THE IMPORTANCE OF REALLOCATIONS IN CYCLICAL PRODUCTIVITY AND RETURNS TO SCALE:

EVIDENCE FROM PLANT-LEVEL DATA

by

Yoonsoo Lee *
Federal Reserve Bank of Cleveland

CES 07-05 March, 2007

Abstract

This paper provides new evidence that estimates based on aggregate data will understate the true procyclicality of total factor productivity. I examine plant-level data and show that some industries experience countercyclical reallocations of output shares among firms at different points in the business cycle, so that during recessions, less productive firms produce less of the total output, but during expansions they produce more. These reallocations cause overall productivity to rise during recessions, and do not reflect the actual path of productivity of a representative firm over the course of the business cycle. Such an effect (sometimes called the cleansing effect of recessions) may also bias aggregate estimates of returns to scale and help explain why decreasing returns to scale are found at the industry-level data.

JEL Classification: D24, E32, O47

Keywords: Entry, Exit, Productivity, Returns to Scale

* I am very grateful to Mark Bils for his valuable advice and encouragement. I thank Eric Bartelsman, Yongsung Chang, Satyajit Chatterjee, Tim Dunne, John Fernald, Joao Gomes, Paul Gomme, Lance Lochner, Toshihiko Mukoyama, Peter Rupert, Mark Schweitzer, as well as seminar participants at the Rochester Wegamans Conference, the Federal Reserve Bank of Cleveland, the Federal Reserve System Committee Meeting in Macroeconomics, the Midwest Macro Meeting, Concordia University, the Wharton School, for useful comments. Randy Becker, Renee Fields, Shawn Klimek, and Christopher Kurz were especially generous with their time in helping me with the Census data. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System. The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Carnegie Mellon University Census Research Data Center and Michigan Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed. Support for this research at the Michigan RDC from NSF (awards no. SES-0004322 and ITR-0427889) is gratefully acknowledged.

1 Introduction

Productivity changes and returns to scale are important components of macroeconomic models. In the past, researchers constructing business cycle models have used aggregate data to estimate a representative firm's production function or short-run productivity changes. Previous findings based on aggregate data have been quite mixed. Hall has written that the procyclical behavior of aggregate productivity is one of the most stylized facts about the United States economy (Hall, 1987). However, Basu et al. (2004) argue that, when measured correctly, technology shocks are not correlated either with output or with business cycles. Using a very different approach to identifying technology shocks, Gali (1998) has reached similar conclusions, finding that technology shocks are very negatively correlated with inputs.

Returns-to-scale estimates based on aggregated data provide similarly puzzling results. While increasing returns to scale are advanced as an explanation for the procyclical behavior of productivity and as an important propagation mechanism in business cycle models, recent studies based on industry-level data find decreasing returns to scale, not increasing returns, as the models would predict (Burnside et al., 1995; Basu and Fernald, 1997).

What matters for macroeconomic models is returns to scale across firms, appropriately aggregated. A serious caveat in the interpretation of studies based on the use of aggregate data is their implicit assumption that the composition of producers remains the same over the business cycle. I show that estimates from aggregated data may not serve as reliable estimates of average firm-level parameters if the composition of producers with different levels of productivity changes over the business cycle. I use detailed plant-level data on U.S. manufacturing to show that there are such countercyclical reallocations across producers, and as a result, estimates of productivity changes that are based on aggregate data may not reflect the true cyclicality of productivity that a representative firm experiences during the course of the business cycle. As predicted in theoretical models of the cleansing effect of recessions (Caballero and Hammour, 1994), I find that output shares are reallocated from less-productive to more-productive plants during recessions.

Furthermore, plants entering and exiting during a recession are more productive than those entering and exiting during a boom. As less productive firms are driven out of or kept from entering the market during recessions, overall total factor productivity (TFP) may rise, reducing the procyclicality of aggregate TFP. While the countercyclical effect of this reallocation (or cleansing effect) is relatively weak for the manufacturing sector as a whole, I find that in certain disaggregate industries, such as durables, the effect is statistically significant.

Having established the importance of composition changes over the business cycle in understanding the cyclical behavior of aggregate TFP, I further show that such changes may help explain the finding of decreasing returns to scale at the industry level in disaggregated data. Reallocations that have a countercyclical effect on productivity tend to bias estimates of returns to scale that are based on aggregate data for the following reason. As inputs are reallocated toward less-productive firms during booms, the marginal response of output to input changes may appear lower in aggregate data than the marginal increase in the output of a typical firm, leading to smaller estimates of returns to scale in aggregate data. In most two- and four-digit SIC industries with significantly decreasing returns-to-scale estimates, I find that estimates of returns to scale *decrease* as the plant-level data are aggregated to the industry level.

The finding of decreasing estimates of returns to scale stands in opposition to the findings of previous research that is based on industry-level data, which document higher returns-to-scale estimates at higher levels of data aggregation (e.g., Caballaero and Lyons, 1992; Basu and Fernald, 1997). In fact, my findings on the countercyclical effects of reallocations between plants sharply contrast with previous findings on the effects of reallocations between industries, such as those of Basu and Fernald (1997, 2002) and Basu et al. (2004). Basu and Fernald (2002) claim that the reallocation between two-digit industries with different marginal products explains much of the cyclicality of aggregated productivity. However, the plant-level evidence in this paper points to some potential problems in such studies, which are based on aggregated data. First, the differences in true marginal products across industries, once corrected for composition bias, may not be as large as they appear in industry-level data. Compared to industry-level estimates, which vary substantially across industries, plant-level estimates are rather closer to constant returns to scale. Furthermore, the previously mentioned studies, which are based on macro-aggregates built from industry-level data, ignore the important reallocations that occur within a detailed industry. The evidence from plant-level data clearly shows that, when corrected for aggregation effects running from the plants to the manufacturing industry, aggregated productivity appears more procyclical.

As pointed out in Barlevy (2002), previous empirical work has not found strong evidence of the cleansing effect of recessions. In agreement with Baily et al. (2001)'s study of the cyclical behavior of labor productivity, I find that the cleansing effect is relatively weak for manufacturing as a whole. However, in certain industries I find the effect is statistically significant and may change the implication of returns-to-scale estimates. This finding is consistent with the view that reallocations are a within-industry phenomenon (Davis et al., 1996; Haltiwanger, 1997) and suggests that the cleansing effect of recessions may be more important within industries. It also implies that it can be problematic to assume productivity change is exogenous at a lower level of data aggregation (two- or four-digit), a common assumption when using industry-level data to estimate a production function.

