The MIT Press

Computing Productivity: Firm-Level Evidence

Author(s): Erik Brynjolfsson and Lorin M. Hitt

Source: *The Review of Economics and Statistics*, Nov., 2003, Vol. 85, No. 4 (Nov., 2003), pp. 793–808

Published by: The MIT Press

Stable URL: https://www.jstor.org/stable/3211806

COMPUTING PRODUCTIVITY: FIRM-LEVEL EVIDENCE

Erik Brynjolfsson and Lorin M. Hitt*

Abstract—We explore the effect of computerization on productivity and output growth using data from 527 large U.S. firms over 1987–1994. We find that computerization makes a contribution to measured productivity and output growth in the short term (using 1-year differences) that is consistent with normal returns to computer investments. However, the productivity and output contributions associated with computerization are up to 5 times greater over long periods (using 5- to 7-year differences). The results suggest that the observed contribution of computerization is accompanied by relatively large and time-consuming investments in complementary inputs, such as organizational capital, that may be omitted in conventional calculations of productivity. The large long-run contribution of computers and their associated complements that we uncover may partially explain the subsequent investment surge in computers in the late 1990s.

I. Introduction

In advanced economies, productivity growth depends both on technological innovation and on the organizational changes enabled by technological innovation. The increasing computerization of most businesses is a case in point. Rapid technological innovation in the computer industry has led to a quality-adjusted price decline of 20% or more per year for several decades (Berndt and Griliches, 1990; Gordon, 1999), and these declines are likely to continue for the foreseeable future. Meanwhile, nominal investment in computers has increased even in the face of precipitous price declines, reflecting the myriad new uses that have been found for computers and related technologies. In recent years, companies have implemented thousands of large and small innovations in software applications, work processes, business organization, supply-chain management, and customer relationship management. Research using a growth accounting approach has documented that rapidly rising computer investment in the United States has contributed significantly to output growth, especially in the late 1990s (Oliner and Sichel, 2000; Jorgenson and Stiroh, 2000). Yet the computerization of firms involves far more than merely their investment in computers.

In this paper, we build on previous research on the growth contribution of computerization, exploiting the advantages of measurements at the firm level. Panel data on outputs and inputs (including computers) are available for large numbers of firms, enabling the use of econometric techniques to estimate the contribution of computerization to several measures of multifactor productivity growth. In addition, firm heterogeneity may be used to obtain a more accurate estimate of the true contribution of computerization, especially where these contributions are in the form of intangible benefits (such as quality, variety, or convenience), which are often poorly measured in output statistics (see a formal treatment of this in appendix B and section II). Furthermore, firm-level data enable us to understand the private returns of computerization that ultimately drive decisions by managers to invest in the first place.

Our focus on the firm level also enables us to investigate the process by which computerization contributes to multifactor productivity growth. Computers are best described as a general-purpose technology whose primary contribution is to make new production methods possible when combined with complementary investments such as new work systems, organizational redesign, and business process reengineering (Milgrom & Roberts, 1990; Malone & Rockart, 1991; Bresnahan & Trajtenberg, 1995; Greenwood & Jovanovic, 1998; Bresnahan, Brynjolfsson, & Hitt, 2002). These changes, in turn, yield substantial productivity improvements and perhaps even structural changes in the economy over longer periods of time (David, 1990; Greenspan, 1997; Brynjolfsson and Hitt, 2000).

Indeed, the business and academic literature on computerization emphasizes the importance of large and small complementary changes, including changes in business processes, organization structure, and innovations in customer and supplier relations.[1] These changes can be thought of as complementary investments in organizational capital that may be up to 10 times as large as the direct investments in computers (Brynjolfsson & Yang, 1999; Brynjolfsson, Hitt, & Yang, 2002). Because these complementary investments take time, a testable implication of this argument is that the long-run benefits of computerization should exceed the short-run contribution. These additional benefits from computerization arise as firms implement complementary changes in the rest of the business. Therefore, the resulting effects of computerization on output may be greater than the factor share of computer capital. We can exploit our panel data to test for this relationship by varying the time horizon over which we calculate input and output growth.

A number of previous studies have found a positive relationship between IT investment and firm productivity levels (Brynjolfsson & Hitt, 1995, 1996b; Lichtenberg, 1995). These studies used production-function estimates and found that output elasticities for computers significantly

[1] See Brynjolfsson and Hitt (2000) for a review, and Bresnahan et al. (2002) and the studies cited therein for empirical evidence on this point.

exceed their capital costs.[2] However, no previous economet-ric study on computers and productivity at the firm level has examined multifactor productivity growth, most likely on account of data limitations.

In this paper, we apply standard growth accounting and productivity measurement approaches to examine the rela-tionship between growth in computer spending and growth in output and multifactor productivity for 527 large firms over 1987–1994. Our results suggest that over short hori-zons (such as 1 year), estimated contributions of computers are roughly equal to their costs—they contribute to output growth but not productivity growth. However, as the time horizon increases (increasing the difference length used in the growth calculation), the contribution rises substantially above capital costs, suggesting that computerization in the long run contributes to multifactor productivity (MFP) growth as conventionally measured.

The quantitative results are consistent with qualitative arguments that computers complement other long-term productivity-enhancing investments, including innovations in business methods and organization, which are carried out over a period of several years. Without a direct measure of the cost and timing of complementary investments, we cannot determine whether correlations between computers and MFP represent a true correlation with MFP growth (if the complements were appropriately included) or simply an equilibrium return on a system of investments of computers and their complements. Nonetheless, they do suggest that computers are related to a broader set of assets and that the long-run contribution of computerization to growth is po-tentially much larger than would be expected from the quantity of direct investment in computer capital.

We provide further background on our theoretical frame-work in section II and present the basic models and data in section III. Section IV presents the results using a variety of specifications, section V discusses the main explanations for the findings, and in section VI we conclude with a brief summary and some implications.

---

[2] In contrast, previous research at the industry level has been relatively inconclusive. Morrison (1997) found a zero or even negative correlation between computers and productivity; Siegel (1997) found a positive relationship after correcting for measurement error in input and output quantity. Other studies showing mixed results in industry data include Berndt, Morrison, and Rosenblum (1992), Berndt and Morrison (1995), Morrison and Berndt (1990), and Siegel and Griliches (1991). Even studies which simply assume that computers were earning a normal rate of return have come to contrasting conclusions about what this implies for their overall contribution to economic growth. See Lau and Tokutsu (1992), Jorgenson and Stiroh (1995), Bresnahan (1986), Brynjolfsson (1996), and Oliner and Sichel (1994). More recently, Oliner and Sichel (2000) and Jorgenson and Stiroh (2000) concluded that computers were a major contributor to the productivity revival in the late 1990s, and Gordon (2000) emphasized the role of other factors. Brynjolfsson (1993), Bryn-jolfsson and Yang (1996), and Brynjolfsson and Hitt (2000) provide more comprehensive literature reviews.

II. Background: The Growth Contribution of Computerization

A. Changes in the Production Process in Unmeasured Inputs

Computers are primarily an investment good, so their effect on economic welfare depends on how successfully they support the production of other goods and services. Companies have substantially increased both nominal and real investments in computers over time, and this trend accelerated further in the 1990s. Presumably, companies perceive that exploiting these new technologies will result in a significant potential increase in profits. In part, this trend reflects the substitution of computers for labor or other types of capital along a given production possibility frontier for computer consumers. Users of ever cheaper computer equipment can thereby achieve greater output for a given cost of inputs. However, after properly allowing for the deflation of computer prices, this type of substitution-driven output growth reflects investment growth, not necessarily multifactor productivity growth by computer *users* (Jorgen-son & Stiroh, 1995; Stiroh, 2002). Nonetheless, the welfare effects ascribed to the decline in computer prices (due to productivity growth by computer *producers*) have amounted to a sizable fraction of recent output growth in the United States (Brynjolfsson, 1996; Jorgenson & Stiroh, 1995, 2000; Oliner & Sichel, 2000).

Computers may affect the multifactor productivity growth of the firms that use them by changing the produc-tion process itself and engendering complementary innova-tions within and among firms—the act of *computerizing* a business process or collection of processes. Rather than merely substituting a cheaper input (e.g., computers) for another input (e.g., labor) in the context of a fixed produc-tion process, companies can combine computers with other innovations to fundamentally change their production pro-cesses. This could lead to an output elasticity that is greater than computers' input share and to the appearance of excess returns on computer capital stock. Viewed another way, the complementary innovations can themselves be thought of as a kind of input, or organizational capital (Brynjolfsson et al., 2002). In this interpretation, the presence of seemingly excess returns to computers, especially in the long run, may suggest the presence of unmeasured complementary factors and provide some indication of their output-growth benefits. Although there is substantial case evidence of a wide variety of these complementary factors, including human capital (Murnane, Levy, & Autor, 1999), internal firm organization (Bresnahan et al., 2002; Davenport & Short, 1990; Or-likowski, 1992), and supply chain management systems (Short and Venkatramen, 1992), few studies have consid-ered the broader economic implications of these factors or measured their presence.

B. Unmeasured Output

In addition to unmeasured inputs, computers have also been associated with unmeasured outputs. A variety of case evidence, as well as a direct survey of managers (Brynjolfsson & Hitt, 1996a), suggests that the provision of intangible outputs such as quality, convenience, variety, or timeliness represent major reasons for investing in computers. These types of benefits are difficult to include in price indices (Boskin et al., 1997), leading to potential understatement of output and productivity growth at the aggregate level. In particular, any purely financial accounting of the return on computing investment will likely understate the true output of firms that invest heavily in computerization to improve intangible aspects of output.

