NBER WORKING PAPER SERIES

TOOLS OR TOYS? THE IMPACT OF HIGH TECHNOLOGY ON SCHOLARLY PRODUCTIVITY

Daniel S. Hamermesh
Sharon M. Oster

Working Paper 6761
http://www.nber.org/papers/w6761

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1998

We thank the Ford Foundation for support provided through the NBER and Robert Mohr, Sarah Libby and especially Cristina Puente for their extremely careful research assistance. We are also grateful to Steve Bronars, George Carter, John McDowell and participants at several seminars and in the SOLE Labor and Population Economics Internet Seminar for helpful suggestions. Most important, we thank the hundreds of scholars who were kind enough to respond to inquiries about their whereabouts and those of their coauthors. The views expressed here are those of the author and do not reflect those of the National Bureau of Economic Research.

© 1998 by Daniel S. Hamermesh and Sharon M. Oster. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Tools or Toys? The Impact of High Technology
on Scholarly Productivity
Daniel S. Hamermesh and Sharon M. Oster
NBER Working Paper No. 6761
October 1998
JEL No. D24, A14

ABSTRACT

Toys. The impact of computers on productivity has been examined directly on macro data and indirectly (on wages) using microeconomic data. This study examines the direct impact on the productivity of scholarship by considering how high technology might alter patterns of coauthoring of articles in economics and their influence. Using all coauthored articles in three major economics journals from 1970-79 and 1992-96, we find: 1) Sharp growth in the percentage of distant coauthorships (those between authors who were not in the same metropolitan areas in the four years prior to publication), as the theory predicts. Contrary to the theory: 2) Lower productivity (in terms of subsequent citations) of distant than close-coauthored papers; and 3) No decline in their relative disadvantage between the 1970s and 1990s. These findings are reconciled by the argument that high-technology functions as a consumption rather than an investment good. As such, it can be welfare-increasing without increasing productivity.

Daniel S. Hamermesh
Department of Economics
University of Texas
Austin, TX 78712-1173
and NBER
hamermesh@eco.utexas.edu

Sharon M. Oster
School of Management
Yale University
New Haven, CT 06520
sharon.oster@yale.edu

I. Introduction and Motivation

In the last several decades investment in equipment and services that make communication easier has increased rapidly in the United States and other developed economies. Since the mid-1980s a combination of technical change and deregulation has also reduced long-distance telephone rates in the United States by 50 percent (Allen, 1995). Fax usage increased by 20 percent between 1996 and 1997 alone (Pitney-Bowes, 1997), while electronic mail, unknown before the mid-1980's, is ubiquitous today. Popular discussion of a wide range of additional examples of rapidly declining prices and explosive growth of the use of telecommunications is provided by Cairncross (1997).

Attempts to measure the impact on aggregate total-factor and labor productivity of these supposedly productivity-enhancing investments in broad-reaching technical improvements such as computing machinery and communications equipment have not met with great success (e.g., Morrison, 1997, but see Greenan and Mairesse, 1996).[1] An alternative, indirect approach has linked the use of high technologies to wages, which presumably reflect productivity (e.g., Krueger, 1993). The difficulty with this latter approach (e.g., Doms et al, 1997; DiNardo and Pischke, 1997) is the potential correlation of unmeasured productivity differences with the propensity to use high technologies.

An alternative to measuring effects on the broader economy is to measure the impact of generalized technical change on productivity in specific activities observed at the micro level. An earlier literature (Griliches, 1958; Trajtenberg, 1989) has clearly traced the effect of specific innovations in raising productivity in specific sectors of the economy. This study expands on that tradition by trying to identify the effect of the recent broad revolution in communications on one activity, scholarly publishing.

We propose here to study scholarly publishing before and after technology greatly lowered communication costs. We examine in particular whether the increasing ease of communication has altered scholars' choices about their methods of production, and whether those methods yielded changes in the productivity of scholarly activity that are consistent with the increasing access to new communications technologies.[2] In Section II we discuss a model of the production process in scholarly writing in relation to the cost of communications, while in Section III we describe the unique data set that we have assembled to examine the relation between technical change and scholarly productivity. Section IV presents the results of using these data to test the hypotheses that we develop, while Section V offers a consistent explanation for most of the results.

II. A Model of Scholarly Production

The example that we use in this study of the impact of technology is the nature and outcomes of the choices of coworkers by authors of scholarly publications in economics. The importance of team research has been stressed by a number of authors studying the economics of innovation (including Dasgupta, 1988, who also presents a summary of some of the research), so that our specific example has broader implications for the study of technical change. Has the decline in the cost of communication altered scholars' choices in a way consistent with these technologies increasing scholarly productivity? In examining scholarly productivity, we focus on research output. In particular, we view scholars as having three choices in producing a scholarly paper: 1) Work solo, s; 2) Work with close-by coauthor(s), c; or 3) Work with distant coauthor(s), d. In the model in this section, the scholar is assumed to choose a production technology that maximizes his/her scholarly productivity, measured as the quality of the paper produced. In a later section of the paper, we adopt an alternative characterization of a scholar's choice of research strategies.

We assume that the scholar has a wide range of potential research activities to choose among and has perfect knowledge of the productivity P (valued in dollars) of all potential matches s, $c_i \in C$ and $d_i \in D$.[3] Each match generates one solo-equivalent article per period.[4] Thus, in our model, productivity is measured entirely by the quality of the paper rather than the number of papers produced. We assume that the production function is linear in the fraction of the time period devoted to scholarly production. Matching with s or c_i takes no time, so that the entire period can be devoted to developing and improving the scholarly product. Matching with one of the known possible distant coauthors takes some resources τ per match. The scholar will choose the match that yields productivity P*:

$$(1) \qquad P^* = \operatorname{argmax}\Big(P(s),\, P(C),\, P(D)-\tau\Big),$$

where P(C) denotes the vector of productivity of the matches c_i , and $P(D)-\tau$ denotes the vector of differences between the productivity of the matches d_i and the cost of making a distant match.

Given that distant co-authorship generates costs not experienced in the production of other research, we can see that if some match d_i is chosen, it must be highly productive to have overcome those added communication costs. This suggests:

Proposition I.1: Distant coauthorships will be more productive than close coauthorships (or than solo work).

With recent improvements in communications technology the cost of a distant match, τ, has decreased over time. Assuming that there have been no offsetting changes in the inherent productivity of the three types of matches over time (in the distributions of the underlying $P(\cdot)$), this decline in cost should increase the likelihood that a distant match will maximize (1). Thus:

Proposition I.2: The fraction of all coauthorships (and all articles) that are distant will increase as communication cost decreases.

When communication cost is high, it pays to match only with the most productive distant coauthors, since only with them will the productivity of the match overcome the cost of matching. As τ decreases, the probability increases that a distant match that is less productive achieves the maximum in (1). Indeed, in the extreme case, when $\tau=0$, distance is costless, and close and distant coauthorships that are chosen will be equally productive. This suggests:

Proposition I.3: The productivity advantage of distant over close coauthorships that are chosen will decrease over time.