The countercyclical effect of composition changes is well documented in studies of the cyclicality of real wages (Stockman, 1983; Bils, 1985; Solon et al., 1994; Chang, 2000). In contrast to the labor market, in which changes in the composition of the workforce are mostly explained by changes that occur on the extensive margin, i.e., the entry and exit of workers, changes in the composition of producers are mostly explained by changes that occur on the intensive margin of production, i.e., reallocations between continuing plants. The effect of the entry and exit of plants is relatively small, because entering and exiting plants account for a small share of output in the industry.

Section 2 of this paper describes the data used in this study and the empirical evidence of composition changes over business cycles. Section 3 examines how these changes in the composition of producers may affect returns-to-scale estimates for different levels of aggregation. Conclusions are presented in the last section.

2 Composition Changes and the Cyclicality of Productivity

2.1 Measurement of Productivity and Data Description

The plant-level data used in this study are taken from the Longitudinal Research Database (LRD) maintained by the Center for Economic Studies at the U.S. Bureau of the Census. In this study, I use the Annual Survey of Manufactures (ASM) portion of the LRD for the years 1972 through 1997. Because the entire ASM comprises a representative sample of manufacturing plants (Davis et al., 1996), the survey allows me to assess the contribution of entering and exiting plants to the cyclical behavior of productivity, as well as the impact of output reallocation across plants.

Plant-level productivity is measured using a standard total factor productivity index similar to that used by Baily et al. (1992) and by Foster, Haltiwanger, and Krizan (2001). The TFP index for plant *j* is computed as follows:

$$\ln tfp_{jt} = \ln Y_{jt} - \alpha_l \ln L_{jt} - \alpha_m \ln M_{jt} - \alpha_k \ln K_{jt} ,$$

where Y_{jt} is real gross output, L_{jt} is labor input, M_{jt} is real materials, and K_{jt} is real capital stock. The input cost shares for four-digit industries are used as the measure of the corresponding factor elasticities.[1]

There are two problems in measuring cost shares in the ASM. First, the ASM only includes the wage and salary costs of labor. In calculating labor's share of total costs, I follow Bils and Chang (2000), magnifying each four-digit industry's wage and salary payments to reflect other labor payments, such as fringe payments and employer FICA payments.[2] Another problem is that capital costs are not available. Given that previous studies by Rotemberg and Woodford (1995) and Basu and Fernald (1997) find small profits in manufacturing, I assume zero profit at the industry

---

[1] This procedure implicitly assumes that all plants in the industry operate with the same production technology, a common assumption in such studies.

[2] Bils and Chang (2000) use information from the National Income and Product Accounts to calculate the ratio of these other labor payments to wages and salaries at the two-digit industry level.

level so that total revenue will be equal to total cost. Next, I calculate the share of costs for input *J* in the total revenue from the four-digit industry-level data, aggregated from the ASM panels. For these computations, I consider capital expenditure shares to be residuals. Basu et al. (2004) used the same strategy and found a result similar to direct attempts at measuring the share of capital expenditures.

Outputs and inputs are measured in 1987 constant dollars. Real gross output is measured as the total value of sales, deflated by the four-digit industry-specific deflator. All output, materials, and investment deflators are from the NBER manufacturing productivity data set (Bartelsman and Gray, 1996).[3] Labor input is measured as the total hours for production and non-production workers. Because hours for non-production worker are not collected, the value for total hours is estimated following the method used in Baily et al. (1992), in which *total hours* represents the total hours for production workers multiplied by the ratio of the total payroll for all workers to the payroll for production workers. Material input is measured as the cost of materials deflated by the four-digit industry materials deflator. Capital stocks for equipment and structures are constructed using the perpetual inventory method.

2.2 Patterns of Entry and Exit over the Business Cycle

Previous studies on the entry and exit of producers document considerable fluctuations in entry and exit rates over the business cycle (Chatterjee and Cooper, 1993; Campbell, 1998). In Figure 1, I report annual entry and exit rates of plants in the manufacturing sector measured as the share of entering or exiting plants to all plants in a given year. In this study, entering plants are either new plants, which appeared in the LRD for the first time, or plants that restarted production after a certain period of inactivity. Similarly, exiting plants include those that stopped producing during

[3] See Bartelsman and Doms (2000) for the drawback to using deflated production to measure productivity. Some caution is needed in interpreting the results. Ignoring any quality improvement in output that is not reflected in the deflator may result in a downward bias in productivity growth. If new plants enter a market with new products having higher prices, and the number of new plants increases during a boom, the use of a single industry-level deflator may lead to overestimating the procyclicality of aggregate productivity.

the following period and stayed inactive (i.e., zero employee or zero output) for a certain period of time, as well as those that permanently shut down. As discussed in detail in Davis et al. (1996), samples in the ASM panels are rotated every five years. Only large "certainty" plants are continuously observed across different ASM panels; moreover, it is very difficult to measure entry and exit between the two years in which the panels are rotated. In order to avoid measurement errors in entry and exit caused by the panel rotations, the results reported exclude entries and exits measured between two different ASM panels, namely for the years 1973–74, 1978–79, 1983–84, 1988–89, and 1993–94. Figure 1 presents the interpolated values for these missing years (i.e., the first ASM years in each rotation, 1974, 1979, 1984, 1989, and 1994).

The entry rate rises during economic booms and falls during recessions. The correlation between the annual entry rate and the annual growth rate of real GDP (excluding the first ASM panel years) is 0.242. The procyclical behavior of the entry rate is consistent with the findings of previous studies. The annual exit rate in Figure 1 covaries positively with the entry rate. This counterintuitive, procyclical behavior of the exit rate is the result of the fact that, in this study, the category of exiting plants includes those that stop production temporarily. Most of these plants enter during the boom part of the cycle, operate for a short period of time, and stop operating until they reenter the market. Such temporary exits increase during booms, explaining more than 50 percent of plant exits in certain years. In contrast, most of the exits during recessions consist of plants that shut down permanently. When measured on the basis of the number of *permanent* shutdowns (i.e., excluding temporary shutdowns), the annual exit rate does not show procyclical behavior.[4]

Table 1 summarizes the share of the total number of plants that entering and exiting plants

[4] These permanent shutdowns are often called plant deaths in the literature. In a similar way, new plants that appear in the LRD for the first time are called plant births. Measuring entry and exit rates as the share of plant births and deaths in the total number of plants in a given year, I find a negative correlation between the annual entry rate and the annual exit rate. The contemporaneous correlation between the entry rate and the log change of real GDP increases to 0.289. The correlation between the exit rate and the log change of real GDP is 0.05.

