The Internet and Local Wages: Convergence or Divergence?*

Chris Forman
College of Management
Georgia Institute of Technology
800 West Peachtree St. NW
Atlanta, GA 30308
chris.forman@mgt.gatech.edu

Avi Goldfarb
Rotman School of Management
University of Toronto
105 St George St
Toronto, ON M5S 3E6
agoldfarb@rotman.utoronto.ca

Shane Greenstein
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston IL 60208
greenstein@kellogg.northwestern.edu

February 2009

Abstract

Did the diffusion of the Internet lead to convergence or divergence of local wages? We examine the relationship between business use of advanced Internet technology and regional variation in US wage growth between 1995 and 2000. We show that business use of advanced Internet technology is associated with wage growth but find no evidence that the Internet contributed to regional wage convergence. Advanced Internet technology is only associated with wage growth in places that were already well off in terms of income, education, population, and industry. Overall, advanced Internet explains one-quarter of the difference in wage growth between these counties and all others.

Keywords: wage growth, convergence, divergence, information technology, Internet
JEL Classification: O33, R11

* We thank Mercedes Delgado and seminar participants at the University of Colorado, Washington University, Arizona State University, the NBER Productivity Lunch, and WISE 2008 for comments. We also thank Harte Hanks Market Intelligence for supplying data. All opinions and errors are ours alone.

1. Introduction

Widespread evidence indicates that investment in information technology (IT) in the 1990s produced gains in US productivity and economic growth at the national, industry, and firm levels. Equally substantial evidence raises questions about whether the benefits of IT investment were experienced everywhere. In particular, new IT investments had the greatest effects on productivity for industries that were already IT-intensive and for workers with more education and skills. Yet, those findings leave open some fundamental questions about the variance in growth of incomes: Did those IT investments contribute to regional inequality in US wages? Specifically, was there a convergence or divergence related to frontier use of IT?

The question arises with special saliency in the 1990s because the new IT investments of that era—and particularly the rise of the commercial Internet—facilitated long-distance communication. One view hails the Internet as a great enabler of economic growth, particularly for low-density regions. This perspective hypothesizes that increased communication breaks the link between local investment, productivity, and wage growth, leading to a convergence of wages across regions. While this view has received widespread attention—e.g. Cairncross's (1997) *The Death of Distance* and Friedman's (2005) *The World is Flat*—and some support from international data on the globalization of services (OECD 2006; Arora et al. 2001), lack of regional data has prevented systematic testing within the United States.

A contrasting perspective casts the Internet as a technology that exacerbates existing inequalities in wages between urban/rural and frontier/mainstream users of IT, and consequently leads to divergence in wages across regions. This view argues that the Internet resembles prior generations of IT (Moss and Townsend 1997; Kolko 2002). Specifically, effective use of frontier IT relies on the presence of skills in the labor market, and wage gains are greatest for workers in skilled occupations, which are more likely to be found in rich urban areas (Kolko 2002). In line with this view, Wellman (2001) argues that the Internet primarily benefits local communication because social networks are local. In addition, Glaeser and Ponzetto (2007) argue that low communication costs help rich, idea-producing areas more than poor, goods-producing areas.

While a lively debate has ensued, little data on business use of the Internet has informed the discussion. We address this gap. Using a comprehensive data set of business Internet use, we construct measures of regional investment in advanced Internet use through late 2000. Our advanced Internet use measure includes a set of frontier applications, such as e-commerce or e-business, going beyond basic applications such as e-mail or Web browsing. In contrast to earlier research, we study a margin of investment and a time period in which IT facilitated communication over long distances and so, potentially, affected economically isolated work. We focus on studying advanced Internet use, a margin of investment that may require the deep labor pools and other complementary resources found primarily in cities. Finally, we connect our Internet data to measures of local economic performance, particularly wages paid by local businesses.

Our econometric approach compares a location's economic performance before advanced Internet technology diffused (i.e., 1995) to its performance after diffusion (i.e., 2000). That is, we use a difference-in-difference econometric estimation approach to identify the relationship between the variance of advanced Internet investment and the variance in regional economic outcomes. Our initial specification assumes that aggregate investment decisions by local establishments are exogenous to wage growth. We find that advanced Internet investment by early adopters is associated with an increase in county-level wages. This positive correlation remains robust to numerous specifications and changes in controls.

We address the assumption that investment is exogenous. First, although we add many controls for factors known to shape investment decisions, the results do not change. Second, we directly address what we consider to be the most likely issue: Omitted variables bias at the location level. The timing of effects points to the Internet as a key driver. We find no positive relationship between early wage growth (during 1990-1994) and the counties that were later adopters of advanced Internet technology. Furthermore, for 1999 to 2005, there is no relationship between later wage growth and counties that were adopters of advanced Internet as of 2000. The early wave of Internet adoption appears to be associated with a one-time change in wage growth across US locations during the late 1990s. We augment this falsification test with instrumental variable estimates that provide further evidence in support of our baseline results.

Our most interesting finding suggests that the Internet caused a divergence of wages. We find a stronger correlation between wage growth and advanced Internet use in counties that already were doing well on a variety of measures. In particular, we find that advanced Internet use is especially correlated with wage growth in the 180 counties that, as of 1990, had a population over 100,000 and were in the top quartile in income, education, *and* fraction of firms in IT-intensive industries. Overall, while the Internet explains just 1% of the wage growth in the average county in our sample, it explains 25% of the difference in wage growth between the 180 counties that were already doing well and all other counties.

Once again, we consider omitted variable bias. We find it difficult to speculate about which unmeasured regional-specific mechanism led to the results we find. Counties that were not in this group—even leading adopters—did not experience wage growth associated with advanced Internet use. At the same time we find it easy to provide an explanation that assigns causality to the Internet.

A scatterplot of the raw data forecasts our core results. Figure 1a shows the relationship between advanced Internet use and local wage growth for all types of counties in the data. Careful observation will show that the regression line is upward sloping (it is also significantly positive), but advanced Internet use is clearly not a core explanation of wage growth in the full sample. In contrast, Figure 1b compares the 180 counties that were already doing well with the other counties. For the 180 counties that were already doing well, advanced Internet use is strongly correlated with wage growth; for the other counties, there is no relationship between advanced Internet use and wage growth in the raw data.[1] Advanced Internet use allowed counties that were doing well (i.e., counties with high income, population, education, and agglomeration of IT-intensive firms) to do even better. In contrast, widespread advanced Internet usage is not correlated with wage growth in smaller, poorer, less educated counties with fewer IT-intensive firms. Thus, although all counties had wage growth, advanced Internet use led to a divergence in wages between affluent and poorer counties. Of course, our analysis goes far beyond this scatter-plot, but the intuition continues to hold after a wide battery of corrections and tests.

Our results have important public policy implications. We find evidence that use of the Internet in business yielded income gains in many locations other than a few locations displaying agglomeration of technology production – such as Santa Clara, Boston, and Seattle. This finding runs counter to portrayals of growth that emphasize agglomeration of production in technology-intensive industries. At the same time, we also reject the hypothesis that Internet investment yielded similar income gains in all locations. We find that investments in advanced Internet use were associated with wage growth only in those counties that are already doing well, *and* we find little evidence that the Internet had much impact in rural areas. In this way, our findings

[1] To construct Figure 1, we truncated the picture and consequently removed some counties with very low and very high Internet use. The results are qualitatively similar when we include these counties, though visually not as clean.

contradict the motivations for a wide array of infrastructure policies subsidizing business Internet use outside of urban areas.

1.1 Related Literature

Our study contributes to a large macroeconomic literature on regional convergence and divergence.[2] We complement recent work that has examined how cross-sectional variance in factors such as education and industry composition contribute to convergence and divergence (e.g., Higgins, Levy, and Young 2006). In a related paper, Glaeser and Ponzetto (2007) show that an increase in the share of skilled occupations is associated with greater local wage growth. They do not empirically focus on IT. Indeed, no research in the convergence/divergence debate has investigated the links between regional convergence/divergence and the Internet-led investment boom of the 1990s. This gap in understanding requires attention because the IT investment boom was an important factor in the late 1990s economic experience. It also has been associated with growth and productivity gains at other levels of aggregation, such as the nation, industry, and firm.[3]

Forman, Goldfarb, Greenstein (2008) found evidence of localization of investment in process innovations using Internet technologies. Establishments substitute away from internal labor into market labor when local wages (supply) are lower (thicker). That evidence suggests that the impact of Internet investment may have an important local component. However, this work focused on firm decision making and did not fully trace the connection between Internet investment and wage growth. That leaves a gap.

---

[2] See Magrini (2004) for a recent survey. For other research on the causes of convergence/divergence in regional growth see, e.g., Glaeser et al (1992), Barro and Sala-i-Martin (1991), and Higgins, Levy, and Young (2006). The literature linking technology to convergence across countries dates from Gershenkron (1962).

[3] IT-using industries and firms had exceptionally good macroeconomic performance in the 1990s, measured at the national (e.g., Jorgenson, Ho, and Stiroh 2005), industry (e.g., Stiroh 2002), firm (e.g., Brynjolfsson and Hitt 2003), and establishment (e.g., Bloom, Sadun, and Van Reenen 2007) levels.

The literature on IT investment and regional growth has only partially addressed this gap by focusing on the connections between IT and regional wage growth, though it has not focused on the role of the Internet in particular. For example, Beaudry, Doms, and Lewis (2006) focus on how local skills and variations in personal computer (henceforth PC) use over 1980–2000 correlate with wage growth. Similarly, Kolko (1999, 2002) finds that IT use over 1977–1995 is associated with the fastest employment growth in agglomerated areas due to the presence of local skills.

In comparison, our paper focuses on the Internet as a distinct factor, and we examine a later time period, because of the overwhelming number of popular writers that assign responsibility to the Internet for the economic prosperity of the late 1990s – such as, mentioned earlier, Cairncross (1997) and Friedman (2005). Moreover, prior work (Forman, Goldfarb and Greenstein 2005) led us to believe it was plausible that the increasing communication capabilities of the Internet would make convergence more likely. In each case, however, the hypothesis was assumed rather than tested.