The problem that we analyze in this paper is similar to the issue of the globalization of business that has been addressed by a number of authors, most recently formally by Lazear (1998). The first global activities that businesses undertake are those that, despite the underlying costs of forming a "team," are so highly productive as to be worthwhile. As the costs of forming global teams fall, the additional hurdle that such activities must overcome also falls; as a consequence the number of global activities will rise and the profitability of the marginal global activity will converge to the profitability of domestic joint ventures. In the case of research production, as the cost of distant partnering falls, the incremental hurdle such papers must overcome will also fall, and the value of the marginal distant and closely authored papers will equalize.

III. Data and Trends in Coauthoring Patterns

A substantial literature on coauthorship by economists already exists (and is summarized by Eisenhauer, 1997). A rising trend of coauthorship since the 1940s is very well documented, while the evidence on the relative productivity of coauthorships compared to single-authored papers is sparse and mixed. No one, however, has considered the nature of coauthorships -- close or distant --

that in this study provides the basis for testing for the productivity-increasing effects of declines in communication cost.

To examine the effects of technology on research production, we designate two periods: 1970-79, years that predate most of the technological and price changes described earlier; and 1992-1996, the post-communication revolution period. To obtain a sufficiently large sample from the early period in which co-authorships of any sort were less common, we used an entire decade. For the two periods we obtained data on all coauthored articles (but not comments or communications) in regular issues of the American Economic Review, Journal of Political Economy and Quarterly Journal of Economics. (Sampling more (lower-quality) journals would have generated many articles, but for too large a fraction of them productivity as we measure it in this study would be zero.) This left us with a sample of 813 coauthored articles, of which 145 had more than two coauthors, with a total of 1879 names appearing on the articles. Among these coauthors were 1309 different individuals.

Using the affiliations listed in the article, we first created a measure of the fraction of coauthors who were at the same institution, or within fifty miles of each other, at the time of the article's publication. The fraction is 0 or 1 for two-authored papers, but it can take different values for multi-authored studies, e.g., 0, 0.2, 0.4, 0.6, 0.8 or 1 for a five-authored article. Clearly, this is a poor measure of propinquity in scholarly production, given publication lags and labor mobility: The apparently distant coauthors could have been together during the entire process of generating the study, but one may have moved during the year or more between final polishing and publication. To account for this difficulty, in most of the analysis we use a broader definition of distant coauthorship: The coauthors were located within fifty miles of each other for fewer than 9 months in the four calendar years preceding their study's publication.[5]

Obtaining information on coauthors' propinquity is not easy. Using electronic mail, and, where necessary, surface mail or telephone, we contacted at least one coauthor of the more than 300 articles that appeared to be distant based on the affiliations listed in the article and on the short biographies in the Survey of Members of the American Economic Association, in the Econometric Society Directory of Members, and in the Royal Economic Society Register of Members.[6] With follow-ups we were able to obtain a definitive classification of the distance or closeness of coauthors of every one of the coauthored studies published in the two sample periods.[7]

In this study we measure productivity as the quality of the paper produced. For our measure we rely on citations, a common metric in this area. In particular, for each article in the sample published in year t, we accumulated data from the Social Science Citation Index on citations by other authors to that article in years t+1, ... ,t+4 for the early studies and for as many years as possible for the recent studies. With citations data available through 1997 this means that at least one year's subsequent citations are available for all years 1992-96.

Because solo-authorship is an alternative to coauthoring that avoids the cost of distance while foregoing the potential benefits of collaboration, we also do some testing of the distance effect and its change over time by making comparisons to solo-authored papers. To do so we collected a sample of articles from the same three journals from 1970-79, and the complete set of solo-authored papers from these journals from 1992-96.[8] Information on subsequent citations and all the other variables on which we obtained data on coauthored articles was also acquired on these articles (except, of course, on their distance status).

Table 1 shows the basic information on the sample of coauthored articles from these three prestigious journals in the early and late periods. The first row presents the well-known fact that the incidence of coauthorship increased sharply over this period, with coauthored articles changing from being a significant minority of papers published to being a substantial majority. The second row reflects the new fact that there has been an accretion of authors even within the (expanding) set of coauthored articles. Indeed, not only did the fraction of multiply-coauthored articles increase: Conditional on multiple coauthorship the probability of four or more authors' names being on the article increased too (from 0.07 to 0.13).

The major purpose of Table 1 is to provide tests of Proposition I.2. The evidence seems absolutely clear: Whether one measures distance as at the time of publication or as more stringently defined over the entire four-year period prior to publication, the recent period saw a substantial increase in distant coauthorships. The data reflect the entire population of coauthored articles published in these journals; but if we view them as samples of articles from the population of published scholarly works, we can test the significance of the increases in the distance measure. The t-statistic on the hypothesis that average fraction of distant coauthorships measured at publication remained unchanged at 0.438 between the two periods is 3.07; that on the hypothesis that average fraction measured more carefully over the four-year pre-publication period remained unchanged at 0.056 is 6.46. Very clearly, the data provide strong support for the proposition that easier communication increased contacts among distant scholars that led to the production of articles published in leading scholarly outlets.[9]

IV. Testing for Productivity Effects

The first six columns of Table 2 describe the sample of coauthored articles, while the two right-hand columns describe the solo-authored articles. In the top part of the table we list the means of the productivity measures describing the articles in the sample. These are presented for the early and recent periods separately, and for all coauthored articles and for two-authored articles classified by authors' distance over the four pre-publication years. (In many cases the measure of distance prevents the multiply-coauthored articles from being categorized as entirely distant or entirely close.)

The evidence on the relative productivity of close versus distant partnering is not generally supportive of the model.[10] In particular, Table 2 shows that mean citations to distant-coauthored articles are typically lower than the citations received by close-coauthored articles. This provides a hint that Proposition I.1, that distant coauthorships are more productive than close-coauthorships, is not supported by the data. Indeed, combining the two-authored papers from both time periods, the t-statistics testing the hypothesis that the distant two-authored articles are as productive as the close ones are t = -1.87, t = -0.22, t = -0.89 and t = -0.64, for Years t+1 through t+4 respectively. Nor is there strong evidence for Proposition I.3, that distant coauthorship approximated the productivity of close coauthorship more closely during the recent period: The gap between distant and close authored productivity shows no compelling pattern, with $\{[P_{t+j,d,LATE} - P_{t+j,c,LATE}] - [P_{t+j,d,EARLY} - P_{t+j,c,EARLY}]\}$ yielding the test statistics t = 0.09, t = -1.43, t = -0.49 and t = -0.53.

The major difficulty with looking at sample means of the articles' citations is that the distributions of citations are highly skewed. To obtain a better feel for the relation of articles' citations to their authors' distance, in Figures 1 and 2 we plot the frequency distributions of subsequent citations for Years t+1 through t+4 for all two-authored articles in the early and late periods separately, with the studies categorized by distance (measured over the four pre-publication years). In all of the eight samples more mass of the distribution of citations is concentrated in the highest category (10+ citations) among the close-coauthored articles than among the distant-coauthored articles. In other words, coauthored "blockbuster" papers tend more commonly to be the result of close-partner collaboration than distant partnerships. Similarly, in seven of eight comparisons the distribution is denser in the two lowest categories combined for the distant-coauthored articles. Looking at entire distributions rather than just the means strengthens the inference that Proposition I.1 is rejected by the data.