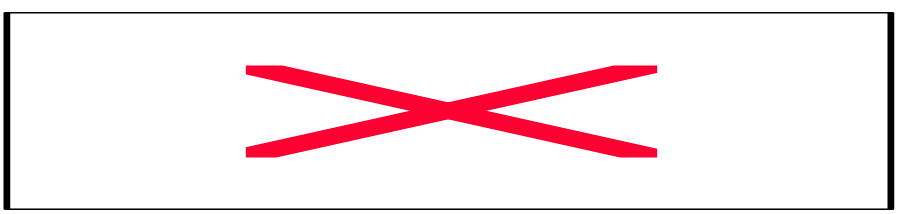
account for over the entire sample period, as well as the share they contribute to total employment and output, and their relative TFP indexes. Entering plants account for about 7 percent of plants in a given year, while 10 percent of plants in a given year stop producing during the following year. Entering and exiting plants tend to be smaller than continuing plants, as reflected in their generally smaller shares of employment and output (2~3 percent). These small contributions contrast with previous studies, such as that of Foster, Haltiwanger, and Krizan (2001), which find that entries and exits make significant contributions to aggregate productivity growth over a longer (five- or 10-year) time horizon. This difference results from the different time horizons over which entries and exits are measured. As the period becomes longer, the number of plants that have entered or exited during the given time period increases. Consequently, the output and labor shares accounted for by entering and exiting plants can be much larger when measured over a longer time horizon.

The last column of Table 1 reports the relative TFP indexes for entering and exiting plants. These indexes consist of the weighted averages of TFPs for entering or exiting plants, divided by the weighted averages of TFPs for continuing plants in the same four-digit industry during the same year. I find that within the same four-digit industry, entrants are relatively more productive than continuing plants, while exiting plants are less productive than continuing plants. While the finding of lower productivity among exiting plants is consistent with the findings of previous studies, such as Foster, Haltiwanger, and Krizan (2001), the finding of higher productivity among entrants differs from previous research. This difference is primarily explained by the following facts: First, the relative productivity of new plants in the ASM panel years is slightly higher than in Census of Manufactures years. Second, new plants that entered in the 1990s have relatively higher productivity than earlier cohorts of entrants. These recent cohorts were not examined in Foster, Haltiwanger, and Krizan (2001).

The same statistics are separately reported for economic boom and recession periods in order to illustrate how the contributions of entering and exiting plants change over time. The second row of Table 1 summarizes the shares and relative TFP for entering and exiting plants for periods when the growth rate of real GDP was greater than 4 percent (i.e., a boom). The third row provides the same statistics for periods when the growth rate of real GDP was less than 1 percent (i.e., a recession). While the number of entering plants increases during a boom, the output share accounted for by entering plants does not increase to a significant degree. This finding is partly explained by the relatively low productivity of entrants during a boom. Overall, plants that enter during a recession or in normal times are more productive than continuing plants. In contrast, plants that enter during a boom are less productive than continuing plants in the same industry. Although the magnitudes are relatively small, these differences in productivity over the business cycle are also found for exiting plants. Plants that exit during a recession are more productive than those that exit during a boom.[5] This difference in the relative productivity of plants that move in and out of production suggests that aggregate productivity is subject to composition effects.

2.3 Decomposition of Aggregate Productivity Changes

Using plant-level data, I examine the extent to which such changes in the composition of producers or shifts in the share of outputs across plants affect the cyclical patterns of aggregate productivity. Following Baily et al. (2001), I have decomposed the time series changes in aggregate productivity into components that reflect a within-plant effect and other effects that reflect the reallocation of shares across plants including the effect of entry and exit:[6]



where $\ln tfp_{jt}$ is TFP index for plant *j* at time *t*, $\ln TFP_t$ is the aggregate TFP index at time *t*, s_{jt} is

[5] The results were similar when I classified booms and recessions in a different way (i.e., boom (recession): years in which the growth rate of real GDP was greater (less) than the average growth rate of real GDP during the sample period (1972-1997).

[6] See Foster, Haltiwanger, and Krizan (2001) for excellent reviews of previous studies using different decomposition methodologies and measurement issues.

the share of output at plant *j* at time *t*, and a bar over a variable indicates the average of the variable over the base and end years ($t - 1$ and t). Because a sample of plants from the ASM is used, the share is further inflated by the ASM sampling weight. The first term in the equation reflects changes in productivity from continuing plants, holding output shares fixed (often interpreted as a "within" effect). This term is measured as the weighted sum of productivity changes with the weights equal to the average output shares across time. The second term reflects changes in output share from continuing plants for fixed levels of productivity (often interpreted as a "between" effect). The last two terms represent the contribution of entering and exiting plants, respectively. These two terms, together constituting the net entry effect, along with the second term, the between effect, represent the effect of reallocations across plants on aggregate productivity changes.

In this decomposition, the change in shares in the second (between-plant) term is weighted by the deviation of plant-level productivity from the average of aggregate productivity, so that an increase in the share of output for a plant contributes positively only if the plant has higher productivity than the average aggregate productivity. In a similar manner, a new plant contributes positively to the aggregate change only if its productivity is above the average, while an exiting plant contributes positively only if its productivity is below the average.

The results of these decompositions are reported in Figure 2. As in Section 2.2, the results reported exclude the first ASM years in each panel in order to avoid measurement errors due to sample rotations in the ASM panels. The values for these missing years in the figure are interpolated. The decomposition components for total factor productivity reveal cyclical patterns similar to those found for labor productivity in Baily et al. (2001). The within-plant component shows clear procyclical behavior. It increased sharply during the booms of 1976 and 1983 and decreased markedly during the recessions of 1980 and 1991. Excluding the first ASM panel years, the correlation between the within-plant component and the contemporaneous real GDP growth is 0.69. Whereas the within-plant term is very procyclical, the between-plant term moves in a countercyclical direction. The between-plant component increased during the recessions of 1975, 1982, and 1991 and decreased during the recovery years of 1976, 1983, and 1992. Although the contribution of plant entry and exit to the annual change in aggregate productivity growth was relatively small (with the exception of the 1990s), the net entry component also moved in a countercyclical direction. These countercyclical reallocation terms suggest that output shares shift from less-productive toward more-productive plants during recessions. As a result of these countercyclical tendencies, aggregate productivity may look less procyclical than the true procyclicality of productivity that is typically observed among individual plants.