In another related paper, Aral, Wu, and Morabito (2007) use firm-level data from Italy to show that enterprise resource planning (ERP) is most associated with productivity gains in regions with weak human capital and technological infrastructure. We find contrasting results to theirs at the local level, perhaps because (1) we focus on the Internet not on ERP, (2) our econometric strategy leads to sharper estimates, or (3) our results simply reflect differences between Italy and the United States.

While prior work has not examined whether the Internet leads to convergence, a number of recent papers have demonstrated how communicating over the Internet may lower the costs of engaging in economic activity in geographically isolated regions. For example, it has been shown that use of the Internet lowers the costs of retail shopping for isolated consumers (Forman, Ghose, and Goldfarb 2009) and stimulates greater job migration (Stevenson 2006). Furthermore, it is widely accepted that lower communication costs enabled by IT and the Internet have enabled the delivery of a set of tradable services at significant distances from the point of final demand (Arora and Gambardella 2005; OECD 2006).

Our paper shares similarities with the literature on skill-biased technical change, but there are also significant differences.[4] Research on skill-biased technical change generally tests the premise that changes in technology use alter the demand for skilled labor, thereby shifting the wage distribution in favor of skilled occupations (e.g., Autor, Katz, and Krueger 1998; Autor, Levy, and Murnane 2003). By demonstrating that wage gains are greatest in locations with more highly skilled workforces, our results are consistent with this prior work. Yet, the variance in our data across regions explores a dimension often overlooked by prior studies. Namely, the variance arises from cross-sectional differences in average incomes, education, and other factors across locations, rather than variance in skills across occupations. In addition, our results suggest that a highly skilled labor force alone is insufficient for a location to realize wage gains from Internet investments; there also must be other factors that shape local labor markets, such as the right combination of high population, income, and industry composition.

While we make this novel connection, we do not fully link it to the literature on biased technical change, mainly due to data limitations. For example, we cannot examine whether wage gains are greatest for high- or low-skilled occupations within a county, nor can we examine how Internet use changes the wage distribution within a location. Our findings will raise questions about such connections, and we leave that for future work.

[4] See Autor (2001) for a survey and Xiang (2005) for a recent example.

2. The Internet and the Localization of Growth

We measure the effects of advanced Internet use on convergence by proceeding in two broad steps. We begin by measuring the average relationship between Internet use and wage growth across all counties. Next, we identify convergence or divergence by examining whether advanced Internet investment led to faster growth in areas that were doing well.

Step 1: Advanced Internet use and local wage growth: Basic economic reasoning combined with findings from a range of prior studies suggests that use of the Internet may be associated with accelerated local wage growth through two mechanisms. First, productivity advances at the establishments using advanced IT raise the demand for labor at those establishments. If enough local firms become more productive through Internet use, local labor demand rises and local wages should be higher. Second, Internet use may lead to wage growth through narrower mechanisms. For example, increases in IT demand may put pressure on local markets for skilled IT workers such as programmers, database administers, and consultants. If the local supply of these occupations is inelastic, then increases in the demand for IT may translate into higher wages for these occupations. Over time, this mechanism may also put upward pressure on wages in other occupations that draw from a similar labor pool.[5]

To measure the impact of the Internet on local wages, we use a difference-in-difference identification strategy, comparing the wages of a time period before advanced Internet technologies diffused (1995) to those of a period when we observe use (2000). We take advantage of the fact that many regional features that shaped labor markets and enterprises in 1995 had not changed by 2000. Our endogenous variable will be Y_{it}, which represents the level of wages in a particular region (i) and year (t). Only a small set of research establishments

[5] We are unable to identify which of these explanations is most likely due to data constraints: We could find no reliable county-level data on 1995 wages for IT workers.

employed advanced Internet use in 1995, and therefore we set this variable to zero in 1995. Our approach yields the following panel regression:

$$(1) \qquad Log(Y_{it}) = \alpha_1 X_{it} + \alpha_2 Z_{it} + \beta Internet_{it} + \tau_t + \mu_i + \varepsilon_{it},$$

Here, τ_t is a time dummy that captures average changes to wage levels over time; μ_i is a location-specific fixed effect that gets differenced out in the estimation; and $Internet_{it}$ measures the extent of advanced Internet use by businesses in region i at time t.[6] We have assumed that ε_{it} is a normal i.i.d. variable.[7] We include two kinds of controls: First, X_{it} are controls for pre-existing factors that may affect wage growth such as income, population, and education. We set these to zero in 1995 so that they are not differenced out of the regression. This allows us to control for the degree to which the *levels* of the variables affect the *changes* in wages. Second, Z_{it} are controls for changes in the factors not directly related to income over time and for which we have data, including Internet use by local households (see Table 1b for the full list).[8]

Our hypothesis is that increases in local business use of advanced Internet will be associated with growth in local wages: A test of $\beta > 0$ against the null of $\beta = 0$.

We assume that the unobservable determinants of wages can be decomposed into an additively separable fixed component and a time-varying component that is constant across counties. We also assume that no systematic factors in ε_{it} are correlated with the unobserved differences in Y_{it}.

[6] As in Athey and Stern (2002) and Hubbard (2003), we treat the diffusion of a new technology as an exogenous factor that leads to a change in economic outcomes, and then examine the plausibility of the exogeneity assumption.
[7] Since we estimate the standard errors using heteroskedasticity-robust methods, the two-period framework is especially appealing. Stock and Watson (2008) show that the standard fixed-effects heteroskedasticity-robust variance matrix estimator is inconsistent if T is fixed and greater than 2.
[8] Taking the first difference yields a standard growth equation of changes in wages on levels and changes in the covariates. Since we treat the standard errors appropriately, this means that the coefficients and standard errors are exactly equivalent to those estimated using a growth equation.

One potential concern in this model is that unobservable changes to local firm or worker characteristics may be correlated with both wage growth and Internet use. We provide considerable suggestive evidence that, when combined, shows that advanced Internet use by firms is strongly correlated with local wage growth. First, as noted above, we include many controls for the initial (1990) conditions of the county to address omitted variables bias at the county level. Additionally, we include controls for changes in county characteristics such as population and age. We also show results with controls for changes in closely related margins of consumer and business IT investment—such as basic Internet investment, PCs per employee, and Internet use at home— changes that vary considerably across regions. If advanced Internet use is associated with wage growth but not with these other margins of IT investment, then omitted variable bias must be specific to advanced Internet. Further, we present instrumental variables regressions that use measures of local telecommunications infrastructure costs and the programming capabilities of related locations as instruments for local Internet investment. As we describe in greater detail below, changes in the values of these instruments will proxy for variance in the local costs of advanced Internet but are unlikely to be systematically correlated with local wage growth. Our results are robust to the use of these instruments, even when used in our difference-in-difference specification with the complete set of controls.

The Internet's unusual history gives us an additional test for the role of region-specific omitted variables: It enables us to employ a useful falsification test. Because the Internet was originally developed to facilitate research collaboration, it did not become clear until late 1994 that commercial implementation of the Internet could have a wide economic impact, and even at that point it was only apparent to Internet insiders. Most incumbent vendors and users in communications and computing markets, as well as IT-intensive users in the economy as a whole, were caught by surprise in late 1994 and early 1995 as the commercial Internet rapidly diffused. That (almost sudden) realization by so many firms contributed to a non-gradual response from both vendors and users. As a result, we should not see any affiliation between Internet investment and local economic activity before 1995.[9] If our assumptions of the orthogonality between the Internet and changes in local unobservables are violated, then our data will produce false positive associations between future use and growth in a period prior to 1995. If we find false positives, then it suggests that violations of our identification assumptions are artificially inducing $\beta > 0$. If not, then it boosts confidence in the exogeneity assumption embedded in (1).[10]

Step 2: Use of the Internet and convergence/divergence. The Internet's rapid diffusion pattern motivates examining divergence and convergence. The Internet achieved near ubiquity in the United States over a short period of time. By the end of the 1990s, over half the households in the United States had Internet access, as did over 90% of medium and large establishments. Since adoption was widespread, it is constructive to ask whether the changes in economic outcomes were too.

We test between the two starkest predictions of the impact of the Internet on wage growth. One view predicts that the Internet would improve growth prospects in many regions, and especially in low-density, economically isolated regions where basic Internet services were particularly valued for lowering communications costs. The contrasting view casts the Internet as

[9] We think it is safe to date the beginning of the investment boom in the Internet to 1996. Dating the rise of the commercial Internet is not an exact science, but a few well-known events provide useful benchmark. The first non-beta version of the Netscape browser became available in early 1995. The Netscape IPO occurred in August 1995 and it went spectacularly well for the founders and investors. Microsoft announced its change in strategy on December 7, 1995. Certainly no serious vendor in IT markets was ignoring the commercial Internet by December 1995. Neither was any large-scale investor in IT applications by this point, but major investment lags behind planning, and therefore changes slowly. As Forman (2005) discusses, while "adjustment costs" slowed down deployment, experimentation and learning about new uses also caused some lag in business investment.

[10] We provide further details on our exploration of endogeneity and omitted variables bias in the results section.

a technology that exacerbated existing inequalities in wages between urban/rural and frontier/mainstream users of IT.[11]

We begin by examining whether the benefits of Internet use are greater in areas with low or high income. Advanced Internet use may contribute to divergence due to an overheating effect on local wage growth from local economic prosperity. One example can illustrate. During this time period, establishments supplying parts for electronics goods and automobiles had a high propensity to invest in advanced Internet usage. Yet, an electronics parts supplier in San Jose, CA, faces a local labor market with a higher average income and education than an automobile parts supplier in Akron, OH, faces. Thus, we examine whether the rise in wages is greater for the higher income region than that for the lower income region, despite both regions investing equally in the Internet. While Internet use in the face of tight local labor demand will contribute to rising wages, similar investments in environments without tight local labor demand conditions will not.