Simply examining means and comparing percentage distributions does not provide a complete test of the alternative models, since other factors may be correlated with the propensity to coauthor with distant scholars and with the inherent quality of the articles that a coauthor produces. As a more complete test of distance effects, we estimate an equation in which the productivity of a given coauthored article is a function of a variety of characteristics of the authors, including their distance from one another, as well as a number of the article's other characteristics. The estimating equation is:

$$(2) \qquad P_{a,t+j} = \gamma_0 + \gamma_1 LATE_a + \gamma_2 DISTANT_a + \gamma_3 LATE \cdot DISTANT_a + \sum_i \gamma_{4i} CITES_i + \gamma_5 EMPIRICAL_a$$

$$+ \gamma_6 AERPAGES_a + \gamma_7 JOURNAL_a + \gamma_8 MULTAUTH_a \text{ , } j=1,...,4 \text{ ,}$$

where a denotes an article, i denotes the first, second or third coauthor, LATE and DISTANT are self-explanatory indicator variables, and P is the article's productivity, measured by its post-publication citations.[11] Because of the skewness of P and its concentration at zero, we estimate (2) using Poisson regression. This method restricts predictions on the dependent variable to be nonegative integers, which must be the case, since P counts citations. The estimate of γ_2 indicates whether distant articles are more or less-cited than close coauthored papers in the early period, while γ_3 shows the direction and magnitude of the change in this effect over time. We expect $\gamma_2 > 0$, $\gamma_3 < 0$, and $\gamma_2 + \gamma_3 > 0$.

There are a number of characteristics of scholars that determine both their propensity to work with distant coauthors and the citation rate of the articles that they subsequently produce. In particular, more able authors may have more opportunities for travel that allow them to work with distant coauthors, and they also may produce inherently higher-quality research. To account for this factor, we compiled each coauthor's citations by other scholars in year t and denote them by CITES. In this way we can hold constant each individual's base-line productivity in looking for distance effects. Thus we are interested in how much a particular coauthorship adds to the individual productivity of the members of the team. Because the distribution of citations to individuals is highly skewed, in all estimating equations we include this covariate in the form of a vector of indicator variables, 10-49 citations, 50-99 citations, 100+ citations (with 0-9 citations the excluded category).

An article's characteristics, other than its inherent quality, may also affect its subsequent citations and do so differentially over time. For example, empirical research may be more or less cited than theoretical work, and it may be less likely to be accomplished via distant production. In the econometric analysis we thus classify each study as EMPIRICAL or not based on whether it included tables, or figures that depicted data. The particular journal where an article appears may both affect P and be an indicator of the inherent quality of an article. In the formal estimation we thus also include indicator variables for each journal. An article's length may be an indicator of its quality, so that we also obtained data on the number of pages in each article. To account for differences in page length across the journals, we adjust these and use the number of AER-equivalent pages.[12] Finally, since multiauthored articles may be cited differentially from two-authored papers, we included the indicator MULTAUTH.

The remainder of Table 2 lists the means of these other variables describing the coauthored and solo-authored articles. Comparing distant and close coauthorships, several differences are apparent. Distant coauthorships are less likely to be empirical in nature than are close-coauthored articles in these journals. Also, remembering that the excluded category is authors who receive fewer than 10 citations per year, it is noteworthy that both the first and second coauthors in distant partnerships are more heavily cited than the first or second partners in close coauthorships. Coauthored articles are somewhat longer than solo-authored papers; and solo authors are less well cited than first or second coauthors, especially distant coauthors.[13]

Before discussing the estimates of the parameters in (2), it is valuable simply to compute γ_2 and γ_3 without the other covariates, especially since some of the article-specific variables in (2) may affect P but may also be affected by distance (such as, for examples, the probability of the article being accepted in the particular journal, or its length). These estimates are shown in the first two columns of Table 3, first for the entire sample, then for a subsample excluding articles with more than two coauthors. The impact of coauthors' distance on an article's productivity is negative in all but the second post-publication year in the 1970s. At the very least we can conclude that this simple Poisson regression provides no evidence that distant coauthorships are more productive, rejecting Proposition I.1. If anything, this negative effect of distance becomes even stronger in the 1990s, directly contradicting Proposition I.3.

Columns (3) and (4) of Table 3 present Poisson regression estimates of γ_2 and γ_3 based on the complete specification in (2).[14] Estimates of the other coefficients from (2) are shown in Table 4, and most of them are unsurprising. Subsequent citations to a study, for example, are increasing in the number of citations that its first and second authors have previously received. Whether this means that the studies that those highly-cited authors produce are inherently more worthy, or whether there is simply a "Matthew Effect" (Merton, 1968) is unclear, but is not an issue. Only the citations to higher-order authors of multiply-authored studies generally have no significant impact on the study's recognition. In this sample, empirical research receives more subsequent citations, even though we hold constant the authors' prior citations; and prolixity pays off in productivity: Extra pages add to the article's eventual citations. There are distinct differences in subsequent citations to articles published in different journals even after adjusting for the authors' own citations. Finally, all else equal, studies published in the 1990s received fewer citations in their first two post-publication years, but more in the next two years, than those published in the 1970s. This may reflect a lengthening of publication lags and thus of the age of citations to published works in published articles.

The central questions in this study revolve around the impact of the ease of communication, as proxied by distance in coauthorship, on productivity measured by subsequent citations to the coauthored article. The results shown in Table 3 are absolutely clear: Holding constant a large variety of other measures, most important the coauthors' productivity, distant coauthorships are significantly less productive than close coauthorships (except for Year t+2 in the early period). Both γ_2 and $\gamma_2 + \gamma_3$ are negative, contrary to the theory. Moreover, while the estimates of γ_3 are either zero or negative, that makes little sense in light of the strongly negative effect of distance overall. Proposition I.1 is strongly refuted by the data, while the productivity of close and distant coauthorships showed no sign of converging during a period when the cost of communication fell dramatically. The estimates of (2) are striking evidence that distant coauthorship by otherwise identical coauthors, publishing articles of the same length and type in the same journals is less productive than coauthorship by near neighbors. Moreover, while easier communication between academics has generated additional collaborations, there is no evidence that their average productivity changed over time (so that implicitly the additional distant coauthorships that were undertaken were no more, and perhaps even less productive than those fewer that would have been written had the cost of communication not fallen).

