{med}} = 0.88$. Its mean is $\bar{\gamma}^H = 1.00$ and its weighted average is $\tilde{\gamma}^H = 1.00$ with standard errors of 0.11 and 0.09, respectively. The externality parameter has median $\hat{\eta}^H_{\text{med}} = -0.03$. Its mean is $\bar{\eta}^H = -0.07$ and its weighted average is $\tilde{\eta}^H = 0.00$ with standard errors of 0.12 and 0.11, respectively. The restriction only seems to play a role of eliminating the heterogeneity across industries rather than introducing a bias. The dispersion of the parameter estimates is similar to the gross output case.

Of the five industries that display apparent positive external effects, all display decreasing internal returns. Of the five industries that display increasing internal returns, all display negative external effects. So while there is heterogeneity in the sample, as in the gross output case, this is most likely due to a severe multicollinearity problem.

The results are moderately sensitive to the instrument set and the sample period as will be shown below.

4. Robustness

This section considers the sensitivity of the results presented above to two considerations: the instrument set and the sample period. In some cases the findings presented above are surprisingly sensitive to the choices made in regard to these considerations. In light of the fact that cross-industry restrictions are generally rejected, this section also discusses whether further disaggregation is necessary. In particular, are cross-industry restrictions to the two-digit level rejected when the data are measured at the three-digit level?

4.1. The sample period

To assess robustness, two sample periods are considered. The first is the baseline sample period of 1953–1984, from Basu and Fernald (1994). The second is 1959–1984, from Caballero and Lyons (1992). Table 7 compares estimates obtained using the two sample periods and the baseline instrument set.

Estimates of returns to scale are moderately sensitive to the sample period. In general, restricted point estimates based on the entire sample are larger, though the differences are quite small when compared with the standard errors. Much the same can be said of the median unrestricted estimates.

The external effects regressions are more sensitive to the sample period chosen. This is consistent with the argument of the next section, which is that separate returns to scale and external effects parameters are difficult to identify from these regressions. Thus, small changes in the sample period or the instrument set can move point estimates considerably. In general, the effect of the shorter sample period is to reduce the estimate of internal returns to scale while raising the magnitude of the external effect. However, in each case the overall degree of returns to scale, $\gamma/(1-\eta)$, is virtually unaffected.

Table 7
Alternative sample periods

Returns to scale regressions

	γ^F		γ^G		γ^H	
Sample	Restricted	Median	Restricted	Median	Restricted	Median
1953–84	1.12* (0.05)	1.02	0.96 (0.02)	0.96	0.94 (0.03)	0.90
1959–84	1.06 (0.04)	1.00	0.95 (0.01)	0.91	0.90 (0.03)	0.84

External effects regressions

	Restricted		Restricted		Restricted	
Sample	γ^F	η^F	γ^G	η^G	γ^H	η^H
1953–84	0.94 (0.06)	0.33† (0.09)	0.90 (0.02)	0.10† (0.03)	0.94 (0.04)	0.00 (0.04)
1959–84	0.84 (0.03)	0.44† (0.05)	0.83 (0.01)	0.21† (0.02)	0.82 (0.03)	0.11† (0.04)

* Significantly greater than 1 at the 5% level. † Significantly greater than 0 at the 5% level.

Rejection of the cross-industry restrictions is robust to the sample period. The p-values for all relevant restrictions are 0.000. Again, the main effect of the restriction seems to be to induce an upward bias in restricted estimates based on value-added.

4.2. The instrument set

Five alternative instrument sets are considered:

1. The instruments from Hall (1990): the growth rate of real military expenditure, the political party of the president, and the growth rate of the world price of oil, plus one lagged value of each of these variables (this is the baseline set for this paper and is the set used in Basu and Fernald, 1995a).
2. Only contemporaneous values of the Hall instruments.
3. One of the instrument sets used by Caballero and Lyons (1992): the growth rate of real military expenditure, the political party of the president, the growth rate of the world price of oil relative to the price of nondurables, and the growth rate of the world price of oil relative to the price of durables.
4. An instrument set similar to the one used by Burnside, Eichenbaum, and Rebelo (1995a): It consists of the current and lagged growth rates of the world price of oil, as well as four lagged values of a measure of monetary shocks. The measure of monetary shocks is based on a quarterly VAR model including the following variables: real GDP, the GDP deflator, the three-month T-bill rate, the monetary base, and an index of industrial commodity prices from the Journal of Commerce (mnemonics GDPQ, GDPD, FYGM3, FMBASE, and FCJM from Citibase).[4] The three-month T-bill rate is assumed to be the policy variable, determined according to lagged information on all five variables as well as contemporaneous information about GDP, the GDP deflator, and the industrial commodity price index. The error term from the fitted policy rule is the measure of the monetary shock.[5] As in Burnside, Eichenbaum, and Rebelo (1995a), for year t, the four shocks from year $t-1$ are used as instruments.
5. An instrument set consisting of current and three lagged values of the growth rate of the world oil price. This choice is motivated by the fact that the oil price instrument is the most powerful of the Hall instruments.