Advanced Internet use could similarly lead to convergence, or a “Robin Hood” forecast for the economy-wide impact from the diffusion of the Internet. That is, as a communications technology with nearly instant universal availability, the Internet might lead to widespread productivity advances across many facets of the economy. We do not dismiss this view out of hand for two reasons. First, it received considerable popular attention at the time, though no systematic test has ever confirmed or refuted any related prediction that goes beyond anecdote. Second, systematic statistical study of the diffusion of basic Internet use to businesses is

[11] Cairncross (1997) was among the earliest to forecast that Internet technology would lead to significant changes in the spatial distribution of economic activity. By reducing the costs of economic isolation, Internet technology could shift economic transactions from locations in urban areas where average wages are relatively high (Glaeser and Mare 2001) to rural locations where wages are relatively lower. Forman, Goldfarb, and Greenstein (2005) investigate the diffusion of the Internet to business, but find evidence consistent with both views. Some evidence shows that basic Internet technology diffused first to rural and small urban areas, but there is also evidence that urban areas were leading users of the advanced Internet. For reviews of this literature, see Forman and Goldfarb (2006) and Greenstein and Prince (2007).

consistent with a premise behind this view: Other work suggests that basic Internet use had high marginal benefit to businesses in isolated and low-density areas, as well as low adaptation costs (Forman, Goldfarb, and Greenstein 2005).

In our first approach we estimate the simplest version of this hypothesis:

(2) $Log(Y_{it}) = \alpha_1 X_{it} + \alpha_2 Z_{it} + \beta Internet_{it} + \phi(Internet_{it} * HighIncome_i) + \tau_t + \mu_i + \varepsilon_{it}$,

Here, ϕ measures the difference for higher and lower income counties in the relationship between wages and advanced Internet. Divergence caused by the Internet will produce $\phi > 0$ and convergence $\phi < 0$ against a null of $\phi = 0$. Rejecting the null does not depend on β, but the estimate for β (combined with the estimate for ϕ) does shape the inference about the economic importance of the Internet for wages.

After showing the result for income, we examine several additional factors that influence local labor market conditions, namely, local skills, population, and industry composition. We focus on each factor for the following reasons. (1) Skills: Considerable evidence points toward complementarities between the use of advanced IT and a skilled (and/or educated) labor force, which implies there are higher wages due to these complementarities.[12] (2) Population: Larger cities have thicker labor markets for complementary services, specialized skills, or specialized vendors. The presence of complementary resources also increases the net returns to IT-based process innovations, thereby increasing the returns to productivity and growth from Internet adoption for enterprises in such locations.[13] (3) Industry composition, or IT-intensity: The clustering of IT-intensive enterprises in the same industry, accentuated by similarities in

[12] See e.g., Bresnahan, Brynjolfsson and Hitt (2002), Autor (2001), and Corali and Van Reenen (2001).

[13] A rich literature in urban economics has provided evidence on the presence of increasing returns and productivity benefits associated with location in an urban area (e.g., Rosenthal and Strange 2004). The reasons go back to Marshall's (1920) initial insights: Urban markets have thicker labor markets, greater input sharing, and more knowledge spillovers. The same reasoning applies to urban areas with an agglomeration of IT-using industries and why they might have advantages over areas without such IT-using industries.

investment behavior within industries (and therefore within regions), can lead to a simultaneous increase in demand for labor in agglomerated productive enterprises.[14] We examine the extreme position that all these factors matter, and we divide counties by the following factors in our approach: *education*, *population*, and *IT-intensity.* We use this extreme position because it provides a way to simplify the underlying five-way interaction (including the previously defined *income* and *Internet*) into a single interaction term. Those counties that score high on all factors are termed *HighAllFactors.* To investigate whether these other factors affect the divergence in incomes associated with Internet adoption, we estimate the following:

$$(3)\quad Log(Y_{it}) = \alpha_1 X_{it} + \alpha_2 Z_{it} + \beta Internet_{it} + \phi_1(Internet_{it} * HighIncome_i) + \phi_2(Internet_{it} * HighEducation_i) + \phi_3(Internet_{it} * HighPopulation_i) + \phi_4(Internet_{it} * HighITIntensity_i) + \phi_5(Internet_{it} * HighAllFactors_i) + \tau_t + \mu_i + \varepsilon_{it},$$

Here, ϕ_1 measures the difference between counties with high and low incomes, and ϕ_5 measures differences between counties with *HighAllFactors* and other counties. If $\phi_1 = 0$ but $\phi_5 > 0$, then divergence in incomes is isolated to regions with high income, education, population, and IT-intensity.

A finding of $\phi_1 = 0$ but $\phi_5 > 0$ also has implications for identification in the presence of potential omitted variables. If this result is a false positive caused by positive covariance between changes in ε_{it} and advanced Internet use, then it suggests this covariance is isolated only to locations that have a high income. While it is always possible that such unobservables may exist, we find it difficult to identify an economic mechanism that produces such unobservables in just a limited number of places.

[14] Furthermore, if IT-intensive enterprises earn greater productivity benefits from new IT use (Stiroh 2002), then industries with such enterprises will have the largest associated wage gains from Internet use. When these productivity benefits spill over to other enterprises (e.g., Greenstone, Hornbeck, and Moretti 2008), then locations with the "right" industries will experience broad-based wage gains that are greater than other equivalent locations that have invested in advanced Internet use.

3. Data

To measure how Internet investment influenced growth in wages, we combine several data sources about medium and large establishments and about US counties. Our IT data come from the Harte Hanks Market Intelligence Computer Intelligence Technology database (hereafter CI database). The database contains establishment- and firm-level data on characteristics, such as the number of employees, personal computers per employee, and use of Internet applications. Harte Hanks collects this information to resell as a tool for the marketing divisions of technology companies. Other economic researchers have used this source as a fruitful way to learn about enterprise IT use.[15] Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 2000.

Harte Hanks tracks over 300,000 establishments in the United States. Because we focus on advanced Internet applications, we exclude government, military, and nonprofit establishments. Our sample from the CI database contains commercial establishments with over 100 employees—in total 86,879 establishments.[16] While the sample only includes relatively large establishments, we do not view this as a problem because very few small establishments employed advanced Internet use in the late 1990s. The primary investors were large establishments making large-scale enterprise-wide investments worth tens of millions of dollars, and, in some large multi-establishment organizations, hundreds of millions of dollars per year.[17]

We focus on those facets of Internet technology that became available only after 1995 in a variety of different uses and applications. Our raw data include at least twenty different specific

[15] There is an increasingly long list of papers that have built on this data source and its predecessor from CI, including our own prior work.

[16] Establishments were surveyed at different times from June 1998 to December 2000. To control for increasing adoption rates over time, we reweight our adoption data by the ratio of average adoption rates in our sample between the month of the survey and the end of 2000.

[17] All our available evidence suggests that adoption monotonically increased in firm size, even controlling for many other determinants. Hence, our sample represents the vast majority of adopters, and this procedure leads to the best possible estimate of use in a location.

applications, from basic access to software for Internet-enabled ERP business applications software. Advanced Internet use involves frontier technologies and significant adaptation costs. We identify advanced Internet use from the presence of substantial investments in e-commerce or e-business applications.[18]

We stress that the investments we consider include several aspects of an enterprise's operations, not just the most visible downstream interactions with customers. These often involve upstream communication with suppliers and/or new methods for organizing production, procurement, and sales practices. We look for commitment to two or more of the following Internet-based applications: ERP, customer service, education, extranet, publications, purchasing, or technical support.[19] Most often, these technologies involve inter-establishment communication and substantial changes to business processes. We also experimented with a variety of alternative measures of business Internet use and our results are qualitatively similar under these alternative definitions.

To obtain location-level measures of the extent of advanced Internet use, we compute average rates of use for a location. Because the distribution of establishments over industries may be different in our sample from that of the population, we compare the number of establishments in our database to the number of establishments in the Census. We calculate the total number of establishments with more than 50 employees in the Census Bureau's 1999 County Business Patterns data and the number of establishments in our database for each two-digit North American Industry Classification System (NAICS) code in each location. We then

[18] In previous work this was labeled *enhancement* because it enhanced existing IT processes and contrasted with *participation*, that is, the use of basic Internet technologies, such as email or browsing (e.g. Forman, Goldfarb, and Greenstein 2002, 2005). In this paper, the contrasts are not the central focus, so we simply call it advanced Internet use, and, when necessary, we will contrast it with basic Internet use and PC use.

[19] Additional details on the construction of this variable can be found in Forman, Goldfarb, and Greenstein (2002).

calculate the total number in each location. Therefore, to account for over- and under-sampling in the Harte Hanks data, we weight a NAICS-location by

$$\frac{\text{Total \# of census establishments in location-NAICS}}{\text{Total \# of census establishments in location}} \times \frac{\text{Total \# of establishments in our data in location}}{\text{Total \# of establishments in our data in location-NAICS}}$$

We sum the weighted county-NAICS-level rates of use across NAICS within a county to obtain county-level estimates of the extent of advanced Internet use.

Prior research has shown that this measure has several attractive properties. For example, when aggregated to the industry level, this measure positively correlates with Bureau of Economic Analysis measures of industry-level differences in IT investment, as we would expect. Examples of industries that tend to have high advanced Internet use are Electronics Manufacturing, Automobile Manufacturing and Distribution, and Financial Services (Forman, Goldfarb, and Greenstein, 2002). Yet, it captures more than just the industry, varying considerably across establishments in different firms and regions. Among the biggest cities, areas with high use are those where a high fraction of local employment is in Internet-intensive (as well as IT-intensive) industries, such as the San Francisco Bay Area, Seattle, Denver, and Houston (Forman, Goldfarb, and Greenstein 2005). Thus, both the industry composition and the features of local areas shape use in the direction that economic intuition would forecast.

We obtain data from businesses on county-average weekly wages paid, total employment, and total establishments from the Quarterly Census of Employment and Wages, a cooperative program of the Bureau of Labor Statistics and the State Employment Security Agencies. Matching these data to our Internet data leaves a total of 2743 county observations. We drop 372 of the total 3115 counties because we lack data on Internet investment. We retain almost every urban and suburban county, as well as most rural ones. The vast majority of the dropped counties come from the lowest quartile of the population distribution.