Table 3 contains the principal empirical findings of this study. The results deserve more than the usual number of robustness checks, given how surprising they are. One possibility is that distant coauthors are increasingly using the time savings generated by lower communication cost to generate research that is not included in our sample (not published in these three journals) but that is published elsewhere and is of high quality (as proxied by our citation measures). If that were true, we would observe that $CITES_{i\,d}$ - $CITES_{i\,c}$ is increasing over time. Taking the fraction of coauthors receiving at least 10 citations in Year t as one measure of this difference, it equaled 0.205 in the 1970s, but fell to 0.082 in the 1990s.[15] While higher-impact authors are more likely to engage in distant coauthorship than other scholars, the impact of the distant coauthors' entire oeuvres fell relative to those of close coauthors between these periods.[16]

One problem with the results in Table 3, especially for the crucial interaction term LATE·DISTANT, is that there are relatively few articles in the sample, especially for years t+3 and t+4. While there are no other obervations to add, we can remove some of the sampling variance by estimating (2) with the dependent variables $P_{t+1}+P_{t+2}$, $P_{t+1}+P_{t+2}+P_{t+3}$ and $P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$, essentially measures of each article's "lifetime" citations.[17] The estimates of γ_2 and γ_3 from this revised specification are presented in the top panel of Table 5. The results underscore and strengthen the conclusions from Table 3. Distant coauthorships are uniformly less productive than close ones, with the differences being significant; and the productivity disadvantage of distant coauthorships increased between the 1970s and 1990s, at the same time that the declining cost of communication induced a sharp increase in distant coauthoring.

Another possibility is that we have defined the recent period too broadly, including in it some articles that could not have taken advantage of changes in communication technology because they were begun before those changes occurred. We can examine this potential problem by reestimating (2) over reduced samples that exclude the largest possible number of earlier years from the 1990s. Thus we estimate (2) to describe P_{t+1} excluding articles from 1992-95, describing $P_{t+1}+P_{t+2}$ excluding articles from 1992-94, $P_{t+1}+P_{t+2}+P_{t+3}$ excluding 1992-93, and $P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$ excluding 1992. In each case only one year from the 1990s is included, making the samples from the LATE period quite small. The results of this reestimation are presented in the second panel in Table 5. Despite the small samples from the 1990s, except for the comparison for $P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$ the effect of distance is generally negative, no less so for articles published in the single year from the 1990s than for those published in the 1970s.[18]

Tables 3 and 5 present the results of regressions based on the distance measure that emerges from our survey of co-authors. It is interesting to consider whether we would have obtained the same results had we simply relied on proximity at the time of the article's publication as the distance measure. The bottom panel of Table 5 answers this question. Equation (2) is reestimated using P_{t+1} and the same cumulative productivity measures as dependent variables, but substituting distance at publication for pre-publication distance.[19] The surprising overall negative effect of distance found in the previous specifications persists in the early period even with this incorrect measure of distance. In the later period, however, use of publication proximity alone yields evidence of a very weak positive effect of distance in the 1990s. It is possible that as publication lags have grown over time and the mobility of economists increased, the use of proximity at time of publication as a measure of distance for the entire production period has become less and less reliable.

The analysis thus far has implicitly considered the impact of only those changes, like the declining cost of communication, that have affected close and distant coauthorships differentially. During the same period, however, there have been other changes in technology that have likely affected the ease of co-authoring more generally. Moreover, access to these changes (e.g., word processing) may in turn be correlated with the covariates in (2) and the productivity measures. To circumvent these problems we expand the sample to include both the coauthored articles and the sample of solo-authored papers. We respecify (2) by adding the main effect, SOLO (equaling 1 if the article is solo-authored, 0 otherwise) and this variable interacted with LATE.[20]

The estimates of this expanded equation are presented in Table 6. Given the authors' and articles characteristics, self-matches (solo-authored papers) were less productive than coauthorships in the 1970s; and there is some evidence that their relative productivity decreased further in the 1990s. The appropriate comparisons of the main results in this table (the estimates of γ_2 and γ_3) are to the estimates in the righthand columns of the first row of Table 3 and to the top two panels of Table 5. They make it very clear that accounting for technological and other changes that might have altered incentives to choose coauthorship has essentially no impact on our conclusions: Distant coauthorship is less productive than close coauthorship, a deficiency that may have increased over time.

The data set contains additional information that we have not used: A substantial fraction of the authors are included in the sample two or more times, and in some cases at least one of their coauthorships is distant and one close (or at least the distance measure is not identical in all the person's appearances in the sample). This information allows us to control even more carefully for factors that might affect an article's impact. In essence, for each person whose distance status varies across his or her coauthored articles we can hold constant for the unobserved productivity that is not accounted for by the author's prior citations. We thus estimate fixed-effect Poisson regressions using subsamples of articles by authors who meet these criteria for inclusion in these reduced samples.

The parameter estimates are presented in Table 7, both for the P_{t+j} measures and for the cumulative measures $P_{t+1}+P_{t+2}$, $P_{t+1}+P_{t+2}+P_{t+3}$, and $P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$. These are based on estimates of the various models weighted by the importance of the observation in the samples.[21] Because of the stringent sampling criteria the numbers of articles included becomes quite small as j increases, as does the number of individuals whose works are included in these subsamples. The small samples make it quite unlikely that the parameter estimates will be highly significant. Despite these difficulties the estimates in the upper panel of Table 7 generally confirm the findings of Tables 3, 5 and 6. While the parameters on distance in the equations describing the P_{t+j} are typically insignificantly different from zero, all the interactions are negative and all of those in the equations describing cumulative citations are significantly negative. These results suggest that holding constant all the scholar's characteristics, both observable and unobservable, the choice d, while becoming more common in the 1990s, was still less productive than choosing c, and the gap was actually increasing.

The restricted samples that exclude multiply-authored articles are smaller still than those on which the estimates in the upper panel of Table 7 are based; and the number of authors included becomes tiny. Nonetheless, even in these small samples the parameter estimates shown in the bottom panel of Table 7 indicate a generally negative effect of coauthors' distance on the productivity of their article. While the interaction terms are sometimes positive, sometimes negative, overall they suggest little diminution in the productivity disadvantage of distant coauthorships.

The estimates presented in Tables 3 through 7 account for a variety of observable characteristics of the articles included in our sample and for both observable and unobservable characteristics of the authors of those studies. Despite our inclusion of this substantial array of controls, the results suggest quite strongly that Proposition I.1 is roundly rejected by the data. There is little doubt from the evidence of this sample that when communication is more difficult the productivity of those who overcome this difficulty and work together is lower than that of others who work together without the need to bear these costs. Moreover, while a decline in these costs increases scholars' propensities to work together at a distance, it certainly does not raise the productivity of those who choose to do so, and it may even have made their relative productivity lower than it was when communication cost was higher.

V. An Alternative Explanation

The empirical results from this unique test of the impact of communication cost and its decline on productivity in a particular activity are clear-cut and striking. Regrettably, except for the facts that the activity that relies on reduced communication cost has increased in frequency, and that its value may have declined relatively, they are also quite inconsistent with the predictions of our model. Proposition I.1, that distant coauthorship is more productive, was soundly rejected. While Proposition I.3, that distant coauthorship will show a relative decrease in productivity over time, may be supported, falling communication cost should generate convergence in productivity in the two activities. The implications of our results are not totally dissimilar to the depiction of the uses of technology in Figure 3. This is all quite disturbing. We might pose the problem revealed by our data as follows: Why are economists engaging in expensive distant partnering when it does not appear to be as productive as a less costly alternative?