[4] Burnside, Eichenbaum, and Rebelo's measures were obtained from Christiano, Eichenbaum, and Evans (1994). Data availability prevented the use of the same variables used by them: real GDP, the GDP deflator, an index of sensitive prices, the Fed Funds rate, total reserves, and nonborrowed reserves. The VAR is estimated using four lags over the period 1949Q1–1992Q4.

[5] The measure of monetary shocks obtained here is highly correlated with Burnside, Eichenbaum, and Rebelo's over their common sample period.

Table 8
Alternative instrument sets

Returns to scale regressions

Instrument set	J	γ^F Restricted	γ^F Median	J	γ^G Restricted	γ^G Median	J	γ^H Restricted	γ^H Median
1	0.16	1.12* (0.05)	1.02	0.05	0.96 (0.02)	0.96	0.15	0.94 (0.03)	0.90
2	0.18	1.24* (0.07)	1.20	0.05	0.98 (0.02)	1.00	0.17	0.93 (0.04)	0.99
3	0.13	1.28* (0.07)	1.16	0.04	0.96 (0.02)	0.95	0.10	0.93 (0.04)	0.95
4	0.14	1.15* (0.05)	1.04	0.03	0.95 (0.01)	0.99	0.12	0.93 (0.03)	0.92
5	0.07	1.03 (0.06)	0.85	0.05	0.94 (0.02)	0.93	0.20	0.91 (0.04)	0.87

External effects regressions

Instrument set	J	Restricted γ^F	Restricted η^F	J	Restricted γ^G	Restricted η^G	J	Restricted γ^H	Restricted η^H
1	0.12	0.94 (0.06)	0.33† (0.09)	0.06	0.90 (0.02)	0.10† (0.03)	0.14	0.94 (0.04)	0.01 (0.04)
2	0.09	0.93 (0.07)	0.62† (0.18)	0.21	0.76 (0.03)	0.54† (0.09)	0.17	0.83 (0.05)	0.17† (0.07)
3	0.06	1.06 (0.07)	0.39† (0.17)	0.16	0.81 (0.03)	0.45† (0.08)	0.10	0.86 (0.05)	0.15† (0.07)
4	0.11	0.73 (0.05)	0.71† (0.07)	0.03	0.88 (0.02)	0.10† (0.03)	0.10	0.98 (0.04)	−0.05 (0.03)
5	0.10	0.79 (0.06)	0.52† (0.09)	0.06	0.84 (0.02)	0.13† (0.03)	0.20	0.86 (0.04)	0.07 (0.04)

The numbers reported in columns labelled J are the p-values associated with tests of the over-identifying restrictions from the restricted regressions.

* Significantly greater than 1 at the 5% level. † Significantly greater than 0 at the 5% level.

To assess robustness, all regressions were rerun using the different sets of instruments. The results are summarized in Table 8. They indicate that there is a great deal of sensitivity to the instruments. First, focusing on the returns to scale regressions we find that the value-added results are the most sensitive. Estimates based on the contemporaneous Hall instruments (set 2) and the Caballero and Lyons instruments (set 3) are much larger than estimates based on the other instruments. This is true for both restricted and unrestricted regressions. The gross output results are quite robust to the different instruments, while only the unrestricted energy-corrected regressions show much sensitivity, with sets 2 and 3 leading to larger median estimates of γ^H.