To examine divergence, we use our previously defined set of factors as variables to interact with our measure of advanced Internet use. We focus on the roles of *income*, *education*, *population*, and *IT-intensity*. The data on *population, education,* and *income* come from the 1990 US Census. For IT-intensity, we measure the fraction of firms in IT-using and producing industries in the county as of 1995 from the US Census County Business Patterns data. National aggregate data shows that such industries have unusually high returns from investment in IT in the 1990s. We define these industries using the classification reported in Jorgenson, Ho, and Stiroh (2005, p. 93).

We combine these data with county-level information from a variety of sources. This information allows us to control for the underlying propensity of the counties to grow and innovate. First, the 1990 US Census provides county-level information on population, median income, net migration to the county (from 1995 data), and percentage of university graduates, high school graduates, African Americans, persons below the poverty line, and persons over age 65. We also use the 2000 US Census to control for changes in non-income-related factors: population, net migration to the county, and percentages of university graduates, high school graduates, persons over age 65, and African Americans. The 2000 Current Population Survey (CPS) Computer and Internet Use Supplement provide our data on the percentage of county households adopting the Internet at home. We use four measures of county-level propensity to innovate: (1) The number of students in Carnegie rank 1 research universities in 1990, (2) The fraction of students enrolled in engineering programs, (3) The percentage of the county's workforce in professional occupations in 1990 (which we will term employment for the remainder of the paper), and (4) The number of patents granted in the 1980s in that county, as found in the NBER patent database.[20]

Table 1a includes descriptive statistics on IT use and our measures of local wages, establishments, and employment. Table 1b includes a description of control variables.

4. Empirical Results

We initially establish a link between advanced Internet use and wages. We then show that no such link exists between advanced Internet use and employment or between advanced Internet use and the number of establishments. Next, we show that advanced Internet use differs from basic Internet and PC use. Finally we show that advanced Internet use contributes to divergence; in particular, it is only associated with wage growth in counties with high levels of income, education, population, and IT-intensity industry. We find no evidence that advanced Internet use has led to convergence.

4.1 Baseline Results

In Table 2, we show the baseline results across counties. Column 1 shows the correlation between advanced Internet use and wages at the county level without any controls. As suggested by the scatterplots in Figure 1, there is a strong positive correlation with wage growth. Column 2 includes county and year fixed effects, which alters the key estimates considerably, as we would expect. Column 3 provides what we view as our main specification: Namely, it includes controls for presample demographics (such as county income and population in 1990), changes in non-income demographics (such as population from 1990 to 2000 and net migration from 1990 to 2000), measures of presample innovativeness (such as patents granted in the 1980s) as well as

[20] Downes and Greenstein (2007) showed that the first three factors help explain availability of Internet infrastructure such as ISPs.

changes in home Internet adoption (effectively zero in 1995). Columns 4 though 7 provide a number of other specifications to show robustness. In every specification, we reject the null that advanced Internet use is not associated with wage growth.

In the main specification the coefficient on advanced Internet use is 0.0252. That is, regions with an average level of Internet use (8.9%) experienced 0.224% wage growth above what regions with no Internet use experienced. A one standard deviation increase in the use of the Internet is associated with 0.336% increase in wage growth. The data are skewed, so it is also interesting to look at the top decile, which is 0.216. That leads to a 0.32% increase in wage growth above the mean. Consistent with Figure 1a, this suggests that the Internet was not the primary force behind the 20% wage growth across all counties in our data from 1995 to 2000. Still, it is related to growth. As we show below, examining the average effect obscures variation across regions.

Even with such a small coefficient, omitted variable bias is an important concern in this analysis. As was previously described, we take several steps to address this concern. First, we have included several controls for the initial conditions of the county. For example, if counties with a more educated population are more likely to experience a wage increase from 1995 to 2000 and are more likely to adopt advanced Internet technologies, then we may observe a positive correlation between Internet technology and wage growth only because of this underlying correlation. For this reason, in Table 2 and subsequently, we control for several county-level characteristics including pre-sample population, racial composition, education, income, poverty, professional workers, enrollment in research universities, enrollment in engineering programs, age, and innovativeness as measured by patents granted in the county.

Second, we have included controls for changes in county characteristics: population, migration, racial composition, education, and age. We also control for Internet use at home as measured by the 2000 CPS Computer and Internet Use Supplement. Therefore, we are examining the question of whether advanced Internet use at work affects wages independent of individual-level propensities to use the Internet. Comparing Columns 3 and 5 of Table 2 shows that this control has no substantive impact on the qualitative results.[21]

Third, we examine the timing of the relationship between advanced Internet use and wage growth to look for false positives. That is, advanced Internet investment should only help firms in the latter half of the 1990s. Prior to 1995, the Internet had not diffused. Therefore, to explore whether our measure of Internet investment is simply capturing county-level propensity to grow, we must show that our measure of Internet investment is not correlated with wage growth prior to 1995.

Three versions of this falsification check are presented in Table 2, Columns 8 and 9, and in Figure 2. Columns 8 and 9 replicate Columns 3 and 4 but use logged wages in 1990 and 1994 as the dependent variables rather than logged wages in 1995 and 2000. Significance is lost, and the coefficient magnitudes fall substantially. Thus, counties with high levels of advanced Internet use in 2000 do not appear have grown faster prior to 1995. Figure 2 replicates the regression in Column 3 using data for all years from 1990 to 2000. The controls are the same as in Column 3 and the dependent variable is logged wages. We interact the measure of advanced Internet use (as of 2000) with year dummies from 1991 to 2000 in the same way and therefore get a measure of the association between advanced Internet use and wage growth over the period. Figure 2 clearly shows that advanced Internet use is not correlated with wage growth until 1997 (when the

[21] This data is only available for a subset of counties that can be identified in the CPS. For this reason, we also include a dummy variable that captures when this information is not available. We also show robustness to running the analysis on this subset of counties.

Internet began to diffuse widely). Between 1991 and 1996 the coefficient is statistically indistinguishable from zero in every year except 1992.

To further examine the robustness of our results in column 10 we present the results of instrumental variable regressions. We instrument for local advanced Internet adoption using several variables that should be correlated with local costs of Internet investment.[22] We include the number of local county connections to ARPANET and BITNET—two wide area data communications networks that were predecessors of the Internet—to proxy for local data communications infrastructure and expertise. We add two additional variables to proxy for local deployment costs: expense-related costs per telephone line in the wire centers of incumbent local-exchange telephone carriers (ILECs) and the year in which the local state capped the prices that ILECs could charge entrants.[23] By influencing local costs of Internet deployment—either through lower telecommunications costs or by increasing the supply of workers skilled in computer networking—these variables should be correlated with local county Internet adoption. However, they are unlikely to be correlated with unobservables influencing local wage growth. ARPANET and BITNET reflect historical decisions (from the 1970s and 1980s) about connectivity to Department of Defense or US university networks; telephone line costs are shaped by the costs of deploying telephone switches, which exhibit economies of scale but should be uncorrelated with wage growth given county fixed effects and other controls; and price caps reflect exogenous state government decisions. Finally, we look at the establishments in multi-establishment firms in the county. We measure the total number of programmers in other establishments and other counties, but in the same firm. We use the average as an instrument. Forman, Goldfarb, and Greenstein (2008) show these variables are correlated with Internet

[22] Here we provide a brief overview of each of these variables. Further details on sources, construction, and the first stage regressions are available in Appendix Tables 3 and 4.
[23] The wire center costs are derived from data used in Chen and Savage (2008).

adoption. They are also likely to uncorrelated with local wage growth; our programmers variable reflects the presence of IT skills in linked counties. In these regressions, we also include a control for multi-establishment firms, since the previous variable is defined only for such firms.[24] In total, we have five instruments for local advanced Internet adoption.

The results of these regressions remain qualitatively similar to our main specification in column 3 (a Hausman test retains the null that the coefficients in columns 3 and 10 are the same with a p-value of 0.999), and statistical significance is retained at the 10% level. While the point estimate on the coefficient of advanced Internet (0.3167) is higher than in some of our other models, this is because the effect is imprecisely measured: the standard error of 0.1705 is much higher than in other specifications.

Individually, these instruments are weak. Together, they become somewhat stronger because they capture different aspects of the variance in advanced Internet usage across counties. While on their own these results may not provide convincing evidence that omitted variable bias is not a concern, in combination with the numerous other robustness checks and controls they generate further confidence in our results.

In Table 3, we examine endogenous variables other than wages. In contrast to the results for wages, our findings show no consistent measurable relationship between advanced Internet use and employment or the number of establishments. The results in these columns are representative of our more general finding through many more analyses that neither total employment nor establishments are correlated with advanced Internet use in any systematic way. Our results suggest that the increase in wages is not related to a substantial negative effect on employment or number of establishments. Since we find no systematic pattern in the data for the

[24] Including multi-establishment firms as a control in all regressions makes no difference to the qualitative results.

relationship between Internet use and employment or number of establishments, for the rest of the study we focus on our wage results.

Finally, we ask whether advanced Internet use proxies for other kinds of IT. Thus, in Table 4, we examine how county-level wages change with advanced Internet investment, basic Internet investment, and PCs per employee. These are all measured using the Harte Hanks database and aggregated to the county level. Forman, Goldfarb, and Greenstein (2005) use the same measure of basic Internet investment and found it to be widely adopted by 2000. The measure of PCs per employee resembles that found in Beaudry, Doms, and Lewis (2006).

The results suggest that advanced Internet use is distinct from other measures of IT. While PCs per employee are positively correlated with wage growth, this relationship is no longer significant once the controls are included. Furthermore, including PCs per employee and basic Internet use as controls does not substantially change the marginal relationship between advanced Internet use and wages. This table suggests that advanced Internet investment is not simply a surrogate measure of IT intensity but that the relationship between wage growth and advanced Internet use is related to advanced Internet technology in particular. Indeed, since the correlation between advanced Internet use and PCs per employee at the county level is 0.20 and the correlation between basic and advanced Internet use is 0.18, the table also suggests that wages in some areas could especially diverge from others when both are high.

We have investigated the robustness of this finding and found no systematic relationship between basic Internet technologies (e.g., email and web browsing) and growth in employment, establishments, or wages. This is surprising because levels of participation were high across establishments and locations by 2000. Revealed preference suggests the benefits were high, especially for a technology with so little use only five years earlier. We speculate that our intuition about revealed preference applies to an inframarginal adopter: When the technology is almost universally adopted, the data may be identifying an uninteresting margin in the benefits of participation. In other words, with basic Internet technology there simply is too little variation in the independent variable.