A number of the responses that we received from some of the over 300 scholars whose coauthorship distance we ascertained using the low-cost communications methods of the 1990s provide a clue to this puzzle. While these comments are clearly self-selected in ways that we cannot determine, they are highly suggestive. Among them are:

> X and I were not in the same city/institution for 9 months in any of the preceding 4 years prior to publication of the articles. X and I were friends at Y during our graduate student days.
>
> X and I were not at the same institution or within 50 miles between 1988 and 1992. He was at Z and I was in W. We did, however, start working together in 1982 and 1983 when we were both at Y.
>
> No. But we were in the same city for 9 months 6 years prior to publication. (Specifically, we were graduate students together at Y.)
>
> If it is of any interest and help, X and I were in graduate school together at Y in the early 1980's.
>
> X and I started working on the paper in 1988, when we were both at Y. That same year we left for other institutions, and continued with the project, on and off, until the publication of the paper.
>
> No, not in the four years preceding publication. However, the paper evolved out of a project that was started while both of us were at Y.
>
> The answer is no. I was at M and then N, while X was at Y then at Z. HOWEVER, we were both at Y together for several years, and had coauthored papers prior to 1992.

These responses typify the unsolicited comments that we received from the authors whom we classified as distant during the four years before publication. In many cases the eventual coauthors had been friends in graduate school and welcomed the chance to resume their friendship <u>on a professional activity</u> years later. In other cases an idea that had been hatched in graduate school or in an earlier collegial relationship and then shelved was revived and brought to fruition. Presumably there are many such friendships and many such ideas that are renewable, not all of which do get renewed. The issue is how easier communication leads to their renewal and what that implies for the observed relative level of and change in the productivity of distant coauthorships.[22]

The comments that we received suggest that scholarly activity creates two streams of benefits: Production benefits, measured by citations to the work, and the consumption benefits realized by the co-authors in the production process. (The notion that personal relationships play a role in academic activities, and even in coauthorship, is not new, e.g., McDowell and Smith, 1992.) Part of the consumption value that is generated derives from interactions with colleagues who may be enjoyable intellectually and/or personally. These interactions take time away from the purely productive aspects of scholarship. This characterization suggests the following model of research. The scholar ranks the infinity of potential utilities from research projects as:

(3) $U(Y_i , C_i)$,

where i is a potential coauthor, Y is the expected income stream yielded by the project chosen, C is the consumption stream and U has the standard properties. Coauthorship alone yields consumption value: A solo-authored project is described by $U(Y_0 , 0)$. The scholar's production/consumption is constrained by the fact that working with a distant coauthor i takes time and money, which we represent by τ_{it} per period, with τ_{it} decreasing in t. In each period the scholar can only spend $1-\tau_{it}$ of the time in productive scholarship. If the utility-maximizing scholar chooses a close coauthor over a distant one, she will thus be able to devote more productive time to the project. Any distant coauthorship that we observe being chosen must, if it yields the maximum U, be one that also yields sufficiently high consumption value to overcome the time (and money) costs that it engenders.

Coauthored articles yield both research and consumption benefits. In the case of distant coauthorship, creating research benefits has a cost that is not experienced in close-coauthored papers. As a consequence, on the margin, while the net value to the researcher of time spent on close and distant coauthorship will be equal, the distant papers will have lower research content and higher consumption value. Given that our measure of productivity is research oriented, this suggests our first proposition:

Proposition II. 1: Distant coauthorships will be less productive in research terms than near ones.

Our second proposition relates to the dynamics of the research process. As τ_{it} decreases with falling communication cost, more distant coauthorships that are desirable for consumption purposes overcome their cost disadvantage and are taken up, leading to:

Proposition II. 2: The fraction of all coauthorships that are distant will increase as communication cost decreases (Same as Proposition I.2.)

Moreover, as technology improves and $\tau_{i\ t}$ decreases, the consumption costs of the distant coauthorship decrease, so that:

Proposition II. 3: The observed productivity of distant coauthorships will rise and approach that of close coauthorships.

This Proposition clearly follows, since as $\tau_{i\ t} \rightarrow 0$ the near and distant coauthorships become economically indistinguishable, and the incremental consumption value of distant partnering can be done with no loss in production.

The predictions of this production/consumption view of the choice of coauthors clearly accord with the fact of rising distant coauthorship, as did the pure production Model I. Unlike the earlier model, however, Model II is consistent with the greater scholarly return to close than to distant coauthorship that is the fundamental fact discovered in Section IV. This explanation fails only in that we found no evidence that the returns to distant and close coauthorship became more equal in recent period.

While the production/consumption model still leaves us with the puzzle of the failure of productivity to converge, it does fit the data better than an approach based on production alone. In addition, the richer model implies another prediction: Distant coauthorship, because it takes money, should be more prevalent among those with higher full-earnings. In effect, distant work is in part a good purchased by researchers. This prediction is borne out in the data. A probit that explains distance in two-authored articles by the measures of authors' citations (measures of their prior productivity and thus proxies for their professional earnings) suggests a significant positive relationship between prior citations and the propensity to write with a distant coauthor.[23]

Given that distant coauthorship appears to be a normal good, what do our results reveal about authors' willingness to pay for it? Using the estimate in Table 5, column (3), and the means of the articles' citations, and assuming that a scholar accrues half the citations to his/her joint work, scholars choosing distance implicitly forego 1.76 citations over a four-year period. Based on estimates of the effect of citations on salaries (Hamermesh et al, 1982), this deficit reduces a scholar's academic-year salary by 0.20 percent per year over the four years, other things equal. Assuming that the average salary is $80,000 per year, this calculation implies that scholars implicitly forego at least $640 when they choose to consume distant coauthorship.[24] By their choices scholars have revealed that working with a distant coauthor has substantial consumption value.

VI. Conclusion and Justications for Economic Welfare

Our results suggest that joint research in economics is increasingly being conducted by authors who work at long distance from one another. In the 1990's approximately one-fifth of all coauthored work was accomplished by authors who lived in different cities for the entire production period. In the 1970's, only 5 percent of the joint research had this property. It is interesting to speculate about the potential effect of this “distant research” on the equilibrium distribution of scholars across universities. From a researcher’s perspective, one of the advantages of being at a large central university is the contact provided with other productive scholars. But as distant research is facilitated, this competitive advantage of the large university may fall. A more even distribution of scholars across universities is likely to result. This result is similar to the prediction (not yet formally tested) by urban economists that the agglomeration benefits of cities have begun to fade under new communication technologies (Glaeser, 1994). Our results suggest, however, that this diminution in agglomeration may not increase productivity.