Without detailed corrections of output price indices to allow for changes in the intangible component of performance levels, it is difficult to capture these effects directly. But we can indirectly measure the value of intangible performance improvements by examining the measurable variations in output among competing firms. In particular, firms that invest more heavily in computers than do their competitors should achieve greater levels of intangible benefits. In turn, customers will recognize and value these benefits. Thus, we can hypothesize that firms that invest in computers for competitive advantage will be able to charge a higher price, force competitors to lower their prices, or both. In aggregate industry or economy-wide data, this type of firm-level variation will be averaged out, making it difficult or even impossible to measure. However, at the firm level, this variation will result in variation in measured revenue and output, enabling at least some of this intangible value to be detected econometrically (see a formal treatment of this issue in appendix B). However, even firm-level data may miss important industry-wide improvements of intangibles and underestimate the contribution of computerization to performance. If two or more competitors simultaneously introduce computer-supported intangible benefits, some or all of these benefits will be passed on to their customers and elude detection in revenue or output data.

III. Models and Data

A. Estimation Framework

We apply the standard growth accounting framework that has been used extensively for studying the productivity of inputs such as capital, labor, energy, and research and development (R&D) (Berndt, 1991). We assume that the production process of the firms in our sample can be represented by a production function (F) that relates firm value added (Q) to three inputs: ordinary capital stock (K), computer capital stock (C), and labor (L). In addition, we assume that the production function is affected by time (t) and by the industry (j) in which a firm (i) operates. Thus,

$$Q_{it} = F(K_{it}, L_{it}, C_{it}, i, j, t). \tag{1}$$

Following common practice, we assume that this relationship can be approximated by a Cobb-Douglas production function.[3] For most of our analyses, we implement this function with three inputs—ordinary capital, computer capital, and labor—written in levels or logarithms of levels (lowercase letters for factor inputs denote logarithms; firm and time subscripts on inputs and output are omitted except when needed for clarity):

$$Q = A(i, j, t) K^{\beta_k} L^{\beta_l} C^{\beta_c}, \tag{2a}$$

or

$$q = a(i, j, t) + \beta_k k + \beta_l l + \beta_c c. \tag{2b}$$

We will also sometimes consider a four-input specification that uses gross output as the dependent variable and includes materials as an additional input.[4]

The term a, often referred to as the multifactor productivity level or, more ambitiously, total factor productivity level, captures differences in output across firms and over time that are not accounted for by changes in the input use. It contrasts with labor productivity by also taking account of changes in capital inputs. Because we hypothesize the potential existence of additional unmeasured inputs, such as organizational capital, we will generally use the more precise terms "two-factor productivity" (2FP) and "three-factor productivity" (3FP) in this paper, depending on whether computers, as well as capital and labor, are explicitly included as inputs. This allows us to highlight the inclusion of these inputs, but not necessarily the totality of all inputs, in our main estimating equations.[5]

This type of productivity framework is usually implemented in time series or panel data settings by taking the time differences of variables in logarithms to yield growth rates. This is usually a single-time-period difference, but longer multiperiod differences (n years) can also be used. If input variables are measured without error, and factor adjustment to price and other exogenous changes is instantaneous, then the short- and long-difference estimates should be identical. However, as noted by Bartelsman, Caballero, and Lyons (1994), when adjustment is not instantaneous,

---

[3] The Cobb-Douglas functional form has the advantage that it is the simplest form that enables calculation of the relevant quantities of interest without introducing so many terms that the estimates are imprecise. More general functional forms such as the transcendental logarithmic (translog) have been utilized in research on the levels of computer investment and productivity (see Brynjolfsson & Hitt, 1995) with output elasticity estimates nearly identical to those for the Cobb-Douglas specification.

[4] Previous work has suggested that the separability assumptions underlying the value-added formulation are often violated in practice, arguing for a four-input output-based specification (Basu & Fernald, 1995). However, the value-added (three-input) formulation has the advantage for econometric estimation that it reduces biases due to the potential endogeneity of materials, the factor input most likely to undergo rapid adjustment to output shocks.

[5] Just as one way to increase labor productivity is through deepening of physical capital, one way to increase three-factor productivity is through deepening of organizational capital.

longer differences can be interpreted as long-run effects of factor input changes. Such changes include not only the direct effect of factor inputs, but also the effects of adjustment of complementary factors. The time-consuming nature of many of the organizational changes that are complementary to computers will make long-run productivity estimates an important part of our analysis.

In addition, when the factor inputs are measured with error, estimates based on longer differences will typically be less biased than estimates based on shorter differences (Griliches & Hausman, 1986). Thus, as we compare elasticity estimates at varying difference lengths, we will need to consider this *errors-in-variables* argument, as well as the long-run elasticity interpretation.

For growth accounting exercises (such as Oliner & Sichel, 2000, or Jorgenson & Stiroh, 2000), the values of the elasticity parameters (β_c, β_k, β_l) are typically assumed to be equal to their theoretical values, thus enabling three-factor productivity growth and the contribution of each input to be computed without econometric estimation. Under standard assumptions (cost minimization, competitive output and input markets, and factor quantities in long-run equilibrium), the output elasticity is equal to the ratio of the current dollar cost of the input to the current dollar value of output. In addition, in growth accounting practice it is common to average these quantities over the growth interval. We denote the price of output and labor by p and w respectively. The rental price of capital (the current dollar value of service flows for a unit of constant dollar stock) is denoted by r^k, and the rental price of computers by r^c; they are typically computed by the approach of Christensen and Jorgenson (1969).[6] This yields the following estimate of three-factor productivity growth:

$$\dot{a}_n = a_t - a_{t-n}$$

$$= (q_t - q_{t-n}) - \frac{1}{2}\left(\frac{r_t^k K_t}{p_t Q_t} + \frac{r_{t-n}^k K_{t-n}}{p_{t-n}Q_{t-n}}\right)(k_t - k_{t-n})$$

$$-\frac{1}{2}\left(\frac{w_t L_t}{p_t Q_t} + \frac{w_{t-n}L_{t-n}}{p_{t-n}Q_{t-n}}\right)(l_t - l_{t-n})$$

$$-\frac{1}{2}\left(\frac{r_t^c C_t}{p_t Q_t} + \frac{r_{t-n}^c C_{t-n}}{p_{t-n}Q_{t-n}}\right)(c_t - c_{t-n}) \tag{3}$$

To econometrically estimate the contribution of computerization, we can proceed in a number of ways. First, we can simply compute three-factor productivity using equation (3) and regress this value on the change in computer stock:

$$\dot{a}_n = \hat{\lambda} + \hat{\beta}(c_t - c_{t-n}) + \epsilon. \tag{4}$$

The estimated parameter in this equation ($\hat{\beta}$) is the contribution of computerization to three-factor *productivity* growth—the excess in the computer output elasticity above its theoretical value. The total *output* contribution could then be calculated by adding this excess amount to the theoretical value derived from the input quantities and the Jorgensonian rental price.

Alternatively, we can utilize a variant of this framework to estimate the output elasticity directly. Here we regress two-factor productivity growth (computed without the computer term $\dot{a}_n^{\backslash c}$) on computer growth. Defining

$$\dot{a}_n^{\backslash c} = (q_t - q_{t-n}) - \frac{1}{2}\left(\frac{r_t^k K_t}{p_t Q_t} + \frac{r_{t-n}^k K_{t-n}}{p_{t-n}Q_{t-n}}\right)(k_t - k_{t-n})$$

$$-\frac{1}{2}\left(\frac{w_t L_t}{p_t Q_t} + \frac{w_{t-n}L_{t-n}}{p_{t-n}Q_{t-n}}\right)(l_t - l_{t-n}), \tag{5}$$

we have the estimating equation:

$$\dot{a}_n^{\backslash c} = \hat{\lambda} + \hat{\beta}^c(c_t - c_{t-n}) + \epsilon. \tag{6}$$

This approach was previously used by Adams and Jaffe (1996) for the study of R&D productivity, and it has the advantage that it enables a direct estimate of the output elasticity and thus the contribution of computerization to *output* growth. A potential disadvantage of the approaches embodied in equations (3)–(6) is that they rely on proper measurement of input quantities (capital, labor, and materials) in deriving the estimate of 2FP and 3FP.

To the extent that computers may be associated with unmeasured complements or intangible assets that might legitimately be part of the productive assets of the firm (for example, organizational capital), the estimates of 3FP and 2FP are likely to be higher than they otherwise would be. In particular, such unmeasured complements can make estimated growth and productivity contributions of computers appear to be larger than the values that theory would predict based on the factor share of computers alone.

In addition to these formulations, we can also consider different approaches to the direct estimation of the production-function relationship [equation (2b)] in differences. The most obvious formulation is to simply estimate the elasticities directly, using either first differences ($n = 1$) or long differences ($n > 1$) of all inputs and outputs. However, this formulation tends to have poor empirical performance on firm-level data, yielding implausibly low estimates for capital inputs and excess elasticities for labor and materials.[7] This is because labor quantity tends to react faster to exogenous shocks and prices than do other quasi-fixed factors such as capital (including ordinary capital, R&D, and computers), and therefore the smaller changes in these other nonlabor factors are more easily overwhelmed by

measurement error. Because computers have a much smaller factor share than capital or labor, it is important that we minimize the estimation bias introduced by these factors. In the context of R&D measurement, Griliches and Mairesse (1984) therefore proposed a *semi-reduced-form* formulation to directly address the endogeneity of labor. Using this formulation in our setting yields the following system:

$$q_t - q_{t-n} = \hat{\gamma}_q + \frac{\beta^k}{1 - \beta^l} (k_t - k_{t-n})$$
$$+ \frac{\beta^c}{1 - \beta^l} (c_t - c_{t-n}) + \epsilon_q,$$

$$l_t - l_{t-n} = \hat{\gamma}_l + \frac{\beta^k}{1 - \beta^l} (k_t - k_{t-n})$$
$$+ \frac{\beta^c}{1 - \beta^l} (c_t - c_{t-n}) + \epsilon_l.$$

(7)

The first equation is simply a direct estimate of the production function in differences of logarithms, omitting the labor input term; the second is a parallel equation for labor. The coefficient estimates (which can be constrained to be equal across equations) are the elasticities of capital and computers relative to the labor elasticity. The actual capital and labor elasticities can be recovered using an estimate of the labor elasticity derived from its factor share.