Overall, our results suggest an association between advanced Internet use and wage growth in the late 1990s. These results are robust to a variety of analyses examining the importance of omitted variable bias. Furthermore, these results suggest that if our results are due to false positives arising from omitted variable bias, then we can circumscribe the features of these omitted variables. They must be correlated with advanced Internet use but not with other margins of IT investment, or other persistent regional features.

4.2 When Was Advanced Internet Use Related to Local Wage Growth?

In this section, we provide evidence that advanced Internet use led to divergence in wages across counties. In particular, advanced Internet use was primarily correlated with county-level wage increases in counties with high income, education, and population, and a large percentage of IT-intensive firms.

Based on equations (2) and (3), our regression results in Table 5 explore this pattern in several steps. Column 1 shows that advanced Internet use is only significantly associated with wage growth in counties in the top quartile of median income as of 1990. This means that counties in the top income quartile with high levels of advanced Internet use grew faster than counties in the top income quartile with low levels of advanced Internet use. In contrast, counties in other quartiles with high levels of advanced Internet use did not experience especially rapid wage growth. In short, advanced Internet use contributes to wage divergence.

Columns 2 through 4 show how the impact of advanced Internet use is influenced by variation in local education, IT-intensity, and population. Like Column 1, Column 2 shows that advanced Internet use is associated with wage growth only for counties in the top quartile of higher education (percentage university graduates as of 1990). The similarity of results is not surprising since 60% of the counties overlap. Column 3 shows that counties with over 100,000 people display a strong association between advanced Internet use and wage growth.

Column 4 examines counties in the top quartile in IT-intensity. In this specification, advanced Internet use is not significantly correlated with wage growth for high IT-intensity counties. Still, we include IT-intensity for three reasons. First and perhaps most importantly, IT-intensity has been emphasized in much of the previous literature linking IT to average productivity (e.g. Jorgenson, Ho, and Stiroh 2005). Second, the coefficient is positive and when added to the coefficient on the main effect in the first row, it is significantly different from zero with 95% confidence. Third, we tried several specifications and the coefficient was sometimes significantly positive and never negative. Therefore, while we are concerned about observational equivalence between IT-intensity and other observable regional attributes, we cannot reject a role for IT-intensity in the relationship between advanced Internet use and wage growth.

Column 5 shows that when we include all four measures of pre-Internet county strength (income, education, population, and IT-intensity), none end up significant. This may not be surprising given that there is considerable overlap between the measures: Each measure contains roughly 680 counties, of which 180 are in the top group in all measures. Column 6 shows that it is in these 180 counties that advanced Internet use is correlated with wage growth. Column 7 shows that it is the combination of more than one factor that drives the relationship between advanced Internet use and wage growth. Thus, there is something different about the 180 counties with high income, education, population, and IT-intensity.

These 180 *HighAllFactors* counties also had higher wage growth than the other 2563 counties in the sample: 29.2% vs. 20.5%. For the 180 counties with *HighAllFactors*, our results suggest that advanced Internet use is related to 8.2% (2.4 percentage points) of the total wage growth. For the other counties, advanced Internet use explains just 1.1% (0.23 percentage points) of overall wage growth.[25] Combined, this means that advanced Internet use explains one quarter of the 8.7 percentage point difference in wage growth between these 180 counties and the other 2563 counties in the sample.

The core results of Table 5 are robust to using continuous measures of income, education, population, and IT-intensity. Income loses significance and IT-intensity gains significance, but the significance and importance of the interaction term remains. Furthermore, adding all two-way interactions to Column 7 (i.e., high income and high education, high income and high population, etc.) does not change the core result: There is a large and significant coefficient for the 180 counties that were already doing well on all four measures.[26] Using the method that generated Figure 2, in Figure 3 we provide a falsification check of the results in Column 6. The results in the figure show that the relationship between advanced Internet and wages begins in the late 1990s.

These results further circumscribe concerns about omitted variables. There is no clear endogeneity story to explain the difference between regions with HighAllFactors and other regions. For example, if otherwise unmeasured regional prosperity causes both wages and

[25] More precisely, these calculations use the coefficient estimates in Table 5, Column 6—the average Internet use for the 180 counties, and the average Internet use in all other counties.

[26] These results are available in the Appendix. Adding the complete set of three-way interactions leads everything to be insignificant. We believe there is too much overlap in the measures to get significant estimates.

investment to rise, why is income growth only leading to Internet investment in places that were already doing well? Income growth is unrelated to Internet investment in other places even if they grew, and even if they were high adopters.

To further address concerns about omitted variable bias, column 8 presents the results of regressions that instrument for advanced Internet and its interaction with HighAllFactors. We interact each of our original instruments with an indicator for being located in one of the HighAllFactors counties. The resulting instruments are combined with the original set to form a total of 10 instruments for 2 potentially endogenous variables. The coefficient of advanced Internet is positive but is now insignificant. However, advanced Internet's interaction with HighAllFactors is positive, statistically significant ($p=0.05$), and of similar magnitude to the related estimate in column 6. Thus, our instrumental variables results provide supporting evidence for the finding that advanced Internet is associated with wage growth, particularly in the 180 counties that were already doing well in all four factors.

An analysis of outliers and "typical" cases among these 180 counties provides further details on the relationship between advanced Internet use and wage growth. Counties among the top 180 that have high advanced Internet use and wage growth (both at least one standard deviation above the mean) include San Mateo and Santa Clara CA (both in San Francisco-Oakland-San Jose MSA); Boulder and Arapahoe CO (Denver-Boulder-Greeley MSA); Fairfax, VA (Washington-Baltimore MSA); Travis TX (Austin-San Marcos MSA); and Washington OR (Portland-Salem MSA). Those with high wage growth (one standard deviation above mean) but relatively low advanced Internet use (below mean) include only Williamson TX (Austin-San Marcos MSA) and Hudson NJ (New York-Northern New Jersey-Long Island MSA). Those with high advanced Internet use (one standard deviation above mean) but relatively low wage growth (below mean) include Madison AL (Huntsville, AL MSA), Lake OH (Cleveland-Akron MSA), Kalamazoo MI (Kalamazoo-Battle creek MSA), and Middlesex CT (New London-Norwich MSA). In short, counties with high advanced Internet use and wage growth are often centers of IT production and use; counties with high advanced Internet use but low wage growth are often small cities where the labor markets are not very tight; and counties with high wage growth but low Internet use are relatively rare.

We also stress these results cannot arise due to the inordinate influence of canonical outliers. For example, we could remove Santa Clara or San Francisco from the data set and the results would not substantially change. In part, this should not be surprising; no single variable, not even advanced Internet use, could possibly explain the anomalous experience in Santa Clara in this time period (i.e., over 80% wage growth in five years). Mostly, however, the robustness of results to the exclusion of observations reflects the pattern in the scatterplot. There was a general experience found in a special set of urban counties outside Santa Clara, as well as inside of it. These 6% of US counties shared similar demographic and industrial traits prior to the Internet's diffusion and reacted to the diffusion of the Internet with similar economic experiences. In short, wage divergence in this time period contained an identifiable regional component.

4.3 Additional Implications of Advanced Internet Use

In this section we investigate whether the benefits of Internet investment persist over time for our 180 top counties and whether these benefits can spill over to adjacent locations. These analyses show that the effects of advanced Internet use were localized in time and space.

In Table 6, by repeating the regressions in Table 5 Columns 6 and 7, but using wage growth between 1999 and 2005 as the dependent variable, we examine whether early Internet-adopting counties continued to have higher wage growth once the diffusion of the Internet slowed. Our results show that the difference between the 180 counties that were already doing well in 1990 and the other counties was coincident with the one-time diffusion of the Internet. Advanced Internet usage is related to rapid growth from 1995 to 2000 in places that were already doing well in 1990. Then, these places maintained their new position in absolute terms. They did not grow faster, but the gains were not reversed either.

In Table 7, we explore whether the benefits of advanced Internet use in *HighAllFactors* counties can spill over to adjacent locations. We examine this question because local labor markets may extend beyond county boundaries. This is particularly likely in metropolitan areas, where workers frequently commute between counties. To investigate this possibility, we reran the regressions in Columns 6 and 7 of Table 5 by adding a new variable that is equal to one when the county is located in an MSA with a *HighAllFactors* county but is not itself one of these counties. The coefficient on this new variable is positive but insignificant and economically small, suggesting that any spillover benefits to being located near a *HighAllFactors* county are, at best, small.

4.4 Open Issues about Biased Technical Change

By introducing regional variation into our discussion of the economic impact of the Internet, we raise questions about local variation in the productivity benefits of IT use both in and out of IT-intensive industries in particular regions. We also raise questions at a variety of levels about local variation in the links between IT use and worker skills and education.

Our findings stress the results for average wage growth, but motivate further research on the mechanisms at work. We do not fully connect our results to the literature on biased technical change, largely due to data limitations. For example, we cannot examine whether wage gains are greatest for high- or low-skilled occupations within a county, nor can we examine how Internet use changes the wage distribution within a location because consistent wage data at the local level for programmers were not kept earlier than 1999. Hence, we have been unable to link programmer and nonprogrammer wages to their changes before and after the deployment of the Internet.

5. Discussion

In this study, we find evidence of an association between use of advanced Internet technology and local wage growth. Furthermore, we find that advanced Internet use is associated with divergence in wages: we only observe a relationship between wage growth and Internet in locations in the top quartile of income. Probing this relationship further, we find that wage gains associated with advanced Internet adoption were isolated to relatively populated locations in which IT production and use were concentrated, and where income and skills were high. This appears to have led to a one-time relative gain in wages for these locations. We also find little evidence that use of advanced Internet technologies was associated with growth in either employment or establishments.

Thus, despite recent evidence that Internet use may lower the costs of geographically isolated economic activity, there is no evidence in our data that advanced Internet use contributed to convergence in wages. In particular, our results suggest the existence of a considerable divide in the benefits of advanced Internet use across urban and rural areas; however, they do not support the use of subsidies to build infrastructure to lower that gap.