The initial motivation for this study was to analyze the effects of general technical improvements, particularly whether their impacts on productivity can be discerned at the micro level. The evidence here suggests that, contrary to expectations, improved technology generates lower-quality output per unit. By facilitating communication, technology has perhaps lowered the relative price of the consumption benefits of joint research. Given the public-good spillovers of research, the consequence of the new technology may be an improvement in the private welfare of economists and a simultaneous fall in social welfare. One can only hope that any such loss in social welfare incurred in the production process will be more than compensated for by the gain in efficiency from technology-induced improvements in the dissemination of the research once produced. In the new equilibrium the wages of academic economists (perhaps all academics) may fall as scholars consume increased psychic income from distant partnering rather than spending time improving the quality of their research.

REFERENCES

Robert Allen, "Testimony," House Subcommittee on Telecommunications and Finance, 102nd Congress, 1st session, May 10, 1995.

Frances Cairncross, The Death of Distance. Boston: Harvard Business School Press, 1997.

Partha Dasgupta, "Patents, Priority and Imitation or, the Economics of Races and Waiting Games," Economic Journal, 98 (1988): 66-80.

John DiNardo and Jörn-Steffen Pischke, "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?" Quarterly Journal of Economics, 112 (1997): 291-304.

Mark Doms, Timothy Dunne and Kenneth Troske, "Workers, Wages and Technology," Quarterly Journal of Economics, 112 (1997): 253-90.

Joseph Eisenhauer, "Multi-Authored Papers in Economics," Journal of Economic Perspectives, 11 (1997): 191-2.

Edward Glaeser, "Cities, Information and Economic Growth," Cityscape: A Journal of Policy Development and Research, 1 (1994): 9-47.

Robert Gordon, "Monetary Policy in the Age of Information Technology: Computers and the Solow Paradox," unpublished paper, Northwestern University, May 1998.

Nathalie Greenan and Jacques Mairesse, "Computers and Productivity in France: Some Evidence," National Bureau of Economic Research, Working Paper No. 5836, November 1996.

Zvi Griliches, "Research Costs and Social Returns: Hybrid Corn and Related Innovations," Journal of Political Economy, 66 (1958): 419-31.

Daniel Hamermesh, George Johnson and Burton Weisbrod, "Scholarship, Citations and Salaries," Southern Economic Journal, 49 (1982): 472-81.

Alan Krueger, "How Computers Changed the Wage Structure: Evidence from Microdata, 1984-1989," Quarterly Journal of Economics, 108 (1993): 33-60.

Edward Lazear, "Globalization and the Market for Teammates," National Bureau of Economic Research, Working Paper No. 6570, May 1998.

John McDowell and Janet Smith, "The Effect of Gender-Sorting on Propensity to Coauthor: Implications for Academic Promotion," Economic Inquiry, 30 (1992): 68-82.

Robert Merton, “The Matthew Effect in Science,” Science, 159 (January 5, 1968): 56-63.

Catherine Morrison, “Assessing the Productivity of Information Technology Equipment in U.S. Manufacturing Industries,” Review of Economics and Statistics, 79 (1997): 471-81.

Sharon Oster and Daniel Hamermesh, “Age and Productivity Among Economists,” Review of Economics and Statistics, 80 (1998): 154-6.

Pitney-Bowes, “Survey of Fax Usage in the United States,” PC Week, 14 (September 22, 1997): 88.

Richard Quandt, “Some Quantitative Aspects of the Economics Journal Literature,” Journal of Political Economy, 84 (1976): 741-56.

Raymond Sauer, “Estimates of the Returns to Quality and Coauthorship in Economic Academia,” Journal of Political Economy, 96 (1988): 855-66.

Gary Stix, “The Speed of Write,” Scientific American (December 1994): 106-11.

Manuel Trajtenberg, “The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners,” Journal of Political Economy, 97 (1989): 444-79.

Jeffrey Wooldridge, “Quasi-Likelihood Methods for Count Data,” in M. Hashem Pesaran and Peter Schmidt, eds., Handbook of Applied Econometrics, Vol. 2. Oxford: Blackwell, 1997.

Table 1. Coauthorship and Distance Probabilities, 1970s and 1990s

	Period	
Outcome	**Early (1970-79)**	**Late (1992-96)**
Coauthored/Total articles	.301	.617
Fraction two-authored|Coauthored	.883	.767
Distant|Coauthored (at publication)	.438	.546
Distant|Coauthored (at publication), 2 authors only	.462	.582
Distant|Coauthored (throughout process)	.056	.203
Distant|Coauthored (throughout process), 2 authors only	.056	.221
Number of coauthored articles	383	430

Table 2. Means of Variables Describing Economics Articles (standard deviations in parentheses)

	Coauthored Articles						**Solo Articles**	
	Early			**Late**			**Early**	**Late**
	All	**Two-authored**		**All**	**Two-authored**			
Variable		**Distant**	**Close**		**Distant**	**Close**		
Article citations:								
Year t+1	1.765	1.263	1.799	1.916	1.369	1.848	1.153	1.277
	(2.316)	(1.240)	(2.410)	(2.729)	(1.712)	(2.664)	(1.668)	(1.567)
Year t+2	3.034	3.789	2.975	3.456	2.745	3.512	2.186	2.339
	(3.436)	(3.326)	(3.398)	(4.340)	(2.945)	(4.793)	(3.095)	(2.642)
Year t+3	3.410	2.789	3.530	5.153	3.767	5.258	2.795	3.043
	(4.057)	(2.511)	(4.210)	(6.912)	(4.275)	(7.329)	(3.838)	(3.183)
Year t+4	3.551	2.789	3.671	6.307	4.292	6.320	2.930	3.433
	(4.976)	(3.066)	(5.184)	(9.858)	(5.353)	(10.382)	(4.097)	(3.641)
Empirical	.517	.421	.498	.549	.425	.533	.419	.596
AER-equivalent pages	12.99	13.97	13.00	18.58	18.56	18.65	12.62	17.52
	(5.62)	(4.88)	(5.75)	(5.31)	(4.63)	(5.70)	(3.82)	(5.17)
Author citations, year t:								
First author:								
10-49	.311	.316	.317	.372	.438	.339	.205	.326
50-99	.063	.158	.050	.153	.123	.163	.037	.071
≥100	.078	.158	.072	.112	.164	.105	.060	.071
Second author:								
10-49	.272	.421	.251	.342	.384	.319		
50-99	.023	.053	.025	.093	.041	.109		
≥100	.034	.053	.034	.088	.123	.074		
Higher-order author:								
10-49	.200	-----	-----	.270	-----	-----		
50-99	.089	-----	-----	.030	-----	-----		
≥100	.044	-----	-----	.120	-----	-----		
N =	383	19	319	430	73	257	215	267

Table 3. Effects of Distance in (2)[1]

	Independent Variable			
Dependent Variable	DISTANT	DISTANT·LATE	DISTANT	DISTANT·LATE
	All coauthorships			
P_{t+1}	-.310	-.037	-.608	.120
	(.199)	(.224)	(.206)	(.231)
P_{t+2}	.208	-.404	-.021	-.300
	(.121)	(.148)	(.123)	(.150)
P_{t+3}	-.240	-.074	-.466	.074
	(.139)	(.161)	(.140)	(.162)
P_{t+4}	-.261	-.127	-.461	.052
	(.138)	(.169)	(.140)	(.172)
	Excluding multiple coauthorships			
P_{t+1}	-.354	.054	-.601	.199
	(.208)	(.236)	(.210)	(.237)
P_{t+2}	.242	-.488	.036	-.383
	(.122)	(.154)	(.124)	(.155)
P_{t+3}	-.235	-.098	-.450	.056
	(.141)	(.165)	(.142)	(.167)
P_{t+4}	-.275	-.112	-.425	-.011
	(.140)	(.176)	(.142)	(.178)

---

[1] Standard errors in parentheses here and in Tables 4-7. The first two columns include only the distance measures and the indicator for LATE. The last two columns include the full set of covariates in (2).