The external effects regressions are even more sensitive to the instrument set. In value-added regressions, estimates of internal returns vary from 0.73 to 1.06, while the external effect parameter varies from 0.33 to 0.71. The estimates imply overall returns to scale which vary from a low of 1.40 (set 1) to a high of 2.52 (set 2). For gross output, sets 1, 4, and 5 yield similar estimates, but sets 2 and 3 are consistent with much larger external effects than the other gross output regressions suggest. Overall returns to scale for sets 1, 4, and 5 are 1.00, 0.98, and 0.97, respectively, while for sets 2 and 3 they are 1.65 and 1.47. The results for the energy-corrected regressions are perhaps the most robust. They indicate no significant external effects for sets 1, 4, and 5, and significant, but small, external effects for sets 2 and 3. Overall returns to scale are never significantly increasing, ranging from a low of 0.92 to a high of 1.01.

One of the criteria in choosing instruments is validity. Validity can be verified by testing overidentifying restrictions when there are more instruments than regressors.[6] Table 8 reports the standard J statistic for testing the overidentifying restrictions in the restricted regressions. For value-added regressions all five instruments pass the test at the 5% level, except for instrument set 5 in the external effects regression where the p-value is 0.02. For gross output regressions most of the tests lead to marginal rejections, with p-values just below 0.05. For energy-corrected regressions there are no rejections at the 5% level. This finding, along with the robustness of the point estimates for energy-corrected regressions, suggests that we may have the most confidence in these results.

A second criterion in choosing instruments is relevance. The instruments must be correlated with the right-hand-side variables. In fact, the argument here is that it is the nature of the correlation between the instruments and the right-hand-side variables that is responsible for the disheartening lack of robustness in the external effects regressions. To measure relevance there are several alternatives. One possibility is to report the R^2 from the first stage regressions. This can be criticized on at least two grounds. First, in univariate regressions, this measure can be maximized by adding arbitrary instruments until a perfect fit is obtained, and the original regressor is obtained as the fitted value. For this reason, an adjusted measure such as $\bar{R}^2$ may be more desirable. Second, as Shea (1996) points out, in multivariate regressions, the R^2 for each right-hand-side variable may be high, even when the second-stage regression is poorly identified. An extreme example of this is when there are two right-hand-side variables and two instruments. Suppose one of these instruments is highly relevant for both variables, while the other is irrelevant to both. This will imply high R^2 in both first-stage regressions but a singularity in the second stage, since the two sets of fitted values will be perfectly correlated. Shea (1996) proposes a partial R^2 measure which adjusts for this issue, which is also discussed in Hall, Rudebusch,

[6] More correctly, such a test should be regarded as test of the entire model specification, including the choice of instruments.

and Wilcox (1994) and Staiger and Stock (1994).[7] The partial R^2 can also be adjusted for the number of instruments to obtain a partial $\bar{R}^2$.

Suppose there are two right-hand-side variables, X_1 and X_2, and a set of instruments, W. Suppose also, for simplicity, that each of these variables has mean zero. The standard R^2 measure for X_1 is the squared correlation between $\hat{X}_1 = W(W'W)^{-1}W'X_1 = P_W X_1$ and X_1:

$$R^2 = \frac{(X_1'\hat{X}_1)^2}{(X_1'X_1)(\hat{X}_1'\hat{X}_1)} = \frac{(X_1'P_W X_1)^2}{(X_1'X_1)(X_1'P_W X_1)} = \frac{X_1'P_W X_1}{X_1'X_1}.$$

The adjusted R^2 is $\bar{R}^2 = 1 - (1 - R^2)(T-1)/(T-k)$, where T is the sample size and k is the number of instruments. To find Shea's partial R^2 measure we replace X_1 and $\hat{X}_1$ by $\tilde{X}_1$ and $\bar{X}_1$, where $\tilde{X}_1$ is the portion of X_1 not explained by X_2 and $\bar{X}_1$ is the portion of $\hat{X}_1$ not explained by $\hat{X}_2$. This implies that $\tilde{X}_1 = M_2 X_1$ and $\bar{X}_1 = \hat{M}_2\hat{X}_1$, where $M_2 = I - X_2(X_2'X_2)^{-1}X_2'$ and $\hat{M}_2 = I - \hat{X}_2(\hat{X}_2'\hat{X}_2)^{-1}\hat{X}_2'$. Denoting the partial R^2 by R_p^2 we have

$$\begin{aligned} R^2 &= \frac{(\tilde{X}_1'\bar{X}_1)^2}{(\tilde{X}_1'\tilde{X}_1)(\bar{X}_1'\bar{X}_1)} \\ &= \frac{(X_1'M_2\hat{M}_2P_W X_1)^2}{(X_1'M_2X_1)(X_1'P_W\hat{M}_2P_W X_1)} \\ &= \frac{X_1'P_W\hat{M}_2P_W X_1}{X_1'M_2X_1}. \end{aligned}$$

This can also be adjusted for the number of instruments as $\bar{R}_p^2 = 1 - (1 - R_p^2)(T-1)/(T-k)$.