Because our work suggests that the returns to IT use may be higher when several factors appear together, we believe the debate about the economic impact of IT must change to focus attention on regional variation. Our work also points to the key role the Internet played in recent experience. That suggests the impact of its diffusion should be treated as a factor quite distinct from other aspects of computing, such as the impact of the PC.

Perhaps more speculatively, we sympathize with movements away from theories of growth in this time period that emphasize agglomeration of production in technology-intensive industries. However, our results also suggest that efforts to subsidize rural Internet development would have little impact. We find little evidence that the Internet has much impact in rural areas. This finding runs counter to the motivations for a wide array of policies encouraging Internet business use outside of urban areas, such as policies to subsidize rural broadband development.

Considerable complementary evidence would be needed to overcome warranted caution about drawing too much from one exercise; however, our results raise many provocative questions in directions that prior research has not explored. We hope this inspires other explorations into understanding the underlying economic mechanisms.

References

Aral, Sinan, Lynn Wu, and Vincenzo Morabito. 2007. Regional Economic Context and the Value of Firm-Level IT Investments. Working paper, NYU.

Arora, Ashish, V.S. Arunachalam, Jai Asundi, and Ronald Fernandes. 2001. The Indian software services industry. *Research Policy* 30(8), 1267-1287.

——— and Alfonso Gambardella. 2005. *From Underdogs to Tigers: The Rise and Growth of the Software Industry in Brazil, China, Ireland, and Israel.* Oxford: Oxford University Press.

Athey, S., and S. Stern. 2002. The Impact of Information Technology on Emergency Health Care Outcomes. *RAND Journal of Economics* 33(3): 399–432.

Autor, David., 2001. Wiring the Labor Market. *Journal of Economic Perspectives* 15(1): 25-40.

———, Lawrence F. Katz, and Alan B. Kreuger, 1998. Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics* 113 (4), 1169–214.

———, Frank Levy, and Richard J. Murnane. 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics* 118(4), 1279–334.

Barro, R. J. and X. Sala-i-Martin. 1991 Convergence across States and Regions. *Brookings Papers on Economic Activity* 1: 107-158.

Beaudry, Paul, Mark Doms, and Ethan Lewis. 2006. Endogenous Skill Bias in Technology Adoption: City-Level Evidence from the IT Revolution. Federal Reserve Bank of San Francisco. Working Paper #06-24.

Bloom, Nick, Raffaella Sadun, and John Van Reenen. 2007. Americans Do IT Better: US Multinationals and the Productivity Miracle. CEP Discussion Paper No. 788.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt. 2002. Information Technology, Work Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics* 117: 339–76.

Brynjolfsson, Erik and Lorin Hitt. 2003. Computing Productivity: Firm-Level Evidence. *Review of Economics and Statistics* 85(4): 793-808.

Cairncross, F. 1997. *The Death of Distance*. Cambridge, MA: Harvard University Press.

Chen, Yongmin and Scott J. Savage. 2008. The Effects of Competition on the Price for Cable Modem Internet Access. Working Paper, University of Colorado at Boulder.

Corali, Eva, and John Van Reenen. 2001. Skill Biased Organizational Change? Evidence from British and French Establishments. *Quarterly Journal of Economics,* 114(4), 1449–92.

Downes, Tom and Shane Greenstein. 2007. Understanding Why Universal Service Obligations May be Unnecessary: The Private Development of Local Internet Access Markets. *Journal of Urban Economics* 62(1): 2–26.

Forman, Chris, 2005. The Corporate Digital Divide: Determinants of Internet Adoption. *Management Science,* 51(4): 641-54.

———, Anindya Ghose, and Avi Goldfarb. 2009. Competition between Local and Electronic Markets: How the benefit of buying online depends on where you live. *Management Science* 54(1): 47-57.

——— and Avi Goldfarb. 2006. Diffusion of Information and Communication Technologies to Businesses. In *Handbook of Information Systems, Volume 1: Economics and Information Systems*, ed. Terrence Hendershott. 1–52. Amsterdam: Elsevier.

———, ———, and Shane Greenstein. 2002. Digital Dispersion: An Industrial and Geographic Census of Commercial Internet Use. NBER Working Paper #9287.

———, ———, and———. 2005. How Did Location Affect the Adoption of the Commercial Internet? Global Village vs. Urban Density. *Journal of Urban Economics* 58(3): 389–420.

———, ———, and———. 2008. Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources? *Journal of Economics and Management Strategy* 17(2): 295–316.

Friedman, Thomas. 2005. *The World is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus, and Giroux.

Gerschenkron, Alexander. 1962. *Economic Backwardness in Historical Perspective.* Cambridge, MA: Belknap Press of Harvard University Press.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer. 1992. Growth in Cities. *Journal of Political Economy* 100: 1126–52.

——— and David C. Mare. 2001. Cities and Skills. *Journal of Labor Economics* 19(2), 316–42.

——— and Giacomo A. M. Ponzetto. 2007. Did the Death of Distance Hurt Detroit and Help New York? NBER Working Paper 13710.

Greenstein, Shane and Jeff Prince, 2007. The Diffusion of the Internet and the Geography of the Digital Divide. In *Oxford Handbook on ICTs*, eds. Robin Mansell, Danny Quah, and Roger Silverstone, pp. 168-95. Oxford: Oxford University Press.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2008. Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants. NBER Working Paper 13833.

Higgins, Matthew J., Daniel Levy, and Andrew T. Young. 2006. Growth and Convergence across the United States: Evidence from County-Level Data. *Review of Economics and Statistics* 88(4): 671–81.

Hubbard, Thomas N. 2003. Information, Decisions, and Productivity: On Board Computers and Capacity Utilization in Trucking. *American Economic Review* 93(4): 1328–53.

Jorgenson, Dale, Mun S. Ho, and Kevin Stiroh. 2005. *Productivity Volume 3: Information Technology and the American Growth Resurgence*. Cambridge, MA: MIT Press. .

Kolko, Jed. 1999. Can I Get Some Service Here? Information Technology, Service Industries, and the Future of Cities. Working Paper, Harvard University.

———. 2002. Silicon Mountains, Silicon Molehills: Geographic concentration and convergence of Internet industries in the US. *Information Economics and Policy* 14(2): 211–32.

Magrini, Stefano. 2004. Regional (Di)Convergence. In *Handbook of Regional and Urban Economics,* Volume 4, eds. J. Vernon Henderson and Jacques-Francois Thisse, 2243–92. Amsterdam: Elsevier North-Holland. .

Marshall, Alfred. 1920. *Principles of Economics.* 8th edition. New York: Porcupine Press.

Moss, Mitchell, and Anthony Townsend, 1997. Tracking the Net: Using domain names to measure the growth of the Internet in US cities. *Journal of Urban Technology* 4(3): 47–60.

OECD. 2006. *OECD Information Technology Outlook 2006.* Paris: OECD.

Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the Nature and Sources of Agglomeration Economies. In *Handbook of Regional and Urban Economics,* Volume 4, eds. J. Vernon Henderson and Jacques-Francois Thisse, 2119–72. Amsterdam: Elsevier North-Holland.

Stevenson, Betsey. 2006. The Impact of the Internet on Worker Flows. Working Paper, University of Pennsylvania.

Stiroh, Kevin J. 2002. Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say? *American Economic Review* 92(5): 1559-76.

Stock, James H. and Mark W. Watson. 2008. Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression. *Econometrica* 76(1), 155–74.

Wellman, B. 2001. Computer networks as social networks. *Science* 29, 2031–34.

Xiang, Chong, 2005, New Goods and the Relative Demand for Skilled Labor. *Review of Economics and Statistics,* 87(2): 285–98.

Table 1a: Descriptive statistics for dependent variables, IT measures, and instruments (for 2000)

Variable	Mean	Std. Dev.	Minimum	Maximum	Number of observations
Log(average weekly wage)	6.153	0.2189	5.4931	7.333	2743
Log(employment)	9.190	1.4695	4.3175	15.08	2743
Log(establishments)	6.699	1.3143	2.7081	12.59	2743
Advanced Internet	0.0890	0.1332	0	1	2743
Basic Internet	0.7869	0.4499	0	1	2743
PCs per employee	0.2253	0.1719	0	1.937	2743
Average number of programmers in other establishments in the same firm	47.32	70.09	0	1137.6	2743
ARPANET connections	0.0215	0.2383	0	7	2743
BITNET connections	0.1597	0.6912	0	15	2743
Year when state adopted a telecom price cap or freeze	1995.0	3.08	1987	1999	2743
Average cost per phone line by state	24.06	3.92	14.92	36.42	2743

Table 1b: Description of control variables

Variable	Definition	Source	Mean
Internet use at home	Percentage of households with Internet at home (2000)	Current Population Survey (CPS) Internet Use Supplement (Census)	0.444
No Internet use at home data for county	Dummy indicating no data on home Internet use	Current Population Survey (CPS) Internet Use Supplement (Census)	0.9213
Total Population	Total Population as of Decennial Census (1990)	Census	89173
% African American	% Population African American as of Decennial Census (1990)	Census	0.0908
% University Graduates	% Population University Graduates as of Decennial Census (1990)	Census	0.1379
% High School Graduates	% Population High School Graduates as of Decennial Census (1990)	Census	0.6996
% Below Poverty Line	% Population Below Poverty Line as of Decennial Census (1990)	Census	0.1622
Median Household Income	Median County Household Income as of Decennial Census (1990)	Census	24493
# enrolled in Carnegie rank 1 research university	Per capita number of Students enrolled in local PhD-granting institutions	Downes-Greenstein (2007)	0.0081
# in Engineering Program	Per capita number of Students enrolled in engineering programs at local Universities	Downes-Greenstein (2007)	0.0010
# Patents Granted in the Country in the 1980s	Total number of patents from inventors located in county, 1980-1989	USPTO	155.7
% professional	% of County's Workforce Employed in Professional Occupations	Census	0.3258
Net Migration	Net migration to county	Census	123.5
% Population over Age 65	% of County Population over 65 as of Decennial Census	Census	0.1452