B. Data Sources and Construction

The data set for this study was created by combining two main data sources: a database of capital stock of computers provided by Computer Intelligence InfoCorp (CII), and public financial information obtained from Compustat II (Compustat). We also employed rental prices for the capital factors from the BLS, and other price deflators from various government and private sources. In some corroborating analyses, we also used a data set of computer hardware and related expenses obtained through surveys conducted by International Data Group (IDG). Appendix A provides additional details on the data sources and construction.

Computer Stock Data: CII conducts a series of surveys that tracks specific pieces of computer equipment in use at approximately 25,000 sites at different locations of the 1000 largest firms in the United States. CII interviews information systems managers to obtain detailed information on each site's IT hardware assets. The site sampling frequency ranges from monthly to annually, depending on the size of the site. CII's interview process includes checking on hardware that was reported in previous interviews to make more accurate time series comparisons. Each piece of hardware is market-valued and aggregated to form a measure of the total hardware value in use at the firm. These data obviate the

need to make assumptions about retirement rates or depreciation, which are typically required when constructing capital series.[8] The CII data provide a relatively narrow definition of computers that omits software, information system staff, and telecommunications equipment. In addition, the CII data represents the wealth stock (market value of the assets) rather than the productive stock (the value of assets based on output capability) of the surveyed firms. Thus, we multiply these wealth stock asset values by the annual aggregate ratio of the productive stock to the wealth stock of computer assets reported by the BLS. This ratio is approximately 1.2 and holds fairly constant across our sample period. Annual computer stock data are available for the Fortune 1000 for the period 1987 to 1994.

We consulted Standard & Poor's Compustat II database to obtain information on sales, labor expense, capital stock, industry classification, employment, and other expenses for all the firms in the CII database. These data were supplemented with price deflators from a variety of sources to construct measures of the sample firms' inputs and outputs using procedures consistent with earlier work (Hall, 1990; Brynjolfsson & Hitt, 1995; Bresnahan et al., 2002). Output, value added, and materials were deflated using the National Income and Product Accounts (NIPA) output deflators at the two-digit industry level in each year.[9] Labor cost was either taken directly from Compustat where reported, or estimated by multiplying employment by a sector-level estimate of average labor expense. Results are similar in magnitude (but often less precise due to the sample size reduction) when we alternatively use employment or restrict the sample to those firms with reported labor expense. Our rental prices for computers and ordinary capital were based on BLS calculations. The computer rental price represents an aggregate for the entire economy for each year, whereas the rental price for ordinary capital is calculated for each industry (at the NIPA two-digit level) in each year. All factor inputs are measured in constant 1990 dollars. The average rental price is 10.3% for ordinary capital and 44% for computers. The large rental price for computer capital reflects the need to compensate for very large negative capital gains due to the deflation of real computer prices each year.

Sample: Using data from the CII database and Compustat, we constructed a nearly balanced panel of 527 firms

---

[8] This method may introduce some error in the measurement of computer inputs because different types of computers are aggregated by stock rather than flow values (weighted by rental price). The direction of such a bias is unclear because it depends on assumptions about depreciation rates of various types of computers at each site.

[9] To the extent that firms that use computers heavily also consume higher-quality materials, this could introduce a downward bias in the materials estimate, because the output deflator may understate quality change in materials. However, this may be offset partially by a bias in the output deflator in the same direction. The effect of this bias is unknown and cannot be directly estimated, but the fact that output-based and value-added-based specifications (reported later) yield similar results suggests that this bias may not be large in practice.

TABLE 1.—REGRESSION ESTIMATES OF MULTIFACTOR PRODUCTIVITY GROWTH ON COMPUTER GROWTH USING VARYING DIFFERENCE LENGTHS AND DIFFERENT CONTROL VARIABLES

Difference Length (years)	No Controls (1)	Year (2)	Industry (3)	Year & Industry (4)	Sample Size
1	0.0198 (0.0088)	0.0141 (0.0095)	0.0166 (0.0082)	0.0107 (0.0089)	3570
2	0.0206 (0.0088)	0.0144 (0.0095)	0.0179 (0.0082)	0.0116 (0.0089)	3043
3	0.0236 (0.0102)	0.0177 (0.0106)	0.0199 (0.0095)	0.0139 (0.0099)	2516
4	0.0236 (0.0093)	0.0158 (0.0103)	0.0237 (0.0087)	0.0162 (0.0097)	1989
5	0.0387 (0.0110)	0.0398 (0.0116)	0.0347 (0.0106)	0.0360 (0.0111)	1462
6	0.0430 (0.0142)	0.0434 (0.0143)	0.0355 (0.0137)	0.0359 (0.0137)	935
7	0.0535 (0.0184)	0.0535 (0.0184)	0.0388 (0.0176)	0.0388 (0.0176)	451

Estimates of the computer coefficient from equation (4) are shown for a range of difference lengths (rows) using different controls (columns); each cell represents a separate regression. Industry controls are used that divide the economy into 10 industries—see footnote 12. Robust standard errors are shown in parentheses.

in the Fortune 1000 over an 8-year period, omitting firms from our raw data which had incomplete data, especially those which had less than 6 of the 8 years present in the sample, and those which had missing data other than at the beginning or end of the measurement period. This left us with a sample of 4097 firm-year observations. We also have corroborating estimates of firm's computer stocks for 1,324 of these observations that were gathered by IDG. IDG gathered data from a single officer in each firm and used a somewhat different definition of computer capital than was used by CII. For the overlapping firms, the computer capital data had a correlation of 73% between CII and IDG data sets.

The firms in the sample are quite large, averaging $1 billion in value added. Within the sample, 57% of the firms are from the manufacturing industry, 41% from service, and 2% from mining, construction, and agriculture. Some service industries—banking, insurance—are largely excluded because many of the firms in these industries do not report ordinary capital stock on Compustat. Because these industries are particularly computer-intensive, the firms in our sample are somewhat less computer-intensive than the economy as a whole. Otherwise, our sample appears to be broadly representative of large firms in the U.S. economy, and firms in the sample account for about 15% of total U.S. economic output over our sample period.

IV. Results

A. Productivity Analyses

In table 1, we report the results of estimating the 3FP contribution of computerization, based on a regression of 3FP growth on computer growth [equation (4)]. We report the results for difference lengths varying from one year to seven years, the maximal difference possible in our data.

Because differences include overlapping data, this introduces a possible correlation between the disturbances for differences with different base years. We therefore perform our estimates weighting the data based on the theoretical form of the within-firm correlation matrix (unique to each difference length), and then use a robust variance estimator to ensure the standard errors are not biased by empirical deviations from this theoretical structure.[10]

Column (1) of table 1 shows that in the base specification, with no time or industry controls, computers are significantly correlated with productivity growth when measured at all difference levels (t-statistics for all estimates are above 2.2). A striking finding is that the estimated coefficients increase monotonically and substantially as we move from a 1-year difference specification to a 7-year difference specification. The 7-year difference estimate is significantly larger than each of the 1- through 4-year difference estimates at $p < 0.05$ or better, and the 6-year difference is significantly above the 1-year and 2-year difference estimates ($p < 0.05$).[11]

[10] The exact form of the within-firm covariance matrix (where each row and column corresponds to a particular year of observation for a single firm) under zero autocorrelation for an observation with a difference length n ending in year t compared to an observation ending in year $t - j$ is given by $cov(\epsilon_t - \epsilon_{t-n}, \epsilon_{t-j} - \epsilon_{t-j-n})$. This yields a matrix with diagonal elements $2\sigma_\epsilon^2$, a jth off-diagonal element of $-\sigma_\epsilon^2$, and zero otherwise, where σ_ϵ^2 is the variance of the disturbance term. Estimates are computed using the STATA xtgee command with this theoretical covariance structure as the weighting input and standard errors computed by the "robust" option, which performs the calculation based on the empirical covariance matrix of disturbances and is thus robust to other forms of correlation or heteroskedasticity.

[11] We also separately investigated the year-by-year coefficients for each regression (results not shown). Although they vary somewhat from year to year, we generally cannot reject the restriction that the elasticities are the same over time for the same difference length (except for one observation in 1-year differences), and we find the general pattern of rising coefficients nearly identical to that shown in table 1.

TABLE 2.—REGRESSION ESTIMATES OF MULTIFACTOR PRODUCTIVITY GROWTH ON COMPUTER GROWTH USING VARYING DIFFERENCE LENGTHS AND ALTERNATIVE SPECIFICATIONS

Specification:	3FP		2FP without IP			
Output metric:	Output		Value Added		Output	
Difference Length (years)	No Controls (1)	Year & Industry (2)	No Controls (3)	Year & Industry (4)	No Controls (5)	Year & Industry (6)
1	0.0039 (0.0038)	0.0018 (0.0040)	0.0289 (0.0087)	0.0198 (0.0087)	0.0076 (0.0038)	0.0055 (0.0040)
2	0.0048 (0.0037)	0.0026 (0.0039)	0.0300 (0.0085)	0.0210 (0.0086)	0.0085 (0.0037)	0.0063 (0.0039)
3	0.0061 (0.0041)	0.0039 (0.0041)	0.0337 (0.0097)	0.0240 (0.0094)	0.0100 (0.0041)	0.0076 (0.0041)
4	0.0058 (0.0038)	0.0038 (0.0040)	0.0339 (0.0089)	0.0266 (0.0093)	0.0096 (0.0038)	0.0076 (0.0040)
5	0.0107 (0.0050)	0.0108 (0.0050)	0.0494 (0.0107)	0.0466 (0.0108)	0.0147 (0.0049)	0.0148 (0.0050)
6	0.0144 (0.0064)	0.0118 (0.0063)	0.0559 (0.0136)	0.0486 (0.0131)	0.0193 (0.0064)	0.0165 (0.0063)
7	0.0182 (0.0081)	0.0143 (0.0082)	0.0668 (0.0179)	0.0518 (0.0169)	0.0234 (0.0081)	0.0193 (0.0081)

Regression estimates of the computer coefficient using a range of difference lengths (rows) for different specifications (columns); each cell represents a separate regression. Columns (1) and (2) represent the regression of computer growth on 3FP growth [analogous to equation (4)] using gross output rather than value added as the output metric and including a materials input term. Columns (3) and (4) represent a regression of computers on 2FP growth where 2FP growth is computed using value added but without including a computer input term [equation (6)]—estimated coefficients are the output elasticities of computers. Columns (5) and (6) represent the equivalent regressions to columns (1) and (2), calculating 2FP. Robust standard errors are shown in parentheses. Sample sizes are as shown in table 1.