The motivation for these measures is as follows. In univariate regressions, for a given variance of the error term, the precision of the estimate of the slope coefficient depends only on the variance of the regressor. In bivariate regressions it depends on the variance of that portion of the regressor which cannot be explained by the other regressor. In univariate instrumental variables regressions the covariance between the regressor and its first-stage fitted values is what is relevant. This is captured in a scale-free way by the standard R^2 measure. However, in bivariate instrumental variables regressions what is relevant is the covariance between the portion of the regressor not explained by the other regressor and

[7] As these authors point out, it is difficult to use relevance measures as a selection criterion. Here, they are viewed more as a diagnostic device.

Table 9
Instrument relevance

Value-added data

Instrument set	Industry input growth				Aggregate input growth			
	Regular		Partial		Regular		Partial	
	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$
1	0.35	0.20	0.24	0.06	0.40	0.25	0.27	0.09
2	0.15	0.06	0.16	0.07	0.10	0.01	0.12	0.02
3	0.16	0.03	0.19	0.07	0.11	−0.02	0.14	0.02
4	0.44	0.30	0.21	0.03	0.50	0.38	0.26	0.08
5	0.30	0.20	0.18	0.06	0.30	0.19	0.19	0.07

Gross output data

Instrument set	Industry input growth				Aggregate input growth			
	Regular		Partial		Regular		Partial	
	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$
1	0.36	0.20	0.28	0.11	0.38	0.23	0.30	0.13
2	0.15	0.06	0.14	0.04	0.10	0.00	0.11	0.02
3	0.15	0.03	0.21	0.10	0.10	−0.04	0.14	0.01
4	0.49	0.36	0.29	0.12	0.57	0.47	0.34	0.18
5	0.35	0.26	0.23	0.12	0.38	0.30	0.25	0.14

Energy-corrected data

Instrument set	Industry input growth				Aggregate input growth			
	Regular		Partial		Regular		Partial	
	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$	R^2	$\bar{R}^2$	R_p^2	$\bar{R}_p^2$
1	0.37	0.22	0.24	0.06	0.40	0.26	0.29	0.12
2	0.17	0.08	0.13	0.04	0.15	0.08	0.15	0.06
3	0.16	0.03	0.17	0.05	0.13	0.03	0.17	0.05
4	0.54	0.43	0.22	0.03	0.63	0.54	0.27	0.09
5	0.37	0.28	0.19	0.06	0.37	0.27	0.20	0.08

the portion of its first-stage fitted values not explained by the other fitted values. This is captured in a scale-free way by the partial R^2 measure.

Table 9 presents measures of instrument relevance for the five instrument sets and three types of regressions. The measures are averages across the 20 industries. Regular R^2 measures are relevant for returns to scale regressions, whereas for external effects regressions partial R^2 measures are needed. On the basis of $\bar{R}^2$, Table 9 indicates that instrument sets 1, 4, and 5 provide the best fit, with instrument set 4 coming out on top. It is perhaps not surprising, then, that returns to scale estimates are, for the most part, similar across these cases. Instrument sets 2 and 3, on the other hand, provide only a weak fit in the first-stage regressions. The key difference between these two sets and the others is the first lag of the oil price, and the monetary shocks. Closer examination of the first-stage regressions indicates that these instruments provide most of the marginal explanatory power. The partial $\bar{R}^2$ measures are very informative. Even though the individual $\bar{R}^2$s for both right-hand-side variables are high for instruments sets 1, 4, and 5 (across all cases they range from a low of 0.19 to a high of 0.54), even these sets do not provide significant partial $\bar{R}^2$s from the first stage regressions (they range from a low of 0.03 to a high of 0.18). This suggests that external effects regressions will be subject to a great deal more sampling uncertainty, which is consistent with the findings above.