Table 2: Wages increase with Internet use

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No fixed effects	County and year fixed effects	Main specification	MSAs only	No controls for home Internet use	Only places where have home Internet use data	Years 1991-2001	Falsification test: Years 1990-94	Falsification test: Years 1990-94 MSAs only	Instrumental variables[a]
Advanced Internet	0.5215	0.0370	0.0252	0.0672	0.0253	0.2393	0.0257	0.0105	0.0165	0.3167
	(0.0481)**	(0.0132)**	(0.0128)*	(0.0364)+	(0.0128)*	(0.1116)*	(0.0141)+	(0.0098)	(0.0341)	(0.1705)+
Observations	5486	5486	5486	1686	5486	432	5486	5488	1686	5486
(Within) R^2	0.05	0.85	0.86	0.92	0.86	0.95	0.94	0.83	0.90	0.83
Fixed effects	No	County	County	County	County	County	County	County	County	County
Controls	None	Year	All	All	All except home Internet use	All	All	All	All	All

Dependent variable is logged wages. Unless otherwise stated, years are 1995 and 2000. Controls include year dummy, population in 1990, median income in 1990, percentage African Americans in 1990, percentage university graduates in 1990, percentage high school graduates in 1990, percentage below poverty line in 1990, per capita number of people attending Carnegie Type 1 schools in 1990, net migration into the county in 1995, number of patents granted to inventors located in the county in the 1980s, per capita number of students in engineering program in 1990, fraction professional in 1995, percentage of persons over age 65 in 1990, change in total population between 1990 and 2000, change in percentage of university graduates between 1990 and 2000, change in percentage of high school graduates between 1990 and 2000, change in percentage of African American between 1990 and 2000, change in net migration into the county between 1995 and 2000, and change in percentage of persons over age 65 between 1990 and 2000. Heteroskedasticity-robust standard errors in parentheses.

[a]Instruments are ARPANET nodes, BITNET nodes, costs per telephone line by county, year when state adopted a price cap or freeze, and the number of programmers residing in other establishment locations within the same firm.

+ significant at 10%; * significant at 5%; ** significant at 1%

Table 3: Employment and Establishments show no clear pattern of correlation with Internet use

Dependent Variable→	*EMPLOYMENT*			*NUMBER OF ESTABLISHMENTS*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	No fixed effects	County and year fixed effects	Main specification with several further controls	No fixed effects	County and year fixed effects	Main specification with several further controls
Advanced Internet	1.2483	-0.0023	-0.0181	1.1220	-0.0026	-0.0031
	(0.2573)**	(0.0201)	(0.0173)	(0.2210)**	(0.0147)	(0.0135)
Observations	5486	5486	5486	5486	5486	5486
(Within) R^2	0.01	0.27	0.44	0.01	0.42	0.58
Fixed effects	None	County	County	None	County	County
Other controls	None	Year	All	None	Year	All

Controls are the same as in Table 2. Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 4: Is Advanced Internet different from other measures of IT use?

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	No Fixed Effects or controls	Compare all three measures of IT use	Compare Advanced Internet and Basic Internet	Compare Advanced Internet and PCs per employee	PCs per employee only	Basic Internet only
Advanced Internet	0.0277	0.0232	0.0229	0.0244		
	(0.0413)	(0.0136)+	(0.0134)+	(0.0133)+		
Basic Internet	0.5447	0.0127	0.0119			0.0153
	(0.0624)**	(0.0108)	(0.0103)			(0.0097)
PCs per employee	0.0702	-0.0014		0.0022	0.0057	
	(0.0185)**	(0.0078)		(0.0074)	(0.0071)	
Observations	5486	5486	5486	5486	5486	5486
(Within) R^2	0.23	0.86	0.86	0.86	0.86	0.86
Fixed effects	None	County	County	County	County	County
Other controls	None	All	All	All	All	All

Dependent variable is logged wages. Controls are the same as in Table 2. Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 5: Effect primarily occurs in places that are already high income, education, IT-intensity, *AND* population

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*[a]
Advanced Internet	0.0150	0.0099	0.0213	0.0191	0.0007	0.0225	0.0029	0.1979
	(0.0138)	(0.0127)	(0.0130)	(0.0160)	(0.0150)	(0.0130)+	(0.0152)	(0.1622)
Advanced Internet and	0.0891				0.0430		0.0378	
High income county	(0.0367)*				(0.0499)		(0.0502)	
Advanced Internet and		0.1068			0.0824		0.0796	
High education county		(0.0447)*			(0.0557)		(0.0557)	
Advanced Internet and			0.1903		0.0927		0.0298	
High population county			(0.0680)**		(0.0756)		(0.0774)	
Advanced Internet and				0.0205	0.0198		0.0155	
High IT-intensity county				(0.0232)	(0.0238)		(0.0241)	
Advanced Internet and High income, education,						0.1785	0.1232	0.1767
IT-intensity, and population county						(0.0530)**	(0.0582)*	(0.0736)*
Observations	5486	5486	5486	5486	5486	5486	5486	5486
(Within) R^2	0.86	0.86	0.86	0.86	0.87	0.86	0.87	0.86
Fixed effects	County	County	County	County	County	County	County	County
Other controls	All	All	All	All	All	All	All	All

Dependent variable is logged wages. Controls are the same as in Table 2. Heteroskedasticity-robust standard errors in parentheses.

[a]Column 8 displays instrumental variables results using the same instruments as table 2 column 5 combined with their interaction with a HighAllFactors indicator.

+ significant at 10%; * significant at 5%; ** significant at 1%.

Table 6: Wage growth from 1999 to 2005 is not related to early use of advanced Internet

	(1)	*(2)*
Advanced Internet	-0.0081	-0.0017
	(0.0136)	(0.0156)
Advanced Internet and		-0.0202
High income county		(0.0439)
Advanced Internet and		-0.0624
High education county		(0.0695)
Advanced Internet and		0.0757
High population county		(0.0791)
Advanced Internet and		0.0127
High IT-intensity county		(0.0273)
Advanced Internet and High income, education,	0.0003	-0.0040
IT-intensity, and population county	(0.0427)	(0.0471)
Observations	5486	5486
(Within) R^2	0.87	0.87
Fixed effects	County	County
Other controls	All	All

Dependent variable is logged wages. Years are 1999 and 2005. Controls are the same as in Table 2.
Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%.

Table 7: Benefits of early Internet use do not spill over to adjacent locations

	(1)	*(2)*
Advanced Internet and in same MSA as High income, Education, IT-intensity, and population county	0.0251 (0.0294)	0.0439 (0.0307)
Advanced Internet	-0.0018 (0.0290)	-0.0406 (0.0328)
Advanced Internet and High income county		0.0410 (0.0499)
Advanced Internet and High education county		0.0815 (0.0554)
Advanced Internet and High population county		0.0315 (0.0901)
Advanced Internet and High IT-intensity county		0.0169 (0.0239)
Advanced Internet and High income, education, IT-intensity, and population county	0.1967 (0.0539)**	0.1500 (0.0579)**
Observations	5486	5486
(Within) R^2	0.86	0.87
Fixed effects	County	County
Other controls	All	All

Dependent variable is logged wages. Years are 1995 and 2000. Controls are the same as in Table 2.
Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%.