We also examine different sets of control variables, one set for year and another for major industry.[12] These control variables remove effects of industry heterogeneity and possibly short-run time productivity shocks common across all firms that might bias the coefficients. At the same time, they also remove the portion of 3FP that is shared by all firms in an industry or across the economy. Thus, the results with these controls are likely to underestimate the true 3FP contribution of computers and their associated complements. In principle, comparing the results with and without controls can provide an indication of how much, if any, of the 3FP growth attributable to computers is common to the economy or industry.

We find that industry and time effects do influence the measured productivity contribution of computerization. Examining the 1-year difference specification (moving across the first row of table 1), we see that time controls reduce the computer excess elasticity (3FP) estimate by 30% and the industry controls by 20%, and combined they reduce it as much as 45%. In the regressions with the controls, we typically cannot reject the null hypothesis of no contribution of computers to 3FP growth in 1-year through 3-year differences, but we consistently find that the estimated elasticity of computers significantly exceeds the computer input share in longer differences. All results continue to show monotonically increasing coefficients as difference length increases.

We also consider a four-input productivity formulation in which we use gross output as the dependent variable of the production function and include materials as a separate input. The results are shown in table 2 for the no-controls regression [column (1)] and the regression with full industry and time controls [column (2)]. Other regressions show comparable behavior to those in Table 1 and are omitted. As expected, given the smaller factor shares of capital and labor in this specification, the precision of the estimates is substantially diminished. However, the magnitudes are comparable to the earlier estimates.[13] With or without controls, short differences are typically not significantly different from 0, but many of the longer-difference results are. Because the value-added specification yields more precise estimates and exhibits no apparent bias relative to the gross output specification, we focus the discussion on value-added specifications in the remainder of the paper.[14]

In the remainder of table 2, we examine estimates of 3FP calculations that omit the computer input term [equation (6)]—the coefficient estimates are thus output elasticities. Applying the Jorgensonian rental formula to the data, the average input share of computers in our sample is 0.84% of value added. Thus, if these results are consistent with table 1, they should be higher in point estimates by 0.0084 (or 0.0034 for the output specifications). As we see from the table, this relationship is approximately true. Although there had been questions about whether computers

[12] Our major industry controls divide the economy into 10 sectors: high-tech manufacturing, process manufacturing, other nondurable manufacturing, other durable manufacturing, mining/construction, trade, transportation, utilities, finance, and other services.

[13] The ratio of value added to output is 40%, so we expect these coefficients to be 40% of the results reported in table 1.

[14] We continue to compute comparable output-based results as a robustness check.

TABLE 3.—REGRESSION ESTIMATES OF THREE-FACTOR PRODUCTIVITY GROWTH
ON COMPUTER GROWTH USING A SEMI-REDUCED-FORM SPECIFICATION,
VARYING DIFFERENCE LENGTHS, AND CONTROLS

Difference Length (years)	No Controls		Year and Industry	
	Computer Coefficient (1)	Capital Coefficient (2)	Computer Coefficient (3)	Capital Coefficient (4)
1	0.0109 (0.0020)	0.1694 (0.0053)	0.0085 (0.0021)	0.1694 (0.0052)
2	0.0236 (0.0025)	0.1914 (0.0056)	0.0197 (0.0026)	0.1915 (0.0056)
3	0.0334 (0.0031)	0.2069 (0.0060)	0.0290 (0.0030)	0.2060 (0.0059)
4	0.0346 (0.0035)	0.2223 (0.0065)	0.0326 (0.0035)	0.2182 (0.0064)
5	0.0395 (0.0043)	0.2329 (0.0073)	0.0401 (0.0042)	0.2277 (0.0072)
6	0.0429 (0.0058)	0.2441 (0.0092)	0.0399 (0.0055)	0.2410 (0.0089)
7	0.0538 (0.0087)	0.2489 (0.0129)	0.0456 (0.0083)	0.2486 (0.0126)

Regression estimates of the computer coefficient using a range of difference lengths (rows) for different specifications (columns). Each row in paired columns (1)–(2) and (3)–(4) represents estimates on the computers and ordinary capital coefficients in a single systems regression. Columns (1) and (2) represent coefficient estimates for computers and ordinary capital in a semi-reduced-form specification [equation (7)] using iterated seemingly unrelated regression (ISUR) constraining the capital and IT coefficients the same across the two-equation system. Columns (3) and (4) represent a second semi-reduced-form system estimate with year and industry controls. Coefficients in columns (1)–(4) are converted to elasticities by multiplying by the sample-average labor input share. ISUR standard errors are shown. Sample sizes are as shown in table 1.

were contributing significantly to output when dates from before our time period were studied (for example, Solow, 1987; Morrison & Berndt, 1990; Loveman, 1994), we can reject the hypothesis that computers do not contribute to output growth in almost all of our specifications. As before, coefficients rise monotonically as difference length is increased in all specifications.

In table 3, we probe the robustness of the results to potential specification errors in capital and labor. System estimates of the semireduced form specification (using iterated seemingly unrelated regression) of the computer and ordinary capital elasticities are reported in column pairs (1)–(2) without controls and (3)–(4) with controls. Because we cannot reject the equality of coefficients across the labor and output equations in the system, we impose this linear restriction for increased efficiency. The results that appear in the table are the elasticities and their standard errors (rather than the ratios of the elasticities to the labor elasticities), calculated using an average labor input share of 0.575.

Consistent with the findings of Griliches and Mairesse (1984) in the R&D context, the semi-reduced-form specifications show considerably greater precision in the estimates, with t-statistics on the order of 10 (compared to 2–3 for the 3FP regressions). However, the results do appear to be slightly different. First, the rise in coefficients is much steeper as we move from 1-year to 7-year differences: there is as much as a fivefold increase in the share of output attributable to computerization. By contrast, on the 3FP

regressions, the corresponding rise was no more than a factor of 3. In addition, the coefficients on the 1-year differences imply that there is an output growth contribution but not a net productivity growth contribution in the short run. Another useful observation from this table is that the rise as we move to longer differences is much more substantial for computer elasticity (+309%) than for the ordinary capital coefficient (+70%), using estimates from the regression with time and industry controls. In addition, the ordinary capital elasticity is relatively unaffected by the presence of time and industry controls, suggesting that there is substantially more cross-industry heterogeneity in the contribution of computers, and that computers may be more strongly correlated with economy-wide changes in output (a correlation attenuated by the use of time controls).

B. Instrumental Variables Estimates

Our earlier results assume that computer investment is determined by exogenous factors and is not correlated with shocks in productivity or output. The time controls remove the effects of shocks common to all firms over time or across industries.[15] However, this approach may be inadequate if the shocks are firm-specific. For example, if firms disproportionately increase investments in computers in years where demand for their products is unexpectedly high, our short-difference elasticity results may be upward biased. Alternatively, if firms change their other expenses in response to demand shocks more than their investments in computers, then our previous panel-data estimators may underestimate the contributions of computerization.

For instruments, we require variables that are correlated with computer investment at the firm level, but not with output shocks. One reason why different firms might have varying levels of computer investments is that, due to historical choices, they have different technological infrastructures, which make incremental investments in computers and their complements more or less difficult. For example, companies with an existing client-server computing architecture may find it faster and less costly to implement modern software systems, such as enterprise resource planning,[16] which typically run in a client-server environment. Alternatively, firms with aging production equipment may find it more difficult to adapt to electronic controls and other computer-enabled production methods. An aging capital base may also represent a firm-specific inability or unwillingness to invest in new technologies. Finally, we might expect, especially in the short run, that capital constraints

[15] Results are also similar when we include controls for the interaction of time and industry (not shown).
[16] Enterprise resource planning systems are integrated software suites that integrate different functional areas of a firm such as production planning, human resource management, and inventory management.

TABLE 4.—INSTRUMENTAL VARIABLES ESTIMATES OF THREE-FACTOR PRODUCTIVITY GROWTH AND OUTPUT GROWTH ON COMPUTER GROWTH USING VARYING DIFFERENCE LENGTHS AND DIFFERENT SPECIFICATIONS

Specification:	Value Added	Semi-Reduced-Form		Output
Controls:		Computer Coeff.,	Capital Coeff.,	
Difference in Length (years)	Year & Industry (1)	Year & Industry (2)	Year & Industry (3)	Year & Industry (4)
1	0.0599 (0.0125)	0.0190 (0.0016)	0.1193 (0.0056)	0.0096 (0.0026)
2	0.0493 (0.0119)	0.0469 (0.0025)	0.1316 (0.0055)	0.0077 (0.0026)
3	0.0668 (0.0117)	0.0846 (0.0036)	0.1557 (0.0059)	0.0112 (0.0028)
4	0.0599 (0.0132)	0.0632 (0.0039)	0.1788 (0.0067)	0.0079 (0.0033)
5	0.0967 (0.0177)	0.0638 (0.0050)	0.1852 (0.0078)	0.0138 (0.0038)
6	0.1151 (0.0220)	0.0583 (0.0078)	0.2032 (0.0107)	0.0181 (0.0048)
7	0.1010 (0.0246)	0.0782 (0.0105)	0.2024 (0.0140)	0.0150 (0.0057)