The magnitude of such countercyclical effects varies across industries from one level of data aggregation to another. Table 2 reports the results from ordinary least squares (OLS), when those reallocation terms are regressed on real GDP growth. For the manufacturing sector as a whole, the coefficients were negative but not statistically significant. Such a mild countercyclical reallocation effect is consistent with previous empirical studies, which overshadowed the cleansing view of recessions (see Barlevy, 2002). However, the results for disaggregated industries suggest that such a cleansing effect can be statistically significant and potentially more important in certain industries. In durables for example, while within-plant TFP increases about 0.75 percent when real GDP grows 1 percent, such reallocations (between effects and net-entry effects taken together) may decrease the response of aggregate TFP growth by as much as one-half percentage point.

3 Composition Bias in Aggregate Estimates of Returns to Scale

The decomposition results suggest that countercyclical reallocations may cause a downward bias in the returns-to-scale estimates. Because the shares of more-productive plants increase during recessions, the extent to which aggregate output decreases would be smaller than the decreases that would have been observed in a representative plant. Furthermore, because the shares of less-productive plants increase more during booms, the increase in output would look smaller in the aggregate data than the marginal increase in the output of a representative plant.

3.1 Estimating Returns to Scale and the Effect of Composition Bias

In this section, I assess the size of the potential bias caused by composition changes. The key equation to estimate returns to scale follows Basu and Fernald (1997):

$$dy = \gamma[c_L dl + (1 - c_L - c_M)dk + c_M dm] + dz \\ = \gamma dx + dz, \tag{3}$$

where *dy*, *dl*, *dk*, and *dm* are the growth rates of, respectively, output, labor, capital, and materials and c_J is the share of costs for input *J* in the total cost. That is, the growth rate of output, *dy*, equals the returns to scale (γ) multiplied by the cost-share-weighted growth in inputs, *dx*, plus the productivity growth, *dz*.[7] Although inputs are plant-specific, I use industry-level input cost shares, averaged over the beginning and ending years of the period of change.

Many researchers, relying on a representative-firm framework, have used aggregated data and run regressions similar to Equation (3) to estimate returns to scale. This procedure implicitly assumes the existence of an aggregate production function similar to that of (2), while *dz* now reflects aggregate productivity changes. As discussed in the previous section, aggregate productivity changes can be decomposed into components that reflect productivity changes within the plant and other components that reflect reallocations of shares across plants, including entry and exit:

$$\begin{aligned} dy &= \gamma dx + dz \cong \gamma dx + d\ln TFP \\ &= \gamma dx + \sum_{j\in Conti} \bar{s}_j \Delta \ln tfp_{jt} + \sum_{j\in Conti} \Delta s_{jt}(\overline{\ln tfp_j} - \overline{\ln TFP}) \\ &\quad + \sum_{j\in Entry} s_{jt}(\ln tfp_{jt} - \overline{\ln TFP}) - \sum_{j\in Exit} s_{j,t-1}(\ln tfp_{j,t-1} - \overline{\ln TFP}) \\ &= \gamma dx + \sum_{j\in Conti} \bar{s}_j \Delta \ln tfp_{jt} + \text{"}Reallocations\text{"} \\ &= \gamma dx + \sum_{j\in Conti} \bar{s}_j \Delta \ln tfp_{jt} + \delta dx + \varepsilon. \end{aligned} \tag{4}$$

Given that reallocations are negatively correlated with aggregate input changes, a regression run on aggregated data may be subject to a bias caused by composition changes.[8]

[7] Cost minimization implies that returns to scale equals the ratio of average to marginal cost.

[8] In equation (4), *dlnTFP* is calculated assuming constant returns to scale. Therefore, the bias, δ should be interpreted with the understanding that the returns-to-scale coefficient, γ, is implicitly restricted to being equal to 1.

In order to assess the size of the bias that may be present in studies using aggregated data, the "reallocations" term is regressed on aggregated input changes, dx, constructed from the LRD. For the manufacturing sector as a whole, the regression coefficient, δ is –0.043 (see Table 3). As expected, because of the relatively small contribution of plant entries and exits to aggregate productivity growth, the bias caused by net entry (Column 3) is much smaller than the bias caused by between-plant reallocations (Column 2).

Although the effects of composition bias on the returns-to-scale estimates for manufacturing as a whole might not be large, the results for durable and nondurable manufacturing suggest that the effect of composition bias may be larger and possibly significant at more disaggregated levels of data.[9] Previous studies, such as Basu and Fernald (1997), which find decreasing returns to scale based on two-digit industry-level data, also suggest that composition bias may be more important at more disaggregated levels of data. Furthermore, they find that reallocations between industries, in contrast with reallocations between plants, are procyclical in the sense that inputs are reallocated toward industries with higher returns to scale during a boom in the cycle. As a result, the estimate of returns to scale is higher at the higher level of data aggregation. Since the effect of between-industry reallocations may offset the effect of the composition bias caused by within-industry reallocations across plants, it is more relevant to examine the effect of composition changes within an industry. For this study, I focus on analysis at two-digit SIC industry level since previous studies find decreasing returns to scale by using data at this level.