Table 4. Coefficient Estimates of the Covariates in (2)

Variable		P_{t+1}	P_{t+2}	P_{t+3}	P_{t+4}
$CITES_1$:	10-49	.418	.367	.302	.292
		(.064)	(.051)	(.049)	(.051)
	50-99	.707	.575	.582	.738
		(.082)	(.069)	(.064)	(.069)
	≥100	.845	.890	1.001	.982
		(.082)	(.065)	(.059)	(.062)
$CITES_2$:	10-49	.251	.192	.136	-.100
		(.060)	(.048)	(.045)	(.049)
	50-99	.228	.441	.240	.312
		(.107)	(.078)	(.077)	(.080)
	≥100	.819	.454	.519	.343
		(.088)	(.079)	(.073)	(.082)
$CITES_3$:	10-49	.091	.201	-.017	-.147
		(.129)	(.108)	(.108)	(.119)
	50-99	.215	.086	-.040	-.072
		(.243)	(.220)	(.207)	(.281)
	≥100	.056	.107	.638	-.155
		(.187)	(.171)	(.121)	(.173)
Multiauthored		.154	-.014	-.061	.095
		(.082)	(.071)	(.066)	(.066)
Empirical		.247	.151	.139	.190
		(.054)	(.043)	(.040)	(.042)
AER Pages		.023	.020	.020	.025
		(.004)	(.003)	(.003)	(.003)
AER		.298	.443	.349	.375
		(.069)	(.057)	(.053)	(.058)
JPE		.124	.159	.159	.274
		(.073)	(.061)	(.056)	(.060)
LATE		.205	-.092	.193	.364
		(.063)	(.050)	(.047)	(.049)
Pseudo-R^2		.091	.099	.130	.135
Number of observations		813	723	645	546

Table 5. Effects of DISTANT and LATE , Alternative Specifications

	Independent Variable			
	DISTANT	**DISTANT·LATE**	**DISTANT**	**DISTANT·LATE**
	All coauthorships		**Excluding multiple coauthorships**	
	Distant during four years before publication			
Dependent Variable				
$P_{t+1}+P_{t+2}$	-.203	-.189	-.162	-.219
	(.106)	(.127)	(.107)	(.131)
$P_{t+1}+P_{t+2}+P_{t+3}$	-.303	-.128	-.272	-.131
	(.084)	(.102)	(.085)	(.105)
$P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$	-.342	-.117	-.298	-.147
	(.072)	(.094)	(.073)	(.098)
	Alternative Definitions of LATE			
Dependent Variable (Sample)				
P_{t+1}	-.620	.124	-.615	.161
(1970-9, 1996)	(.205)	(.283)	(.212)	(.293)
$P_{t+1}+P_{t+2}$	-.215	-.072	-.178	-.348
(1970-9, 1995)	(.106)	(.183)	(.108)	(.213)
$P_{t+1}+P_{t+2}+P_{t+3}$	-.308	.220	-.315	-.084
(1970-9,1994)	(.084)	(.257)	(.086)	(.129)
$P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$	-.352	.290	-.343	.327
(1970-9,1993)	(.072)	(.109)	(.074)	(.117)
	Distant at time of publication			
Dependent Variable				
P_{t+1}	-.126	.171	-.130	.201
	(.081)	(.111)	(.083)	(.120)
$P_{t+1}+P_{t+2}$	-.110	.154	-.099	.103
	(.049)	(.071)	(.051)	(.076)
$P_{t+1}+P_{t+2}+P_{t+3}$	-.102	.284	-.098	.220
	(.037)	(.057)	(.039)	(.061)
$P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$	-.102	.201	-.107	.106
	(.031)	(.053)	(.032)	(.057)

Table 6. Effects of Distance in Equations Including Solo-Authored Studies[1]

	Independent Variable								
	DISTANT	DISTANT ·LATE	SOLO	SOLO ·LATE	DISTANT	DISTANT ·LATE	SOLO	SOLO ·LATE	N
	All coauthorships				Excluding multiple coauthorships				
Dependent Variable									
P_{t+1}	-588	.103	-.180	.062	-.574	.176	-.236	.144	1295
	(.205)	(.230)	(.079)	(.102)	(.210)	(.237)	(.080)	(.186)	
$P_{t+1}+P_{t+2}$	-.185	-.187	-.144	.025	-.132	-.223	-.196	.060	1162
	(.106)	(.127)	(.047)	(.064)	(.106)	(.131)	(.048)	(.066)	
$P_{t+1}+P_{t+2}+P_{t+3}$	-.303	-.124	-.101	-.202	-.264	-.134	-.149	-.166	1022
	(.084)	(.102)	(.035)	(.051)	(.085)	(.105)	(.036)	(.053)	
$P_{t+1}+P_{t+2}+P_{t+3}$ $+P_{t+4}$	-.328	-.123	-.124	-.320	-.281	-.152	-.176	-.291	881
	(.072)	(.094)	(.029)	(.046)	(.073)	(.097)	(.029)	(.048)	

---

[1] Each equation includes all the covariates in (2) plus the indicator LATE.

Table 7. Weighted Fixed-Effects Estimates of the Impact of Distance (Based on distance during the four years before publication)

Dependent Variable	DISTANT	DISTANT·LATE	Dependent Variable	DISTANT	DISTANT·LATE	Number of: Articles (Authors)
All coauthorships						
P_{t+1}	-.269	-.182				203
	(.417)	(.468)				(95)
P_{t+2}	.175	-.754	$P_{t+1}+P_{t+2}$	.107	-.656	161
	(.274)	(.306)		(.228)	(.258)	(72)
P_{t+3}	.140	-.323	$P_{t+1}+P_{t+2}+P_{t+3}$	.112	-.388	130
	(.324)	(.364)		(.187)	(.219)	(61)
P_{t+4}	.042	-.926	$P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$	-.014	-1.261	75
	(.323)	(.550)		(.166)	(.268)	(35)
Excluding multiple coauthorships						
P_{t+1}	-.541	.050				101
	(.313)	(.348)				(45)
P_{t+2}	-.322	-.094	$P_{t+1}+P_{t+2}$	-.393	-.109	76
	(.210)	(.243)		(.176)	(.205)	(31)
P_{t+3}	-.813	.905	$P_{t+1}+P_{t+2}+P_{t+3}$	-.625	.530	62
	(.230)	(.268)		(.142)	(.170)	(26)
P_{t+4}	-.259	-.205	$P_{t+1}+P_{t+2}+P_{t+3}+P_{t+4}$	-.354	-.539	44
	(.264)	(.364)		(.137)	(.194)	(19)

FOOTNOTES

[1] Gordon (1998) provides a thoughtful analysis of why one should not have expected much effect, at least from computers alone.