Instrumental variables (IV) regression estimates of the computer coefficient using a range of difference lengths (rows) for different specifications (columns). Each cell in columns (1) and (4) represent a separate regression; the pair of columns (2)–(3) for each row represents a separate systems regression. Column (1) represents an IV estimate of equation (4). Columns (2) and (3) represent an ISUR systems regression, constraining the computer and ordinary capital coefficients to be the same across equations and normalized by the sample-average labor share [see equation (7)]. Column (4) represents an equivalent regression to column (1) using 3FP calculated with gross output instead of value added and including a materials term. All regressions use the same instrument set (in levels): capital age, ratio of PCs to mainframe terminals, ratio of network nodes to PCs, debt-equity ratio, and stock market beta. All instruments are interacted with time and industry dummy variables. Robust standard errors are shown in parentheses except in columns (2) and (3), where ISUR standard errors are reported.

could be a deterrent to computer investments or investments in computer-related complements.[17]

We therefore hypothesize a principally cross-sectional set of instrumental variables (IV) for computer growth that includes five measures in total. The first and second measures assess the extent of a firms' deployment of a client-server computing architecture (the ratio of personal computers to mainframe terminals and the fraction of PCs connected to a network). The third measure is capital age, which reflects other production technologies. The final two measures concern capital costs and investment constraints (the debt-to-equity ratio, which is a measure of leverage, and beta, which is a measure of the volatility of the firms' stock price that is a key driver of the cost of capital under the capital asset pricing model). These instruments are introduced in levels, and their effects are allowed to vary by sector and time. We also include time dummies and industry

[17] We considered using price data, but decided against it because prices do not vary across firms. We also considered techniques such as those proposed by Arellano and Bond (1991) or Griliches and Hausman (1986), which enable instrumental variables estimation in panel data without external instruments. In general, factor growth rates for a particular firm have little correlation over time (Blundell and Bond, 1999), making it difficult to estimate production functions in differences with internal instruments. In our data, Arellano and Bond's (1991) dynamic panel data estimator did not perform well—point estimates in a first-difference specification were similar to our results (computer coefficients around 0.013), but had very wide confidence intervals, reflecting low first-stage power. The "systems GMM" estimator of Blundell and Bond (1998) performed slightly better and yielded a computer elasticity point estimate of 0.014, but the estimates were still quite imprecise. However, these estimators are not suitable for long-difference estimation, because long differences alter the moment restrictions that can be used in identification.

control variables in the regressions to remove changes in common exogenous factors over time (such as prices) as well as industry heterogeneity. The time dummies also accommodate any possible set of time-series instruments common across all firms.

Instrumental variables estimates were computed by a two-stage procedure to enable us to compute standard errors comparable to those reported in our other productivity estimates. In the first stage, 3FP and computer growth were projected on the instrument set using ordinary least squares. Then, the fitted values from this first-stage regression were used to compute productivity contribution estimates using the same technique to account for within-firm autocorrelation as before (see footnote 10).

Results of this IV approach for various specifications are shown in table 4. The specifications based on the 3FP regression (column 1) show coefficient estimates substantially larger than any of the previous estimates. Both regressions also show the now-familiar rise in coefficients as a function of the difference period, although the rise is not as large (60%–80%) and is no longer monotonic. As one might expect, the estimates of the semireduced form using IV are more comparable to those without IV, both in the magnitude (0.019) of the one-year differences and in the substantial additional rise as the time difference is lengthened. Similar results are found on the output-based specifications [column (4)]. The IV results provide evidence against the alternative hypothesis that endogeneity leads to an upward bias in the estimate of computer productivity (if anything, they suggest the opposite). Similarly, they suggest that the rising

coefficients are not easily explained by an errors-in-variables bias, which would be removed by instrumental variables estimation. Instead, the results are consistent with the accumulation of complementary inputs that enhance the output contributions of computerization over time.

C. Production-Function Estimation and an Alternative Data Set

To examine the possibility that our results are unique to this data set or the modeling approach we employ, we now analyze the data using production functions instead of directly examining productivity, and we compare results from our data with those of an alternative data set from International Data Group (IDG). In addition, we can further examine the effects of measurement error on our estimates by using the IDG estimate of computer capital stock as an instrument. Under the assumption that measurement errors are uncorrelated between the IDG and the CII data sets, using the IDG estimate as an instrument for the CII computer stock should remove bias due to measurement error (although it will do little to reduce the effects of other forms of endogeneity).

Most previous firm-level studies have focused on estimating production functions in which the elasticity of other factors (capital and labor) are estimated from the data but are constrained to be the same across firms. The results from a three-input (computers, capital, labor) production-function estimation are shown in tables 5 and 6, using both our data set and the data set from IDG, which was used in earlier research by Brynjolfsson and Hitt and by Lichtenberg.

Overall, we find consistency both within this study and between this study and previous work. Ordinary least squares (OLS) estimates of the production function in levels with time and industry controls are reported for each data

TABLE 5.—REGRESSION OF VALUE ADDED ON FACTOR INPUT QUANTITY—LEVELS REGRESSION

Specification	CII-OLS	IDG-OLS	CII-IV
Column	(1)	(2)	(3)
Computer capital elasticity	0.0483 (0.0110)	0.0272 (0.0086)	0.0584 (0.0272)
Ordinary capital elasticity	0.1963 (0.0178)	0.1764 (0.0154)	0.1678 (0.0181)
Labor elasticity	0.7189 (0.0281)	0.7791 (0.0216)	0.7556 (0.0283)
Control variables	Year, industry	Year, industry	Year, industry
R^2 (%)	95.0	95.8	95.8
Sample size: Observations	4097	1324	1324
Firms	527	357	357

Levels regression of value added on computers, capital, and labor quantity for the Computer Intelligence InfoCorp (CII) and International Data Group (IDG) data sets. Huber-White robust clustered (by firm) standard errors reported in parentheses. Columns (1) and (2) represent OLS regressions. Column (3) represents the equivalent regression of column (1) instrumenting computer capital with the corresponding estimate from IDG.

TABLE 6.—INSTRUMENTAL VARIABLES REGRESSION OF THREE-FACTOR PRODUCTIVITY GROWTH ON COMPUTER GROWTH USING IDG COMPUTER CAPITAL QUANTITY AS AN INSTRUMENT AND VARYING DIFFERENCE LENGTHS

Difference Length (years)	Estimate	Sample Size
1	0.0093 (0.0192)	779
2	0.0473 (0.0277)	551
3	0.0724 (0.0333)	331
4	0.0938 (0.0228)	183
5	0.0357 (0.0244)	66

IV regression of 3FP growth on computer growth using a range of difference lengths (rows). Identical to regression in table 4, column (1) (specification: value added; control: year and industry), except the difference in IDG computer stock is included in the instrument list. Robust standard errors in parentheses.

set. These estimates were obtained by pooling the data and estimating the coefficients with OLS, with the standard errors corrected for heteroskedasticity and within-firm correlation using the Huber-White method. The CII estimates for the computer elasticity are higher than the corresponding IDG estimate, but they are not significantly different. The difference may be due to better precision in the CII computer stock estimates than in the IDG estimates,[18] which leads to less bias from errors in variables. Estimated coefficients on other factors are comparable. When we run an IV regression, instrumenting the CII computer capital level by the corresponding IDG estimate, we find that the coefficient on computers rises by about 20%. These IV estimates are also remarkably close to the 7-year difference results. This is not surprising, because one can view a levels regression as equivalent to a difference regression where the difference length becomes very large. Altogether, this suggests consistency in our estimate of the long-run measured contribution of computerization.

To further explore the effect of measurement error, we can utilize the IDG estimate (this time in differences) as an instrument for the IV 3FP regressions, such as those reported in table 4. Results of this analysis [comparable to column (1) of table 4 with this additional instrument] are shown in table 6. Due to the substantial reduction in the size of the data set (because IDG is both a smaller and a less complete panel), the confidence intervals on the estimates are quite wide. However, we still see rising coefficients as the difference length increases, at least up until 5-year differences, where only 66 observations remain. This appears to provide further evidence against the alternative hypothesis that our observed pattern of rising coefficients over longer differences is attributable simply to a measurement-error explanation.

[18] Recall that CII uses a more thorough asset-tracking method than IDG's interviewing of a single key employee at the surveyed firm.

V. Discussion and Analysis

A. Potential Explanations for the Results

The principal results from this econometric analysis are: (1) the measured output contributions of computerization in the short run are approximately equal to computer capital costs, (2) the measured long-run contributions of computerization are significantly above computer capital costs (a factor of 5 or more in point estimates), and (3) the estimated contributions steadily increase as we move from short to long differences. These results are robust to a wide range of alternative treatments, including: using productivity growth or output specifications; estimating production functions rather than productivity values; and applying a series of econometric adjustments for the endogeneity of labor and (subject to limitations of our instrument set) the endogeneity or measurement error of computer investment.

One interpretation of these results could be that computers, at least during this period, had excess rates of return (elasticity per unit of capital input). However, in light of the related research on how computers actually affect business organization and processes, a more consistent explanation is that computer investment is complemented by time-consuming organizational changes. We hypothesize that the short-time-difference estimates represent the direct contribution of computer investment—the increase in output associated with the purchase and installation of a computing asset for some narrow, short-term business purpose. We hypothesize that the long-time differences represent the overall value contributed by the combined computers + complements system—the increase in productivity associated with longer-term adaptation of the organization to more fully exploit its computing assets. In this interpretation, the high values of the long-time-difference estimates correctly reflect the total contribution of the computers + complements system and not just the contribution of computers alone.

The presence of the complements complicates any calculations of the return on the original computer investments. In particular, we would likely overestimate the rate of return if we used these estimates of the output contribution and only included measured computer capital stock in the denominator. Such a calculation would ignore the potentially large, if intangible, investments in the complements that drive the productive use of computers. Alternatively, if we are willing to assume that firms are efficient, on average, in their investments in both tangible (computers) and intangible (complements) assets, then we can derive the likely magnitude of intangible investments that complement computer investments.