3.2 Estimates of Returns to Scale: Industry- vs. Plant-level Data

One straightforward method of avoiding composition bias is to measure returns to scale at the plant level, giving fixed weights to the exact same plants over time. In this section, I estimate the baseline model in Equation (3), using an ordinary least squares (OLS) regression at two different

[9] I found a significant reallocation effect in durables when the reallocation term is regressed on dx, constructed from the NBER manufacturing database. Although I report results based on the aggregated LRD to keep the data consistent, dx based on aggregated LRD may understate aggregate input changes, due to sample attrition (Appendix of Davis et al., 1996).

levels of aggregation, i.e., the plant level and the two-digit SIC industry level, focusing on the effect of reallocation within two-digit industries.[10]

A potential problem of plant-level analysis is the attenuation bias caused by measurement errors. Previous studies suggest that plant-level returns to scale might be understated by measurement errors present in plant-level hours or capital stocks.[11] Because the specification requires measuring changes in inputs and outputs, first-differencing variables may magnify the attenuation bias, leading to a much smaller returns-to-scale estimate. As a standard response to errors in variables, I introduce an instrument: the cost-share-weighted growth in inputs, dx, measured over $t + 1$ and $t - 2$. Given that a firm's input decisions are highly correlated, plant-level input changes between t and $t - 1$ and those between $t + 1$ and $t - 2$ should be highly correlated as well. If measurement errors are not serially correlated, an IV estimation using the instrument will yield consistent estimates of returns to scale. Although IV estimation may help reduce the attenuation bias caused by measurement errors, it does not take into account the endogeneity of inputs.

Column (1) of Table 4 presents the plant-level OLS results for a pooled sample, which includes plants that have operated for two consecutive years in which they produced nonzero output. All plant-level regression results are obtained from weighted regressions using the ASM sampling weight, so that the sample is representative of U.S. manufacturing as a whole. The IV

[10] As pointed out by a number of researchers since Marschak and Andrews (1944), production function estimates obtained by the OLS method are subject to a simultaneity bias generated by the relationship between productivity and input demands. However, because this study focuses on the effects of composition bias, and its primary concern is the difference between estimates at different levels of aggregation, OLS estimation serves the primary purpose of this study. Assuming that the specification is correct for both the plant- and industry-level regressions, a direct comparison of plant- and industry-level estimates allows an assessment of the size of aggregation bias in estimates obtained from industry-level data.

[11] See Westbrook and Tybout (1993) for evidence of measurement errors in capital. In the appendix, I present evidence of the attenuation bias caused by measurement errors, following the method of Griliches and Hausman (1986) and Goolsbee (2000). In Table A6, returns-to-scale estimates rise (even within the same set of plants), as changes in inputs and outputs are measured over a longer period of time.

estimates appear in bold if the Hausman specification test rejects the null hypothesis, i.e., consistency of the OLS at the 5 percent level of significance.

In order to examine the effects of aggregation, the same equation, (3), is estimated using the industry-level data, created by aggregating all plants in the industry for a given year. The ASM sampling weight is used to make the aggregated data mimic the data used in aggregate studies representing the entire industry. Because the industry-level estimation is less likely to be subject to measurement errors, the OLS estimates are used as the industry-level estimates.[12]

The industry-level estimates show wide variation, ranging from 0.413 for tobacco (SIC 21) to 1.621 for electrical machinery (SIC 36), whereas the plant-level estimates are rather closer to constant returns to scale. Compared to the industry-level estimates (Column 3), the plant-level estimates are smaller in industries with industry-level estimates larger than 1 and larger in industries with industry-level estimates smaller than 1.[13] Although the results are not reported, the analysis of four-digit SIC industries shows a similar pattern.

The bias implicit in aggregated data might help resolve the puzzling finding of the decreasing returns to scale in previous studies that used industry-level data. For example, the statistically significant, decreasing returns-to-scale estimates found in the petroleum (SIC 29) and leather (SIC

---

[12] As it turns out, the Hausman test suggests that the OLS estimates are not statistically different from the industry-level estimates obtained from IV estimations.

[13] Overall, I find that industry-level estimates are larger than plant-level estimates in most two-digit industries in the durables sector. This result may appear inconsistent with the prediction of the productivity decomposition in the durables sector as a whole. However, the effect of composition bias within the two-digit industry can be different from that at a higher level of aggregation. In fact, I find a positive composition bias effect (i.e., $\delta > 0$) in nonelectrical and electrical machinery (SIC 35 and 36), in which industry-level estimates are larger. Furthermore, such countercyclical reallocations across plants with different *productivity levels* may be dwarfed by the procyclical effects of reallocations across plants with different *returns to scale* (as in Basu and Fernald, 1997). If returns to scale vary significantly across plants within an industry and resources are reallocated toward plants with larger returns to scale during booms, creating procyclical bias, industry-level estimates of returns to scale may be larger than plant-level estimates. Unfortunately, with no appropriate measure of plant-specific returns scale, it is very difficult to measure the effect of such reallocation. After correcting for such reallocations effects at the industry level, Basu and Fernald find smaller returns to scale in durables.

31) industries suggest the existence of relatively large positive profits, which seems to contradict empirical evidence of a low profit level, as found in previous studies (Rotemberg and Woodford, 1995; Basu and Fernald, 1997). However, this does not necessarily imply that a typical plant in these industries has decreasing returns to scale, making positive pure profits. Even if the production of an average plant in these industries exhibits constant returns to scale, aggregation may create a bias in the aggregate estimates and lead to a different implication than would the true returns to scale of an average plant. This finding suggests that differences in industry-level estimates of returns to scale across industries may reflect differences in the size of the bias caused by within-industry reallocations, rather than the between-industry differences in returns to scale of an average plant. Whether the industry-level estimate is larger or smaller than true returns to scale will depend on the cyclical behavior of reallocations within the industry.

Implication and Caveat

In order to address the question of what the estimates of the average firm-level returns to scale are, Table 5 presents returns-to-scale estimates in the manufacturing sector when various instruments are used to deal with measurement error. While I find slightly decreasing returns to scale for continuing plants, the returns-to-scale estimate is not very different from 1.[14]

However, appropriate caution should be used when the parameters in this paper are used for the calibration of a macroeconomic model. First, the production function at the plant level may be different from the aggregate production function at the industry level. While the plant-level is likely to be the better level for measuring the production technology of a representative agent, parameters estimated from plant-level data may fail to capture important macroeconomic mechanisms (e.g., reallocations as a propagation mechanism as discussed in Basu and Fernald, 1997

[14] I follow Griliches and Hausman (1986) in choosing instruments. While using variables with longer lags as instruments helps to reduce the attenuation bias, it limits the sample to longer-lived survivors (e.g., 5-year survivors in Column 4), which tend to be bigger and may have higher returns to scale. While the IV estimate in Column 5 is not subject to this problem, the estimate may react to industry-level variation, rather than plant-level variation, leading to an estimate that is closer to an industry-level estimate.

or entry and exit), which are at work during business cycles.