What explanation is there for this? The instruments are aggregate variables which probably capture the aggregate business cycle components of the industry level variables. Since these aggregate components are a significant part of the variation in the industry variables, the regular R^2 measures are quite large, for at least some of the instruments. Put differently, the fitted $\Delta\hat{x}_{it}$ are probably highly correlated with the aggregate components of the Δx_{it}. By extension, the instruments would also capture the aggregate business cycle component of the aggregate input variable Δx_t. However, to identify the internal returns to scale parameter what is relevant is the squared correlation between that portion of Δx_{it} not explained by Δx_t ($\Delta\tilde{x}_{it}$) and that portion of $\Delta\hat{x}_{it}$ not explained by $\Delta\hat{x}_t$ ($\Delta\bar{x}_{it}$). Like the instruments, Δx_t is likely to capture the business cycle components of Δx_{it} so that $\Delta\tilde{x}_{it}$ will contain mostly idiosyncratic movements in industry input. On the other hand, because the instruments are not industry-specific, both $\Delta\hat{x}_{it}$ and $\Delta\hat{x}_t$ contain mostly aggregate business cycle components, so the part of one not explained by the other is unlikely to contain significant quantities of idiosyncratic information. For this reason, we may expect partial R^2 measures to be low, and identification to be particularly difficult, in external effects regressions.

4.3. Further disaggregation

This section considers whether two-digit data are sufficiently disaggregated given that restrictions across industries at that level are rejected. In particular, this section asks whether restrictions across three-digit industries within two-digit sectors are rejected. To do this an alternative data set is needed. Here, the NBER productivity database documented in Gray (1989) is used. Unfortunately, comparisons between this data set and the baseline data sets are difficult, due to differences in the sample period (1958–91), and major differences in the treatment of specific payments to labor. Most striking is the fact that labor's share is typically much smaller in the NBER productivity database than in the baseline data sets.

For the NBER database the value-added and energy-corrected regressions were rerun for 1959–84 using each two-digit sector as a separate panel. Within each two-digit sector the restriction that all three-digit industries were alike was tested. The results are mixed. For about half of the two-digit industries the restriction is not rejected at the 5% level, for both types of regressions. For those industries where the restriction is rejected, it is at nowhere near the level of significance that was found across two-digit sectors in the previous sections. While the unrestricted three-digit level regressions suggest some heterogeneity, they are also difficult to interpret due to a sharp increase in the amount of sampling uncertainty.[8]

5. Conclusion

This paper has studied the properties of production function regressions prevalent in the literature on market structure, returns to scale, and external effects. The main findings of the paper are:

1. Cross-industry restrictions on the parameters of these regressions are rejected strongly. The restrictions lend considerable precision to parameter estimates and lead to sharp inferences about the nature of manufacturing industries in the U.S. However, the sharpness of these inferences is found to be unwarranted when the restriction is relaxed. Furthermore, in value-added regressions the restriction seems to induce an upward bias in the estimated degree of returns to scale.

2. Correcting measures of input for cyclical changes in the utilization of capital has a significant impact on estimates of returns to scale and external effects. Most of the evidence suggests that, once this correction is made, returns to scale are constant or slightly decreasing, while external effects are negligible, in the typical manufacturing industry.

3. Separate identification of internal returns to scale and external effects parameters is extremely difficult using standard instruments. Some of the instrument sets considered simply do not fit the data very well. Others fit the two right-hand-side variables individually, but appear to identify aggregate business cycle components of both. As a result, the second- and third-stage regressions are ill-conditioned. This is offered as an explanation of the lack of robustness of the point estimates in external effects regressions across different instrument sets.

The conclusions of this paper for future research are twofold. First, greater attention to possible aggregation bias is needed. This has also been advocated in recent work by Basu and Fernald (1995b) and Burnside, Eichenbaum, and

[8] Full details of the results are available from the author upon request.

Rebelo (1995b). Second, alternative instrument sets should be sought which contain industry-specific components. While this is difficult, recent work by Shea (1993) and Bartelsman, Caballero, and Lyons (1994) suggests that it is possible.

References

Abbott, T.A. III, Z. Griliches, and J.A. Hausman, 1988, Short-run movements in productivity: Market power versus capacity utilization, Manuscript (Harvard University).

Bartelsman, E.J., R.J. Caballero, and R.K. Lyons, 1994, Customer- and supplier-driven externalities, American Economic Review 84, 1075–1084.

Basu, S., 1994, Procyclical productivity: Increasing returns or cyclical utilization?, Quarterly Journal of Economics, forthcoming.

Basu, S. and J.G. Fernald, 1994, Constant returns and small markups in U.S. manufacturing, International finance discussion paper no. 483 (Board of Governors of the Federal Reserve System, Washington, DC).