This implies that measured "excess" returns ascribed to computers may provide an indirect estimate of the input quantity of these complementary factors, if one assumes that computers and the complements actually earn only normal returns. In this interpretation, for every dollar of computer capital stock, there are four or more additional dollars of unmeasured complements that are correlated with the measured computer capital. These hidden complements could then account for the additional output we measure. Moreover, the rising coefficients over time imply that the adjustment in complementary factors is not instantaneous. In the remainder of this section, we discuss the evidence regarding three plausible alternative explanations, as well as ways of distinguishing the proposed explanation of organizational complements from the alternative explanations.

Alternative Explanation 1: Random Measurement Error: If computer inputs were measured with random error, we would expect estimates on computers' contribution to be biased downward (Griliches & Hausman, 1986). This bias should be most pronounced in shorter differences, since the amount of "signal" (for example, the true change in computer investment) is likely to be reduced by differencing more than the "noise," because the latter is less likely to be correlated over time. Thus, the signal-to-noise ratio, which is inversely proportional to the bias, is likely to increase as longer differences are taken.[19] Thus, our rising coefficients are potentially consistent with a random-measurement-error explanation.

However, three observations contradict this measurement-error hypothesis. First, errors-in-variables models would predict that the relationship between elasticity and difference length would have a specific, concave pattern. If random error is uncorrelated over time, then the true elasticity is related to the measured elasticity by $\beta(n) = \beta^{true}[1 - \sigma^2_{error}/(n\sigma^2_{signal} + \sigma^2_{error})]$, where σ^2_{error} is the error variance and σ^2_{signal} is the true variance in the input. In our data, no single assumption on the error variance fits the observed pattern of our coefficients well. Second, some of the treatments (using alternative estimates of computer capital stocks and IV) should reduce or eliminate the effects of measurement error and thus suppress the pattern of rising coefficients if measurement error is the cause of that pattern. But the same pattern of rising coefficients appears in the IV regressions, and instrumenting the CII data with the alternative estimate for IDG to reduce the measurement error also preserves the increasing-coefficients result. Third, and perhaps most important, the errors-in-variables explanation implies that even the long-time-difference estimates *understate* the true elasticity. Yet the observed estimates taken at face value suggest that computer investments generate extraordinary returns, so if random measurement error is creating a downward bias, then the true and higher magnitude of the effect of computer investments is still unexplained. Therefore, even though we believe there may be

[19] In addition, because changes in different inputs for the same firm are nearly uncorrelated in our sample, the same downward bias should be evident in our specifications that have multiple regressors, such as the semi-reduced-form estimates. This is a straightforward calculation from the standard results on the effects of errors in variables with multiple regressors (see, for example, Greene, 1993).

substantial random measurement error in our measurements of computer inputs, this does not appear to be the sole, or even the principal, explanation of our findings of excess returns. In particular, random measurement error cannot explain why the measured long-run elasticity is so large relative to the factor share of computer capital.

Alternative Explanation 2: Miscounted Complements: Our main conclusion is that organizational investments are probably the largest and most important complements to computers. However, there are a variety of other, simpler complements to the technical investments measured in the data for this study. Computer hardware and peripherals (measured in our analysis) are only one input of a set of technical complements including software, communications and networking equipment, computer training, and support costs.

The size of these technical complements can be considerable. For instance, the Bureau of Economic Analysis (BEA) estimates that in 1996, current dollar business investment in software was $95.1Bn while business investment in computer hardware was $70.9Bn, a ratio of 1.2 : 1 (BEA, 2000). Whether or not technical complements such as software can influence our estimates of the computer elasticity and productivity contribution depends on whether and how they are included in other capital or labor (and thus measured as other inputs in the growth accounting framework).

Productivity estimation, in which omitted factors appear as either capital or labor, has been studied in the context of R&D (Griliches, 1988, chapter 15; Schankermann, 1981). Of particular concern in these studies was that the labor input devoted to R&D was double-counted, appearing as both R&D expense and labor expense. A similar framework can be extended to cases where omitted factors are simply misallocated between categories but correlated with the primary factor of interest (see Hitt, 1996, chapter 1, appendix D). However, because these misclassifications have offsetting effects—factor productivity estimates of computers are biased upward because the computer input quantity is understated, but are biased downward because the contribution of these complements is being credited to capital or labor—this form of misclassification may not substantially influence our results. For instance, if one assumed that there was $2 each of misclassified capital and labor for each $1 of computer stock, then it would result in only a 20% upward bias in the elasticity estimate, based on the derivation appearing in Hitt (1996).[20] Thus, although this form of misclassification can explain some of the apparent excess returns, it is too small to be the principal explanation. In

[20] This analysis shows that as long as computers are small relative to capital and labor, the measured rate of return of computers (output contribution per dollar of factor input) will be equal to a weighted average of the rates of return of the various inputs, with weights equal to the amount of misclassification.

addition, this type of misclassification does not explain the rising coefficients over longer differences.

Alternative Explanation 3: Uncounted Complements: The same is not true for factors that are complementary to computers but omitted entirely from the measures of other factor inputs. This can arise in two situations. First, it arises if for some reason firms are historically endowed with these complements and they do not require current investment to maintain them (for example, if a set of modern, computer-friendly business processes were present at the outset of our sample period). Second, it arises if firms are actively investing in building these complements, but the costs are expensed against labor or materials rather than capitalized. In either situation, only a small portion of the overall investment appears in the growth accounting estimate. Over our sample period, it was indeed uncommon for many aspects of computing projects to be capitalized according to Financial Accounting Standards Board (FASB) rules, including internally developed software. There were considerable changes in these rules in the late 1990s to better recognize software as an investment, but many other types of project costs—especially organizational change investments—are rarely allowed to be capitalized [see Brynjolfsson et al. (2002) or Lev and Sougiannis (1996) for a discussion].

The effect of this type of misclassification can be large. For instance, if there is $1.2 of unmeasured software stock per $1 of computer stock (as stated by BLS estimates), this could account for a 120% overstatement of the measured rate of return to computers. Since software is likely to represent a considerable portion of the unmeasured technical complements (that do not appear in current expense), it would suggest that any excess returns beyond a factor of 2 are probably due to other complements. The most natural candidates are organizational complements such as business processes and organization.

This explanation also ties in closely with our finding of rising coefficients over longer time differences. If, over the short run, the current cost (appearing in labor or materials) of either technical or organizational investments is large relative to their accumulated stock, then the offsetting effects of misclassification on the elasticity estimate come into play. Over longer horizons the stock is large relative to current expense, so there is no corresponding downward bias in the elasticity estimates and consequently we observe high measured returns to computers.

B. Firm-Level Estimates and Aggregate Output Growth

Using our elasticity estimates for computers and the annual real growth rate of computer capital of about 25% per year, computers and their associated complements have added approximately 0.25% to 0.5% annually to output growth at the firm level over this period. As the factor share of computers has grown, so has the output contribution of computerization, ceteris paribus. This contribution will also

appear as increases in productivity growth as conventionally measured (i.e., including labor and tangible capital), although without estimates of the cost of the complementary investments we do not know whether our system of computers and complements would show productivity growth in a metric which fully accounted for the complements as additional inputs (such as intangible organizational capital). However, because our productivity calculation reflects only private returns, including rent stealing but not productivity spillovers, we also cannot know whether the aggregate effect on the economy is smaller or larger than that of the private returns.

If computers were more likely than other inputs to be used to capture rents from competitors, then the aggregate returns to the economy would be less than the sum of the private returns we measure. Firms that invest in computers would merely displace those that do not. Worse, the net effect would be to lower aggregate profits because redistributing rents is a zero-sum game that has no effect on aggregate profits, whereas computer expenditures are costly. However, aggregate corporate profits do not appear to be any lower in our sample period, and there is some evidence that they grew.

There is more evidence for an effect in the opposite direction—computer investments generate positive returns both for the firm and, in aggregate, for the economy. Some of the private benefits of computerization spill over to benefit consumers and even competing firms. For example, when firms like Wal-Mart demonstrate new IT-enabled efficiencies in computerized supply chain management, their competitors attempt to imitate any successful innovations (with varying degrees of success). These innovations are generally not subject to any form of intellectual property protection and are widely and deliberately copied, often with the aid of consulting firms, benchmarking services, and business school professors. Another example of positive externalities is the improved visibility IT systems provide across the value chain, which reduces the impact of exogenous shocks—companies are now less prone (but not immune) to excessive inventory buildups. Job mobility also disseminates computer-related benefits as IT professionals move from firm to firm or use industry knowledge to create new entrants. As a result, the gains to the economy might plausibly be much larger than the private gains to the original innovator.

Computer investments also lead to increases in less observed—but publicly shared—forms of productivity. When two or more competing firms simultaneously invest in flexible factory automation systems, most of the productivity benefits are passed on to consumers via competition in the form of greater product variety, faster response, and fewer stockouts. As noted earlier, these types of outputs are not measured well, leading to underestimates of aggregate productivity growth.

VI. Conclusion

This paper presents direct evidence that computerization contributes to productivity and output growth as conventionally measured in a broad cross-section of large firms. Furthermore, the pattern of rising growth contributions over longer time periods suggests that computers are part of a larger system of technological and organizational change that increases firm-level productivity over time. This is consistent with the conception of computers as a general-purpose technology. Computerization is not simply buying computer capital; instead it involves a broader collection of complementary investments and innovations, some of which take years to implement.

Specifically, although computer investment generates useful returns in its first years of service, we find that greater output contributions accrue over time. When we examine the data in 1-year differences, we find that computerization contributes to output an amount roughly equal to the factor share of computers. This implies that computers contribute to output growth but not to productivity growth in the short run. Over longer time horizons (between 3 and 7 years), computerization is associated with an output contribution that is substantially greater than the factor share of computers alone—between 2 and 5 times as much as the short-run impact. This implies a substantial contribution to long-run productivity growth as conventionally measured.