Second, plant-level estimates are subject to biases, the direction of which is not always consistent. Overall, measurement errors in plant-level variables may cause returns to scale to be understated. Imperfect input measures may also overstate input changes over the cycle, causing a downward bias. For example, cyclical changes in inputs can be overstated when various types of overhead capitals (e.g., headquarters, intellectual property, organization capital) are omitted, or when cyclical changes in input quality and price are not properly measured (e.g., lower quality of inputs or higher input prices during booms). In addition, as Klette and Griliches (1996) have pointed out, returns–to-scale estimates may be biased downward if firms sell outputs at different prices (i.e., imperfect competition) while firm-level outputs are deflated based on a common output deflator. On the other hand, the estimates of returns will be biased upward if industry-level variation in inputs is correlated with technology changes. Failure to measure variation in utilization may also lead to an upward bias. While both industry- and plant-level estimates are subject to such biases, the size of a given bias may vary across different levels of aggregation.

Finally, differences in plant-level productivity should be interpreted with caution. While a number of studies using establishment data document a wide dispersion of productivity levels, it is not clear how much of the difference in TFP across plants is explained by differences in the quality of inputs. Although inputs are reallocated between plants with different (average) productivity levels, it does not necessarily imply that marginal products differ between plants. Further investigation of this issue will illuminate the ways in which resources are reallocated over the business cycle.

4 Conclusion

In examining longitudinal plant-level data in U.S. manufacturing, I find that actual productivity may be more procyclical than observed aggregate productivity. As reallocations among producers over the business cycle create a countercyclical component of aggregate productivity, aggregate productivity exhibits less procyclicality than the true procyclicality of productivity observed for a typical producer. Without correcting for such a countercyclical composition bias, measures of technology shocks based on aggregated data may understate the cyclicality of the shocks that a representative agent experiences over the business cycle.

The countercyclical effect of reallocations within an industry helps explain the finding of decreasing returns to scale at the industry level of data. However, the lack of evidence for important increasing returns to scale suggests that other factors such as technology shocks or cyclical utilization may be more important factors in the cyclical behavior of productivity.

REFERENCES

Baily, Martin N., Bartelsman, Eric J., and Haltiwanger, John C. 2001. "Labor Productivity: Structural Change and Cyclical Dynamics," *Review of Economics and Statistics, 83* (3), pp. 420–33.

Baily, Martin N., Hulten, Charles, and Campbell, David. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity: Microeconomics*, pp. 187–249.

Barlevy, Gadi. 2002. "The Sullying Effect of Recessions." *Review of Economic Studies, 69* (1), pp. 65–96

Bartelsman, Eric J., and Doms, Mark. 2000. "Understanding Productivity: Lessons from Longitudinal Micro Datasets." *Journal of Economic Literature, 38* (3), pp. 569–94.

Bartelsman, Eric J. and Gray, Wayne. 1996. "The NBER Manufacturing Productivity Database." National Bureau of Economic Research, Inc., NBER Technical Working Paper: No. 205

Basu, Susanto and Fernald, John G. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy, 105* (2), pp. 249–83.

Basu, Susanto and Fernald, John G. 2002. "Aggregate Productivity and Aggregate Technology." *European Economic Review, 46* (6), pp. 963–91.

Basu, Susanto, Fernald, John G., and Kimball, Miles. 2004 "Are Technology Improvements Contractionary?" National Bureau of Economic Research, Inc., NBER Working Paper: No.

10592.

Bils, Mark. 1985. “Real Wages over the Business Cycles: Evidence from Panel Data.” *Journal of Political Economy, 93* (4), pp. 666–89.

Bils, Mark, and Chan, Yongsung. 2000. “Understanding How Price Responds to Costs and Production.” *Carnegie-Rochester Conference Series on Public Policy,* 52, pp. 33–78.

Burnside, Craig, Eichenbaum, Martin, and Rebelo, Sergio. 1995. “Capital Utilization and Returns to Scale.” in Ben S. Bernanke, and Julio J. Rotemberg eds., *NBER Macroeconomics Annual*, *10,* pp. 67–110.

Caballero, R. J. and Lyons, R. K. 1992. “External Effects in U.S. Procyclical Productivity”, *Journal of Monetary Economics 29* (2), pp. 209–225.

Caballero, Ricardo J. and Hammour, Mohamad L. 1994. “The Cleansing Effect of Recessions.” *American Economic Review, 84* (5), pp. 1350–1368.

Campbell, Jeffrey R. 1998. “Entry, Exit, Technology and Business Cycles.” *Review of Economic Dynamics, 1* (2), pp. 371–408.

Chang, Yongsung. 2000. “Wages, Business Cycles, and Comparative Advantage.” *Journal of Monetary Economics, 46* (1), pp. 143–71.

Chatterjee, Satayajit and Cooper, Russell, 1993. “Entry and Exit, Product Variety and the Business Cycle.” National Bureau of Economic Research, Inc., NBER Working Paper: No. 4562.

Davis, Steven J., Haltiwanger, John C., and Schuh, Scott. 1996. *Job Creation and Destruction*, Cambridge: MIT Press.

Foster, Lucia, Haltiwanger, John C., and Krizan, C.J. 2001.“Aggregate Productivity Growth: Lessons from Microeconomic Evidence,” in Hulten, Charles R., Dean, Edwin R., and Harper, Michael J., eds., *New Developments in Productivity Analysis*, Chicago: University of Chicago Press, pp. 303–63.

Galı´, Jordi. 1999. “Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?” *American Economic Review, 89* (1), pp. 249 –71.

Goolsbee, Austan. 2000. “The Importance of Measurement Error in the Cost of Capital.” *National*

Tax Journal, 53 (2), pp. 215–28.

Griliches, Zvi, and Hausman, Jerry. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics, 31* (1), pp. 93–118.

Hall, Robert E. 1987. "Productivity and the Business Cycle." *Carnegie-Rochester Conference Series on Public Policy, 27*, pp. 421–44.

Haltiwanger, John. 1997. "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence." *Federal Reserve Bank of St. Louis Economic Review*, *79* (3), pp. 55–77.

Klette, Tor J. and Griliches, Zvi. 1996. "The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous." *Journal of Applied. Econometrics, 11* (4), pp. 343–61.