Basu, S. and J.G. Fernald, 1995a, Are apparent productive spillovers a figment of specification error?, Journal of Monetary Economics 36, 165–188.

Basu, S. and J.G. Fernald, 1995b, Aggregate productivity and the productivity of aggregates, NBER working paper no. 5382.

Baxter, M. and R.G. King, 1992, Productive externalities and the business cycle, Institute for Empirical Macroeconomics discussion paper no. 53 (Federal Reserve Bank of Minneapolis, Minneapolis, MN).

Beaudry, P. and M.B. Devereux, 1993, Monopolistic competition, price setting and the effects of real and monetary shocks, Manuscript (Boston University, Boston, MA).

Beaudry, P. and M.B. Devereux, 1994, Money and the real exchange rate with sticky prices and increasing returns, Manuscript (Boston University, Boston, MA).

Benhabib, J. and R.E.A. Farmer, 1994, Indeterminacy and increasing returns, Journal of Economic Theory 63, 19–41.

Boldrin, M. and A. Rustichini, 1994, Growth and indeterminacy in dynamic models with externalities, Econometrica 62, 323–342.

Burnside, C., M. Eichenbaum, and S. Rebelo, 1995a, Capital utilization and returns to scale, NBER macroeconomics annual 1995, forthcoming.

Burnside, C., M. Eichenbaum, and S. Rebelo, 1995b, Sectoral Solow residuals, NBER working paper no. 5286.

Caballero, R. and R. Lyons, 1989, The role of external economies in U.S. manufacturing, NBER working paper no. 3033.

Caballero, R. and R. Lyons, 1992, External effects in U.S. procyclical productivity, Journal of Monetary Economics 29, 209–226.

Christiano, L.J., M. Eichenbaum, and C. Evans, 1994, The effects of monetary policy shocks: some evidence from the flow of funds, Review of Economics and Statistics, forthcoming.

Farmer, R.E.A. and J. Guo, 1994, The econometrics of indeterminacy: An applied study, Manuscript (University of California, Los Angeles, CA).

Gali, J., 1992, Multiple equilibria in a growth model with monopolistic competition, CEPR discussion paper no. 751.

Gali, J., 1994, Monopolistic competition, business cycles, and the composition of aggregate demand, Journal of Economic Theory 63, 73–96.

Gray, W.B., 1989, Productivity database, Manuscript (Clark University, Worcester, MA).

Hall, A., G. Rudebusch, and D. Wilcox, 1994, Judging instrument relevance in instrumental variables estimation, Manuscript (North Carolina State University, Raleigh, NC).

Hall, R.E., 1988, The relation between price and marginal cost in U.S. industry, Journal of Political Economy 96, 921–947.

Hall, R.E., 1990, Invariance properties of Solow's productivity residual, in: P. Diamond, ed., Growth/productivity/employment (MIT Press, Cambridge, MA) 71–112.

Hornstein, A., 1993, Monopolistic competition, increasing returns to scale, and the importance of productivity shocks, Journal of Monetary Economics 31, 299–316.

Jorgenson, D.W., 1990, Productivity and economic growth, in: E.R. Berndt and J.E. Triplett, eds., Fifty years of economic measurement (University of Chicago Press, Chicago, IL) 19–118.

Jorgenson, D.W., F.M. Gollop, and B.M. Fraumeni, 1987, Productivity and U.S. economic growth (Harvard University Press, Cambridge, MA).

Kiyotaki, N., 1988, Multiple expectational equilibria under monopolistic competition, Quarterly Journal of Economics 103, 695–713.

Romer, P.M., 1986, Increasing returns and long-run growth, Journal of Political Economy 94, 1002–1037.

Rotemberg, J.J. and M. Woodford, 1995, Dynamic general equilibrium model with imperfectly competitive product markets, in: T.F. Cooley, ed., Frontiers of business cycle research (Princeton University Press, Princeton, NJ) 243–293.

Shea, J., 1993, Do supply curves slope up?, Quarterly Journal of Economics 108, 1–32.

Shea, J., 1996, Instrument relevance in multivariate linear models: a simple measure, Review of Economics and Statistics, forthcoming.

Staiger, D. and J. Stock, 1994, Instrumental variables regression with weak instruments, NBER technical working paper no. 151.

Weil, P., 1989, Increasing returns and animal spirits, American Economic Review 79, 889–894.