The results are consistent with the hypothesis that the long-term growth contribution of computerization represents the combined contribution of computers and complementary organizational investment. Other explanations for our findings, such as measurement error (either random or systematic), do not explain these results as well. Our instrumental variables regressions also suggest that endogeneity does not appear to lead to upward biases in the estimation of computers' contribution. The magnitude of the long-run output elasticity associated with computerization is too large to be explained solely by omitted technical complements (like software). By contrast, computer-enabled organizational investments, such as developing new business processes and inventing new ways to interact with customers and suppliers, are plausibly of sufficient magnitude to account for the additional output growth.

Although the late 1990s saw a surge in productivity and output as well as a corresponding surge in computer investment, it is important to note that our analysis is based on earlier data from the late 1980s and early 1990s. This earlier time period did not enjoy extraordinary growth in the overall economy. If computers indeed require several years to realize their potential growth contribution, the economic performance in the late 1990s, including the surge in measured productivity, in part reflects the massive computer and organizational investments made in the early 1990s. Furthermore, high private returns associated with computerization and the increase stock of organizational capital that we impute for the early 1990s also provide the foundation for

the decision by firms to increase their nominal investments in computers shortly thereafter.

REFERENCES

Adams, J. D., and A. B. Jaffe, "Bounding the Effects of R&D," *RAND Journal of Economics* 27:4 (1996), 700–721.

Arellano, M., and S. R. Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," this REVIEW 58 (1991), 277–297.

Bartelsman, E. J., R. J. Caballero, and R. K. Lyons, "Customer- and Supplier-Driven Externalities," *American Economic Review* 84:4 (1994), 1075–1084.

Basu, S., and J. G. Fernald, "Are Apparent Productive Spillovers a Figment of Specification Error?" *Journal of Monetary Economics* 36 (1995), 165–188.

Berndt, E., *The Practice of Econometrics: Classic and Contemporary,* Reading, MA: Addison-Wesley (1991).

Berndt, E., and Z. Griliches, "Price Indexes for Microcomputers: An Exploratory Study," NBER working paper no. 3378 (1990).

Berndt, E. R., and C. J. Morrison, "High-Tech Capital Formation and Economic Performance in U.S. Manufacturing Industries: An Exploratory Analysis," *Journal of Econometrics* 65 (1995), 9–43.

Berndt, E. R., C. J. Morrison, and L. S. Rosenblum, "High-Tech Capital, Economic Performance and Labor Composition in U.S. Manufacturing Industries: An Exploratory Analysis," MIT working paper no. 3414EFA (1992).

Blundell, R. W., and S. R. Bond, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87 (1998), 115–143.

——— "GMM Estimators with Persistent Panel Data: An Application to Production Functions," Institute for Fiscal Studies working paper no. IFF W99/4 (1999).

Boskin, M. J., E. R. Dulberger, R. J. Gordon, Z. Griliches, and D. W. Jorgenson, "The CPI Commission: Findings and Recommendations," *American Economic Review* 87:2 (1997), 78–83.

Bresnahan, T. F., "Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services," *The American Economic Review* 76:4 (1986), 742–755.

Bresnahan, T. F., and M. Trajtenberg, "General Purpose Technologies: 'Engines of Growth'?" *Journal of Econometrics* 65 (1995), 83–108.

Bresnahan, Timothy, E. Brynjolfsson, and L. Hitt, "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence," *Quarterly Journal of Economics* 117:1 (2002), 339–376.

Brynjolfsson, E., "The Productivity Paradox of Information Technology," *Communications of the ACM* 35:12 (1993), 66–77.

——— "The Contribution of Computers to Consumer Welfare," *Information Systems Research* 7:3 (1996), 281–300.

Brynjolfsson, E., and L. Hitt, "Information Technology as a Factor of Production: The Role of Differences among Firms," *Economics of Innovation and New Technology* 3:4 (1995), 183–200.

——— "The Customer Counts," *Informationweek* (September 8, 1996a), 38–43.

——— "Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending," *Management Science* 42:4 (1996b), 541–558.

——— "Beyond Computation: Information Technology, Organizational Transformation and Business Performance," *Journal of Economic Perspectives* 14:4 (2000), 23–48.

Brynjolfsson, E., L. Hitt, and S. Yang, "Intangible Assets: How the Interaction of Computers and Organizational Structure Affects Stock Market Valuations," *Brookings Papers on Economic Activity: Macroeconomics* 1 (2002), 137–199.

Brynjolfsson, E., and S. Yang, "Information Technology and Productivity: A Review of the Literature," in M. Zelkowitz (Ed.), *Advances in Computers,* volume 43 (1996).

——— "The Intangible Benefits and Costs of Computer Investments: Evidence from Financial Markets," MIT Sloan School of Management working paper (1999).

Bureau of Economic Analysis, *Gross Output and Related Series by Industry,* computer disk (1996).

——— "Recognition of Business and Government Expenditures for Software as Investment: Methodology and Quantitative Impact 1959–1998," working paper (May 2000).

——— *Fixed Asset Tables,* computer disk (2001).

Bureau of Labor Statistics, *BLS Handbook of Methods,* Washington, DC: BLS Publications, Office of Publications and Special Studies (1997).

——— "Capital and Related Measures from the Two-digit Database," computer files (release date April 10, 2001).

Christensen, L. R., and D. W. Jorgenson, "The Measurement of U.S. Real Capital Input, 1929–1967," *Review of Income and Wealth* 15:4 (1969), 293–320.

Council of Economic Advisors (Ed.), *Economic Report of the President,* Washington, DC: U.S. Government Printing Office (1996).

Davenport, T., and J. Short, "The New Industrial Engineering: Information Technology and Business Process Redesign," *Sloan Management Review* 31:4 (1990), 11–27.

David, P. A., "The Dynamo and the Computer: A Historical Perspective on the Modern Productivity Paradox," *American Economic Review Papers and Proceedings* 1:2 (1990), 355–361.

Gordon, R. J., "Has the 'New Economy' Rendered the Productivity Slowdown Obsolete?" Northwestern University working paper (June 1999).

——— "Does the New Economy Measure Up to the Great Inventions of the Past?" *Journal of Economic Perspectives* 14:3 (2000), 49–74.

Greene, W., *Econometric Analysis,* New York: Macmillan (1993).

Greenspan, A., "Monetary Policy Testimony and Report to the Congress," Committee on Banking and Financial Services, U.S. House of Representatives (July 22, 1997). Available from http://www.bog.frb.fed.us/BOARDDOCS/HH/.

Greenwood, J., and B. Jovanovic, "Accounting for Growth," NBER working paper no. 6647 (1998).

Griliches, Z., *Technology, Education, and Productivity,* New York: Basil Blackwell (1988).

Griliches, Z., and J. Hausman, "Errors in Variables in Panel Data," *Journal of Econometrics* 31 (1986), 93–118.

Griliches, Z., and J. Mairesse, "Productivity and R&D at the Firm Level," (pp. 339–374), in Z. Griliches (Ed.), *R&D, Patents and Productivity* (Chicago: University of Chicago Press, 1984).

Hall, B. H., "The Manufacturing Sector Master File: 1959–1987," Documentation, NBER working paper no. 3366 (1990).

——— "The Stock Market's Valuation of R&D Investment during the 1980's," *American Economic Review* 83:2 (1993), 259–264.

Hitt, L. M., *Economic Analysis of Information Technology and Organization,* Doctoral Dissertation, MIT Sloan School of Management (1996).

Jorgenson, D. W., and K. Stiroh, "Computers and Growth," *Journal of Economics of Innovation and New Technology* 3 (1995), 295–316.

——— "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity* 1 (2000), 125–211.

Lau, L. J., and I. Tokutsu, "The Impact of Computer Technology on the Aggregate Productivity of the United States: An Indirect Approach," Department of Economics, Stanford University working paper (1992).

Lev, B., and T. Sougiannis, "The Capitalization, Amortization, and Value-Relevance of R&D," *Journal of Accounting and Economics* 21 (1996).

Lichtenberg, F. R., "The Output Contributions of Computer Equipment and Personnel: A Firm-Level Analysis," *Economics of Innovation and New Technology* 3 (1995), 201–217.

Loveman, G. W., "An Assessment of the Productivity Impact of Information Technologies" (pp. 88–110), in T. J. Allen and M. S. S. Morton (Eds.), *Information Technology and the Corporation of the 1990s: Research Studies* (Oxford: Oxford University Press, 1994).

Malone, T., and J. Rockart, "Computers, Networks and the Corporation," *Scientific American* 265:3 (1991), 128–136.

Milgrom, P., and J. Roberts, "The Economics of Modern Manufacturing: Technology, Strategy, and Organization," *American Economic Review* 80:3 (1990), 511–528.

Morrison, C. J., "Assessing the Productivity Impact of Information Technology Equipment in US Manufacturing Industries," this REVIEW 79 (1997), 471–481.

Morrison, C. J., and E. R. Berndt, "Assessing the Productivity of Information Technology Equipment in the U.S. Manufacturing Industries," National Bureau of Economic Research working paper no. 3582 (1990).

Murnane, R. J., F. Levy, and D. Autor, "Technical Change, Computers and Skill Demands: Evidence from the Back Office Operations of a Large Bank," MIT mimeograph (1999).

Oliner, S., and D. Sichel, "Computers and Output Growth Revisited: How Big is the Puzzle?" *Brookings Papers on Economic Activity* 2 (1994), 273–334.

———— "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives* 14:4 (2000), 3–22.

Orlikowski, W. J., "Learning from Notes: Organizational Issues in Groupware Implementation" (pp. 362–369) in J. Turner and R. Kraut (Eds.), *Conference on Computer Supported Cooperative Work* (Toronto: Association for Computing Machinery, 1992).

Pelaia, E., personal communication with IBM representative (1993).

Schankermann, M., "The Effects of Double-Counting and Expensing on the Measured Returns to R&D," this REVIEW 63 (1981), 454–458.