Marschak, Jacob, and Andrews, William H. 1944. "Random Simultaneous Equations and the Theory of Production." *Econometrica*, *12* (3/4), pp. 143–205.

Rotemberg, Julio, J., and Woodford, Michael. 1995. "Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets." in Cooley, Thomas ed. *Frontiers of Business Cycle Research*, Princeton: Princeton University Press.

Solon, Gary, Barsky, Robert, and Parker, Jonathan. 1994. "Measuring the Cyclicality of Real Wages: How Important is Compositional Bias?" *Quarterly Journal of Economics, 109* (1), pp. 587–616.

Stockman, Alan. 1983."Aggregation Bias and the Cyclical Behavior of Real Wages." Manuscript, University of Rochester, 1983.

Westbrook, Daniel, and Tybout, James. 1993. "Estimating Returns to Scale with Large Imperfect Panels: An Application to Chilean Manufacturing Industries." *World Bank Economic Review, 7* (1), pp. 85–112.

Table 1: Shares and Relative TFP of Entrants and Exiting Plants

	Shares in Total Number of Plants (Entry/Exit Rate)		Employment Shares		Output Shares		Relative TFP	
	Entering Plants	Exiting Plants	Entering Plants	Exiting Plants	Entering Plants	Exiting Plants	Entering Plants	Exiting Plants
All sample years	.074	.096	.029	.023	.024	.031	1.094	.977
Boom	.077	.108	.025	.019	.021	.034	.981	.935
Recession	.054	.075	.026	.026	.020	.026	1.072	.974

Note: Booms are years in which log change of real GDP > 4%. Recessions are years in which log change of real GDP <1%.

Table 2: Countercyclical Effect of Reallocations

– regression of GDP growth on reallocation terms

	[1]	[2]	[3]
Dependent variable:	*Reallocations* (Between-plant and Net Entry)	$\sum_{j\in Conti}\Delta s_{jt}(\overline{\ln tfp_j}-\overline{\ln TFP})$ (Between-plant)	$\sum_{j\in Entry} s_{jt}(\ln tfp_{jt}-\overline{\ln TFP}) - \sum_{j\in Exit} s_{j,t-1}(\ln tfp_{j,t-1}-\overline{\ln TFP})$ (Net Entry)
Manufacturing	-0.123	-0.097	-0.027
	(0.164)	(0.145)	(0.059)
Num. of obs.	20	20	20
Nondurables	-0.076	-0.084	0.008
	(0.255)	(0.239)	(0.056)
Num. of obs.	20	20	20
Durables	-0.497*	-0.442*	-0.055
	(0.148)	(0.145)	(0.062)
Num. of obs.	20	20	20

Note: The dependent variable is a subset of the "*Reallocations*" term in Equation (4), stated in the column heading. The independent variable is the log change in real GDP. The sample period is 1972–97, excluding the years 1974, 1979, 1984, 1989, and 1994. * Significant at 1% level.

Table 3: Composition Bias in Returns-to-Scale Estimates

Independent variable:	[1] *Reallocations* (Between-plant and Net Entry)	[2] $\sum_{j \in Conti} \Delta s_{jt} (\overline{\ln tfp_j} - \overline{\ln TFP})$ (Between-plant)	[3] $\sum_{j \in Entry} s_{jt} (\ln tfp_{jt} - \overline{\ln TFP}) - \sum_{j \in Exit} s_{j,t-1} (\ln tfp_{j,t-1} - \overline{\ln TFP})$ (Net Entry)
Manufacturing			
δ	–.043	–.037	–.006
(Std. Err)	(.096)	(.085)	(.035)
Num. of obs.	20	20	20
Nondurables			
δ	–.096	–.032	–.027
(Std. Err)	(.298)	(.174)	(.040)
Num. of obs.	20	20	20
Durables			
δ	–.118	–.107	–.011
(Std. Err)	(.075)	(.071)	(.027)
Num. of obs.	20	20	20

Note: The independent variable is a subset of the "*Reallocations*" term in Equation (4), stated in the column heading. The dependent variable is the cost-share-weighted change in inputs (dx). The sample period is 1972–97, excluding the years 1974, 1979, 1984, 1989, and 1994. The coefficients of the constant terms are not reported.

Table 4: Returns-to-Scale Estimates at Different Levels of Aggregation, Two-digit SIC

A. Nondurables

SIC code	Industry		[1] Plant level (pooled) OLS	[2] Plant level (pooled) IV	[3] Aggregate (all plants) OLS
20	Food	γ	.688	**.741**	.632
		(Std. Err)	(.021)	(.041)	(.154)
		Num. of obs.	118,839	79,337	20
21	Tobacco	γ	.682	**.659**	.413
		(Std. Err)	(.078)	(.082)	(.214)
		Num. of obs.	1,389	1,030	20
22	Textiles	γ	.876	**.868**	1.022
		(Std. Err)	(.030)	(.036)	(.132)
		Num. of obs.	37,572	25,598	20
23	Apparel	γ	.837	**.869**	.833
		(Std. Err)	(.017)	(.029)	(.110)
		Num. of obs.	62,086	33,192	20
26	Paper	γ	.904	.971	1.053
		(Std. Err)	(.028)	(.040)	(.212)
		Num. of obs.	45,584	32,434	20
27	Printing	γ	.729	**.806**	.584
		(Std. Err)	(.020)	(.054)	(.238)
		Num. of obs.	75,493	38,712	20
28	Chemicals	γ	.864	.972	.465
		(Std. Err)	(.034)	(.096)	(.303)
		Num. of obs.	68,289	45,441	20
29	Petroleum	γ	.874	**.983**	.425
		(Std. Err)	(.056)	(.062)	(.169)
		Num. of obs.	15,455	10,200	20
30	Rubber	γ	.855	.996	1.271
		(Std. Err)	(.020)	(.031)	(.098)
		Num. of obs.	55,865	34,389	20
31	Leather	γ	.897	**1.060**	.836
		(Std. Err)	(.059)	(.082)	(.136)
		Num. of obs.	10,066	6,537	20