[2] That technology has changed the nature of scholarly communication should be trivially obvious to anyone involved in the enterprise. Stix (1994) presents a good discussion of some aspects of this change.

[3] The dollar-valuation is indirect. But there is substantial evidence (e.g., Hamermesh et al, 1982) that the direct measures of productivity that we use in the empirical work here translate into higher pay for the scholars who are objectively more productive.

[4] Throughout we assume, following the evidence in Sauer (1988), that the returns to a scholar from an otherwise identical coauthored article are equal to 1/N times the rewards to solo-authorship, where N is the number of coauthors.

[5] The four-year cut-off is admittedly quite arbitrary. It is possible that a few of the coauthors whom we classify as distant were together for a year or more and were working on the project five years or more before publication. It is highly unlikely that they had completed all the substantive work five years before publication, so that classifying them as distant even in this case allows us to capture the notion that their productivity would have been enhanced by easier communication.

[6] We are aware of the irony of using these methods, which would have been impossible fifteen years ago, to examine the potential impact on scholarly productivity of declining communication cost. Nonetheless, until many years after this paper appears in print one will be unable to determine whether our distant collaboration, which would have been impossible if communication cost were higher, added to our productivity as we define it here.

[7] The letter/email message to each asked, "... we wonder if you could let us know whether you and your coauthor were located at the same institution or in institutions within 50 miles of each other for any 9-month period in the 4 years preceding publication of your paper."

[8] For the 1970s we sampled one-fourth of the issues of each journal. For each journal each month of publication had the same representation in our sample.

[9] Some researchers have proposed the increasing ease of communications to explain the rising propensity of scholars to work together. In the sample of leading publications here, if the rate of distant coauthorship had remained the same, and if all of the distant coauthors would have otherwise published alone, the coauthorship rate would have risen from 0.301 to 0.574. The growing propensity to coauthor with distant colleagues can account for at most 15 percent of the rising fraction of articles that are coauthored.

[10] Comparing solo- and coauthored papers, the former are cited less frequently, with the latter receiving 30 percent more citations per paper in the early period, 26 percent more in the later peiod. Since the only evidence on the issue (Sauer, 1988) suggests that the pecuniary gains to a citation to a coauthored paper are half that to a solo-authored article, one would expect that coauthoring costs proportionately less time. In terms of the model, this would require the scholar choosing coauthorship to be able to write proportionately more articles per time period.

[11] To conserve degrees of freedom, for multiply-authored articles the vector $CITES_3$ is measured based on the citations in year t to the most highly-cited among the third or higher-numbered authors.

[12] The adjustment was done by scanning two pages that contained neither equations nor tables, and doing a word count. The AER-equivalent for the JPE is 0.718, for the QJE, 0.589.

[13] Although not relevant for this study, it is interesting to note the growing verbosity (editorial laziness?) of all types of articles, s, c and d.

[14] Because the distributions of the P_{t+j} in the samples are typically overdispersed one might wish to apply a negative binomial estimator. We did this, with results that do not differ greatly from those presented here. Similarly, ordinary least squares estimates do not alter the general conclusions. We concentrate on the Poisson estimates due to the evidence of the estimator's general robustness (Wooldridge, 1997).

[15] Including only those coauthors with at least 50 citations per year, the difference was 0.12 in the 1970s, 0 in the 1990s.

[16] The failure to find convergence in the quality of coauthored articles classified by distance might be due to the possibility that an increasing fraction of the distant papers involve coauthorship between a senior person and a very junior one (perhaps a former graduate student). While we cannot obtain information on the experience of all the authors, this possibility does not seem supported by the data. Let W be the difference between the fraction of first and second authors who have at least 10 citations in Year t. $W_{d,EARLY} = 0.105$, $W_{c,EARLY} = 0.129$;$W_{d,LATE} = 0.077$, $W_{c,LATE} = 0.105$. The gap between (this one measure) of the citation counts of the first two coauthors fell by 0.028 among distant coauthors, 0.024 among close coauthors. This suggests no relative change in the differences in professional impacts within pairs of coauthors classified by distance.

[17] Obviously we are not measuring the true lifetime of the article's productivity. The right truncation should not be a problem, however, because of the very high autocorrelation of citations to individual articles. Thus, for instance, the rank correlation of P_{t+3} and P_{t+4} is 0.72, while the Pearson correlation is 0.87. The first-order autocorrelations are increasing in j. More important, there is no reason to assume that these autocorrelations differ between distant and close coauthored papers, or that they have changed differentially over time between close and distant coauthored articles.

[18] This basic conclusion is strengthened if we use larger samples from the 1990s, for example, if we describe $P_{t+1}+P_{t+2}$ using data from 1970-79 and 1994-95, $P_{t+1}+P_{t+2}+P_{t+3}$ using data from 1970-79 and 1993-94, etc.

[19] As an example, one of us is included in the sample twice. In both cases the (two-authored) coauthorship is listed at distant at time of publication, but close using the more careful measure. In one case the authors were at the same institution until 5 months before publication, in the other until 16 months before publication.

[20] This specification is essentially a "triple-difference," but one that controls for a host of other variables. It is distinguished from the standard triple-difference method, which reads the effect of interest from the three-way interaction in an equation that also includes three main effects and three two-way interactions, because d and c are both branches of the choice to coauthor, while s and c are alternatives that do not involve distant work. This means that the three-way interaction is identically zero, as is one arbitrarily chosen two-way interaction.

[21] Some articles appear more than once in these samples because, for example, two of their authors are included. Since we assume that the productivity measures reflect each article's impact, to avoid weighting these articles more heavily we calculate the parameter estimates by weighting each article by the inverse of the number of times it appears in the subsample. The signs, sizes and significance (or lack thereof) of the parameter estimates do not change very much if we use unweighted Poisson estimates.

[22] Only one of the many authors who responded said he had never met his coauthor. In that case, however, a third coauthor was the intermediary between the two unacquainted authors. We know of only one two-authored publication, not in our sample, in which the authors never met before the final version of the study was accepted for publication (Oster and Hamermesh, 1998).

[23] Testing the coefficients on the six indicator variables for the citations of authors of two-authored papers in an equation that also includes LATE yields $\chi^2(6) = 14.22$, with a p-value of .03. Except for second authors with between 50 and 100 citations, all the coefficients were positive, implying a greater propensity for distant coauthorship than among scholars with very few citations.

[24] At least, because the calculation only accounts for the first four years of citations to the distant-coauthored article. Given the evidence that the median age of cited articles in economics was six years at one point (Quandt, 1976), the total implicit cost may be at least twice this high. Also, to the extent that the scholar incurs out-of-pocket costs in distant coauthorships, the calculation is understated further.