Short, J. E., and N. Venkatramen, "Beyond Business Process Redesign: Redefining Baxter's Business Network," *Sloan Management Review* 34:1 (1992), 7–21.

Siegel, D., "The Impact of Investments in Computers on Manufacturing Productivity Growth: A Multiple-Indicators, Multiple-Causes Approach," this REVIEW 79:1 (1997), 68–78.

Siegel, D., and Z. Griliches, "Purchased Services, Outsourcing, Computers and Productivity in Manufacturing," National Bureau of Economic Research working paper no. 3678 (1991).

Solow, R. M., "We'd Better Watch Out," *New York Times Book Review* (July 12, 1987), 36.

Stiroh, K. J., "Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?" *American Economic Review* 92:5 (2002), 1559–1576.

APPENDIX A

Variables and Data Construction

The variables used for this analysis were constructed as follows:

1. Sales

Total sales as reported on Compustat [item 12, Sales (Net)] deflated by two-digit industry-level deflators from *Gross Output and Related Series by Industry* from the Bureau of Economic Analysis (1996) for 1987–1993, and estimated for 1994 using the 5-year average inflation rate by industry.

2. Ordinary Capital

This figure was computed from the total book value of capital (equipment, structures, and all other capital) following the method in Hall (1990). The gross book value of capital stock [Compustat item 7, Property, Plant and Equipment (Total − Gross)] was deflated by the capital investment deflator (by industry) reported in "Capital and Related Measures from the Two-Digit Database" (BLS, 2001). The deflator was applied at the calculated average age of the capital stock, based on the ratio of total accumulated depreciation [calculated from Compustat item 8, Property, Plant and Equipment (Total − Net)] to current depreciation [Compustat item 14, Depreciation and Amortization]. The calculation of average age differs slightly from the method in Hall (1993), which made a further adjustment for current depreciation. The constant-dollar value of computer capital was subtracted from this result. Thus, the sum of ordinary capital and computer capital equals total capital stock.

3. Capital Rental Prices (Ordinary Capital)

This series was obtained from the BLS multifactor productivity by industry estimates "Capital and Related Measures from the Two-Digit Database" (BLS, 2001). This publication was also the source of the capital

deflators used in our analysis. These measures are based on calculations of a Jorgensonian rental price (see footnote 6) for major asset classes in each industry and then aggregating to obtain an overall capital rental price for each NIPA two-digit industry, which is then mapped to the two-digit SIC industries in our data. Details on methods and calculation approaches are found in the *BLS Handbook of Methods*, chapter 11 (BLS, 1997).

4. Computer Capital (CII Data-Set Definition)

Total market value of all equipment tracked by CII for the firm at all sites. Market valuation is performed by a proprietary algorithm developed by CII that takes into account current true rental prices and machine configurations in determining an estimate. This value is deflated by the BEA price series for computer capital (BEA, 2001).

5. Computer Capital (IDG Data-Set Definition)

Composed of mainframe and PC components. The mainframe component is based on the IDG survey response to the following question (note: the IDG survey questions quoted below are from the 1992 survey; the questions may vary slightly from year to year):

> What will be the approximate current value of all major processors, based on current resale or market value? Include mainframes, minicomputers and supercomputers, both owned and leased systems. Do NOT include personal computers.

The PC component is based on the response to the following question:

> What will be the approximate number of personal computers and terminals installed within your corporation in [year] (including parents and subsidiaries)? Include laptops, brokerage systems, travel agent systems and retailing systems in all user departments and IS.

The number of PCs and terminals is then multiplied by an estimated value. The estimated value of a PC was determined by the average nominal PC price over 1989–1991 in Berndt and Griliches's (1990) study of hedonic prices for computers. The actual figure is $4,447. The value for terminals is based on the 1989 average (over models) list price for an IBM 3151 terminal of $608 (Pelaia, 1993). These two numbers were weighted by 58% for PCs and 42% for terminals, which was the average ratio reported in a separate IDG survey conducted in 1993. The total average value for a "PC or terminal" was computed to be $2,835 (nominal). This nominal value was assumed each year, and inflated by the same deflator as for mainframes: the BEA price series for computer capital (BEA, 2001).

6. Labor Expense

Labor expense was either taken directly from Compustat (item 42, Labor and Related Expenses) or calculated as a sector-average labor cost per employee multiplied by total employees (Compustat item 29, Employees), and deflated by the price index for total compensation (Council of Economic Advisors, 1996).

The average sector labor cost is computed using annual sector-level wage data (salary plus benefits) from the BLS from 1987 to 1994. We assume a 2040-hour work year to arrive at an annual salary. For comparability, if the labor figure on Compustat is reported as being without benefits (labor expense footnote), we multiply actual labor costs by the ratio of total compensation to salary.

7. Employees

The number of employees was taken directly from Compustat (item 29, Employees). No adjustments were made to this figure.

8. Materials

Materials were calculated by subtracting undeflated labor expenses (calculated above) from total expense and deflating by the two-digit industry deflator for output. Total expense was computed as the difference between Operating Income Before Depreciation (Compustat item 13), and Sales (Net) (Compustat item 12).

9. Value Added

Computed from deflated sales (as calculated above) less deflated materials.

Appendix B

Reconciling Firm and Industry Productivity Estimates in the Presence of Unobserved Output

In the main text, we argue that firm-level data may be better able to capture intangible benefits that arise from computer use to the extent that it is due to firm-specific investments, whereas these benefits may be missed in industry-level analyses due to aggregation error. This section presents a formal treatment of that argument.

Consider a single input production function in which a firm produces output by using computers—this is an assumption of separability and is made for convenience in this discussion. Without further loss of generality, we assume that this function is linear in some measure of computers (C) and output (O), normalized to mean 0 for the sample, plus a conventional error term (i.i.d., mean 0): $O = \gamma C + v$. Assume we have observations on multiple firms (N, indexed by $n = 1, \ldots, N$), in M industries (indexed by $m = 1, \ldots, M$).

Let output and computer inputs for each firm comprise a component common across a particular industry (O_m, C_m) and a firm-specific component (ϵ^o, ϵ^c). These firm-specific components are assumed to be i.i.d. across firms, to have mean 0, and to be uncorrelated with the industry effects, but may have a nonzero correlation within firms. These firm-specific components represent unique IT investments in the firm and the private benefits firms receive from these investments.[21] Thus,

$$O = O_m + \epsilon^o,$$

$$C = C_m + \epsilon^c.$$

Note that we have suppressed the firm and industry subscripts except where necessary for clarity.

We consider two OLS estimators of the production relationship: one in firm-level data (a data set with $M \times N$ observations), and an alternative industry-aggregated data set (a data set with M observations representing the industry mean on each O_m and C_m). The OLS estimator of the productivity term in firm level data is thus

$$\hat{\gamma}_{\text{firm}} = \frac{\text{cov}(O_m, C_m) + \text{cov}(\epsilon^c, \epsilon^c)}{\text{var}(C_m) + \text{var}(\epsilon^c)}.$$

The equivalent industry-level estimate is

[21] One type of private benefit that this formulation captures is errors in firm-specific price deflators—if a firm earns greater revenues for the same level of "physical" output due to unmeasured product quality, it will appear as additional output when revenue is deflated by a common industry deflator and is at least partially captured by ϵ^o.

$$\hat{\gamma}_{\text{industry}} = \frac{\text{cov}(O_m, C_m) + \dfrac{1}{N^2}\text{cov}\left(\sum_n \epsilon^c, \sum_n \epsilon^0\right)}{\text{var}\left(C_m + \dfrac{1}{N^2}\sum_n \epsilon^c\right)}$$

$$= \frac{\text{cov}(O_m, C_m) + \dfrac{1}{N}\text{cov}(\epsilon^c, \epsilon^o)}{\text{var}(C_m) + \dfrac{1}{N}\text{var}(\epsilon^c)}.$$

We are interested in the conditions under which the industry-level estimate is less than the firm-level estimate ($\hat{\gamma}_{\text{industry}} < \hat{\gamma}_{\text{firm}}$). Substituting the equations above and rewriting slightly, we get a condition (assuming that computers have a nonnegative effect on output in these manipulations)

$$\frac{\text{var}(C_m) + \text{var}(\epsilon^c)}{\text{var}(C_m) + (1/N)\,\text{var}(\epsilon^c)} < \frac{\text{cov}(O_m, C_m) + \text{cov}(\epsilon^c, \epsilon^o)}{\text{cov}(O_m, C_m) + (1/N)\,\text{cov}(\epsilon^c, \epsilon^o)}.$$

If we note that $(1/N)\text{var}(\epsilon^c) \geq (1/N)\text{cov}(\epsilon^c, \epsilon^o)$, the inequality is preserved after deleting the right-hand terms in the denominators, although this will tend to understate the differences in elasticity estimates (in the correct direction for our argument).[22] Collecting terms yields

$$1 + \frac{\text{var}(\epsilon^c)}{\text{var}(C_m)} < 1 + \frac{\text{cov}(\epsilon^c, \epsilon^o)}{\text{cov}(O_m, C_m)},$$

or

$$\frac{\text{cov}(O_m, C_m)}{\text{var}(C_m)} < \frac{\text{cov}(\epsilon^c, \epsilon^c)}{\text{var}(\epsilon^c)}.$$

The left-hand side is simply the regression coefficient for the industry-specific components alone ($O_m = \gamma_{\text{ind-only}}C_m + v$), and the right-hand side is an analogous regression on the firm-specific components only ($\epsilon^o = \gamma_{\text{firm-only}}\epsilon^c + v$).

There are two implications of this equation:

1. Whenever the marginal product of the firm-specific component of computer investment exceeds the marginal product of the industry component, industry-level data will understate the benefits of computers.
2. If the data have the industry-specific effects removed (as by differencing or industry dummy variables in the regression), then a positive coefficient on IT is evidence of an incremental firm-specific benefit of computers.

[22] A sufficient condition is that the firm-specific component of computer investment exhibits nonincreasing returns to scale. If N is large, these terms can also be dropped.