B. Durables

SIC code	Industry		[1] Plant level (pooled) OLS	[2] Plant level (pooled) IV	[3] Aggregate (all plants) OLS
24	Lumber	γ	.745	**.803**	.885
		(Std. Err)	(.017)	(.048)	(.127)
		Num. of obs.	67,711	35,644	20
25	Furniture	γ	.919	**1.000**	1.296
		(Std. Err)	(.022)	(.049)	(.105)
		Num. of obs.	28,549	16,290	20
32	Stone, Clay, & Glass	γ	.990	**1.074**	1.109
		(Std. Err)	(.021)	(.067)	(.219)
		Num. of obs.	28,549	29,721	20
33	Primary Metals	γ	.792	.914	1.212
		(Std. Err)	(.037)	(.032)	(.099)
		Num. of obs.	40,854	27,782	20
34	Fabricated Metals	γ	.840	**.952**	1.293
		(Std. Err)	(.016)	(.032)	(.202)
		Num. of obs.	112,483	66,116	20
35	Nonelectrical Machinery	γ	.872	**.943**	1.608
		(Std. Err)	(.016)	(.034)	(.181)
		Num. of obs.	118,588	68,887	20
36	Electrical Machinery	γ	.927	**1.029**	1.621
		(Std. Err)	(.018)	(.035)	(.264)
		Num. of obs.	69,047	45,032	20
37	Transportation Equipment	γ	.889	**.997**	1.207
		(Std. Err)	(.023)	(.066)	(.077)
		Num. of obs.	39,173	25,331	20
38	Instruments	γ	.855	**.845**	.524
		(Std. Err)	(.030)	(.050)	(.122)
		Num. of obs.	31,455	19,733	20
39	Miscellaneous Durables	γ	.905	**.978**	.887
		(Std. Err)	(.043)	(.056)	(.175)
		Num. of obs.	26,491	13,944	20

Note: ASM sample weight is used. The sample period is 1972–97, excluding the years 1974, 1979, 1984, 1989, and 1994. Plant-level IV estimates appear in bold if the Hausman test rejects the consistency of the corresponding OLS estimates at the 5% level of significance.

Table 5: Returns-to-Scale estimates for the Manufacturing Sector

	[1] OLS	[2] IV	[3] IV	[4] IV	[5] IV
γ	.828	.910	.937	1.049	1.077
(Std. Err)	(.006)	(.012)	(.062)	(.091)	(.128)
Num. of obs.	1,078,471	655,350	766,787	512,176	1,078,471
Hausman test		263.73	3.00	3.75	3.81
stat. (p-value)		(.000)	(.083)	(.053)	(.051)

Note: For the IV estimates, the following instrumental variables were used for each column.

Column 2: plant-level changes in the cost-share weighted inputs (*dx*) between $t + 1$ and $t – 2$.

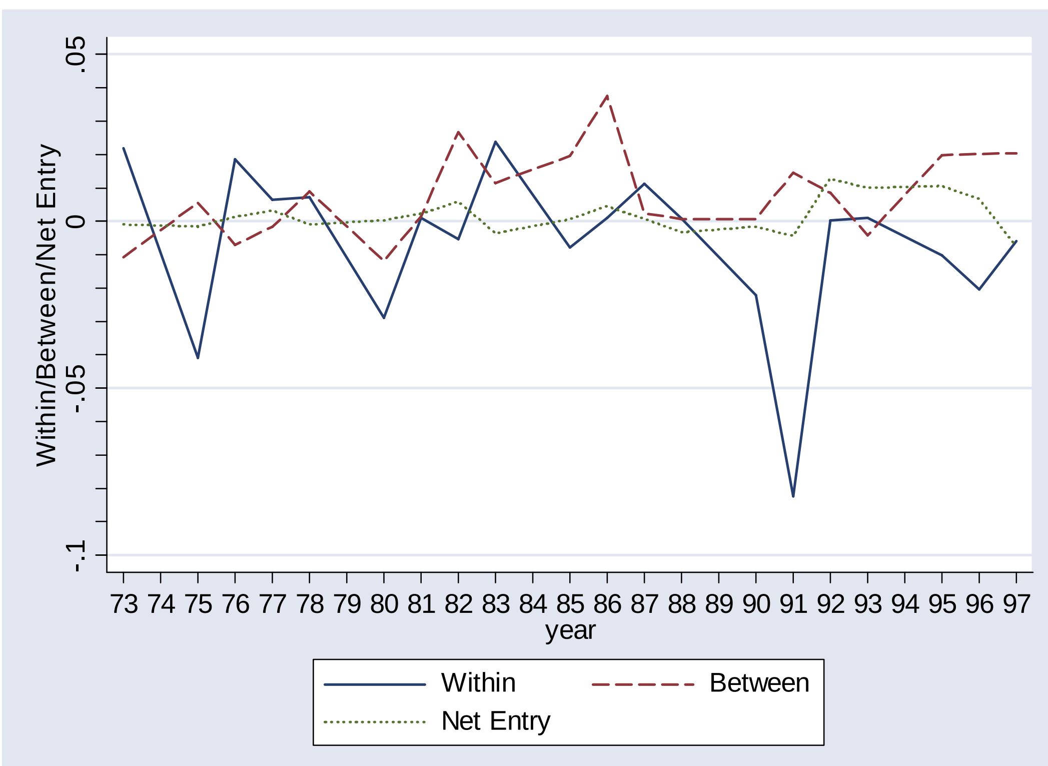
Column 3: 2-year and 3-year lags of cost-share weighted inputs (level).

Column 4: 4-year and 5-year lags of cost-share weighted inputs (level).

Column 5: 4-digit industry-level changes in the cost-share weighted inputs (*dx*) between t and $t – 1$.

Figure 1: Entry and Exit Rates of Manufacturing Plants

Figure 2: Productivity Decompositions



Note: 1974, 1979, 1984, 1989, & 1994 interpolated in the graph.

Appendix

Table A6: Estimates of Returns to Scale over Different Time Horizons

– Plant-level pooled regressions with the same sample of continuing plants

	[1] t & $t-1$	[2] t & $t-2$	[3] t & $t-3$	[4] t & $t-4$
γ	0.889	0.915	0.924	0.930
(Std. Err)	(0.009)	(0.007)	(0.006)	(0.005)
Num. of obs.	632,268	632,268	632,268	632,268

Note: The dependent variable is the log change in real output measured over the stated time period in the column heading. The independent variable is the cost-share-weighted change in inputs over the same stated time period. The regressions are run for the same sample of plants that have operated for at least four consecutive years, to exclude the effects of sample changes due to changes in the time period. ASM sample weight is used.