Figure 1a: Advanced Internet Use and Wage Growth by County

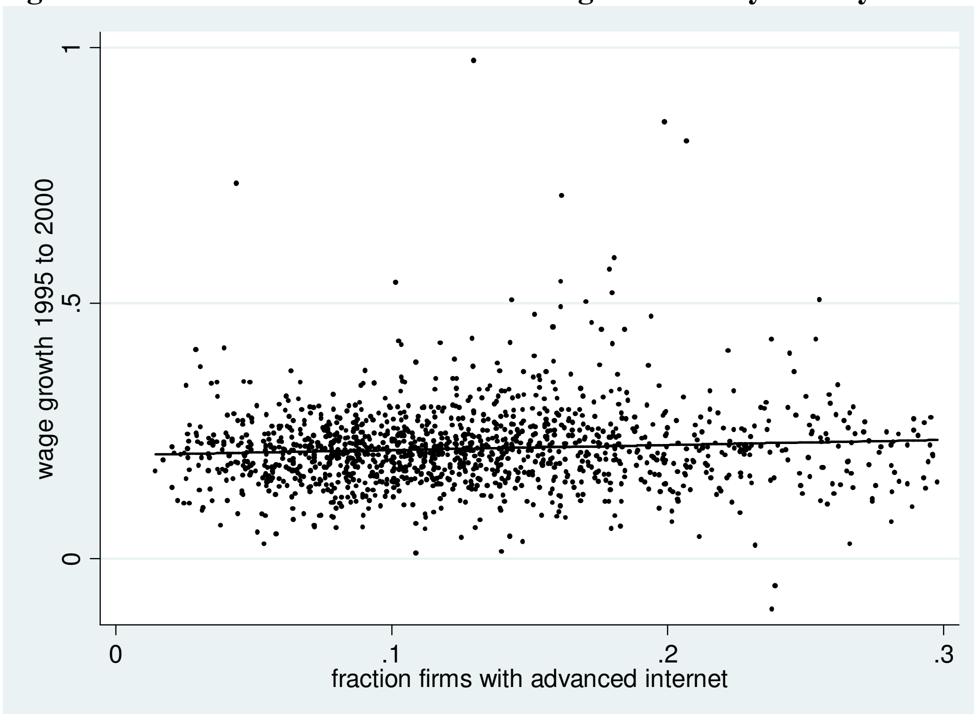


Figure 1b: Advanced Internet Use and Wage Growth by County Type

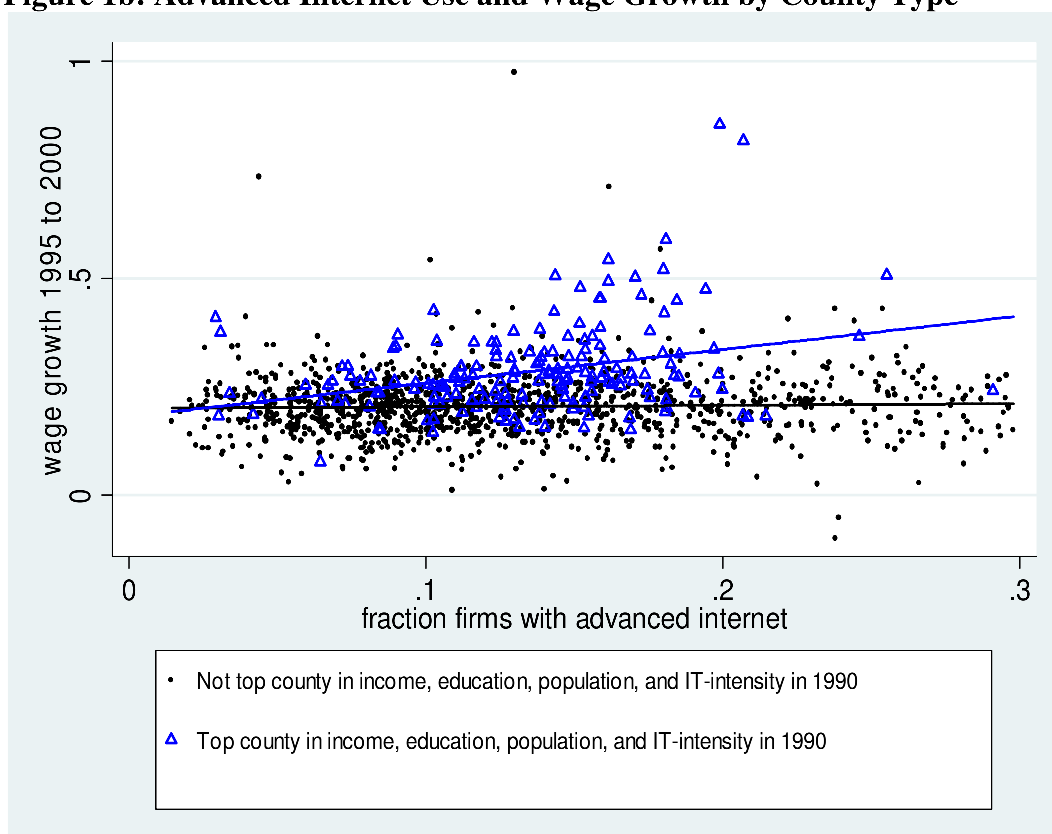
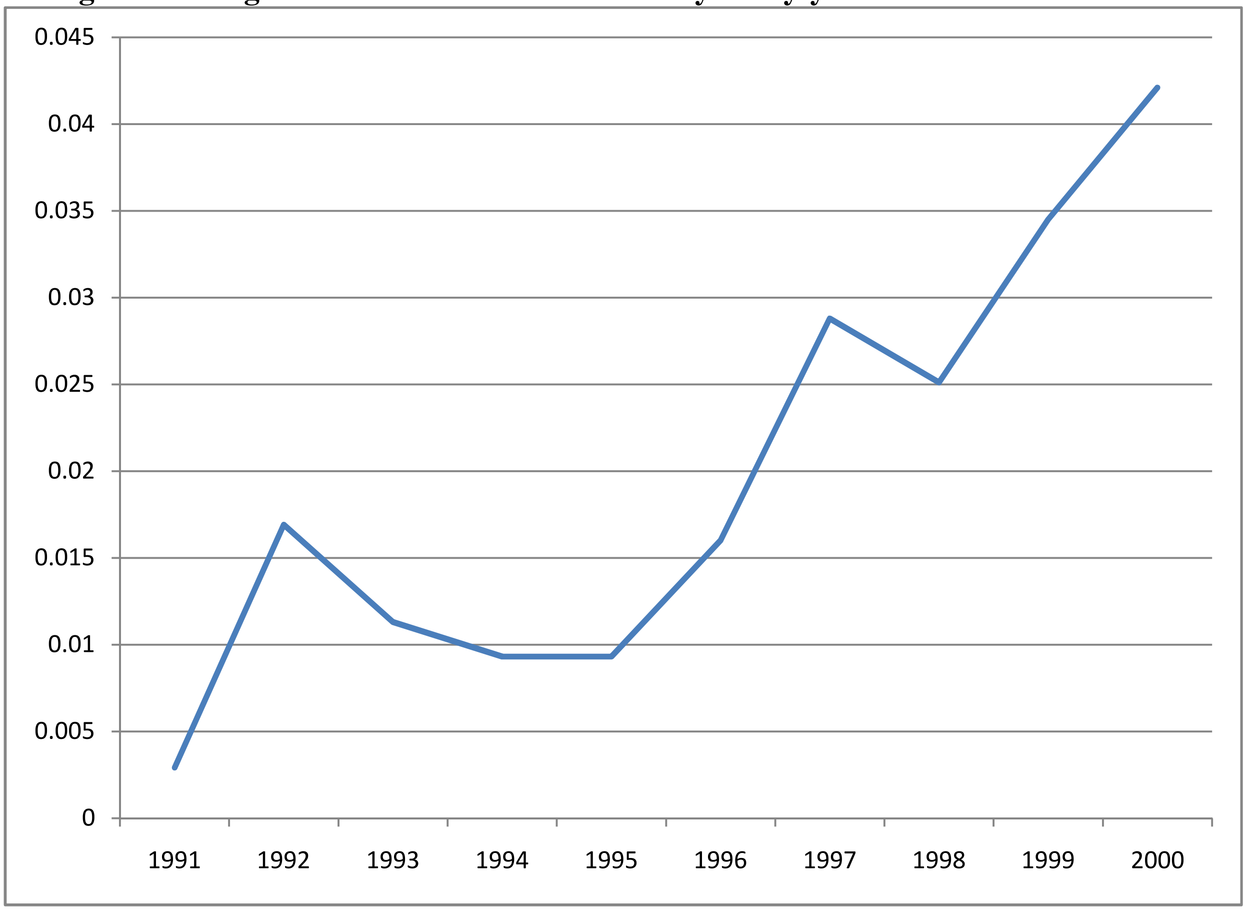
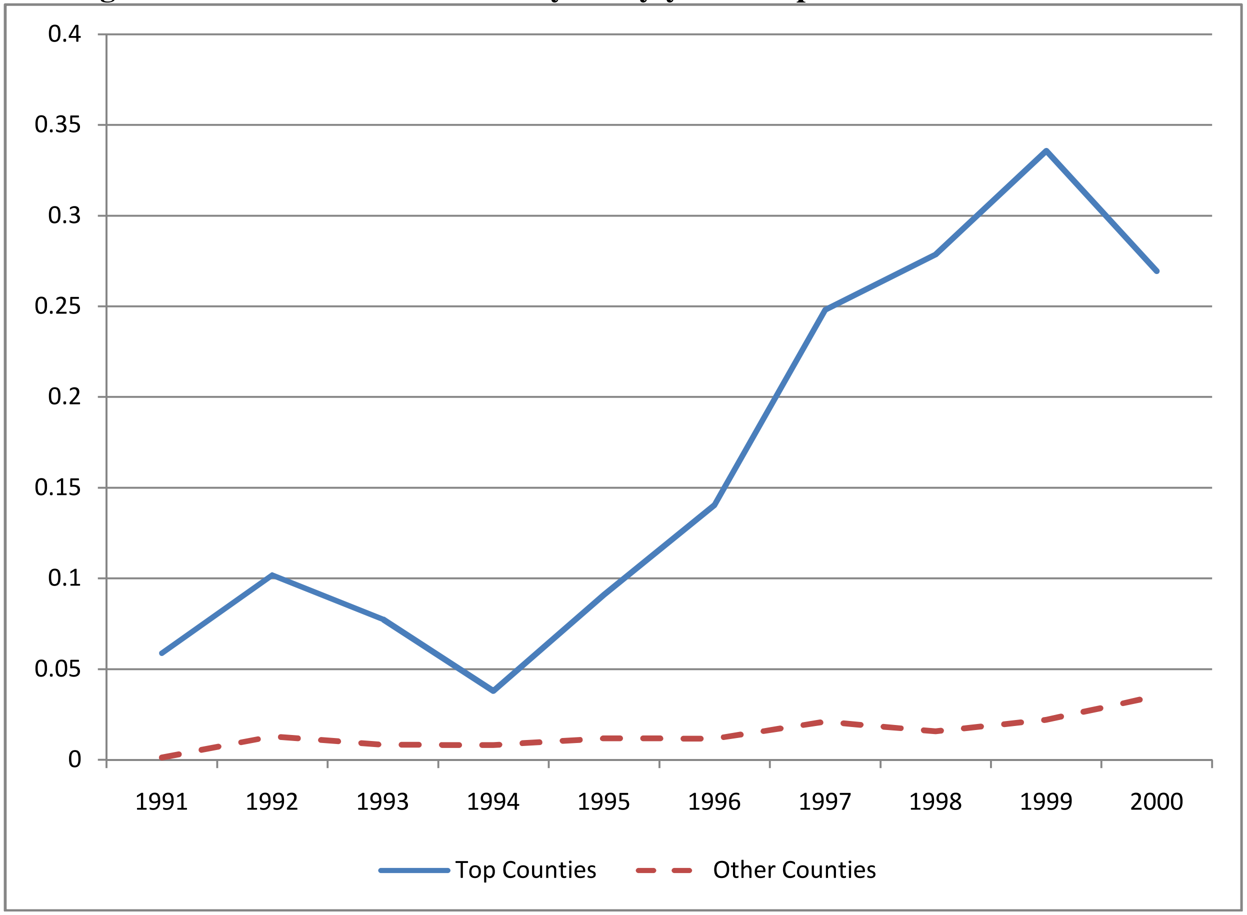


Figure 2: Marginal effect of advanced Internet year-by-year



This is based on the regression model is table 2 column 3 except that each year from 1990 to 2000 is included in the regression and a separate effect of advanced Internet (as of 2000) was estimated for each year using 1990 as the base. Controls are the same as in table 2.

Figure 3: Marginal effect of advanced Internet year-by-year in top counties



This is based on the regression model is table 5 column 6 except that each year from 1990 to 2000 is included in the regression and a separate effect of advanced Internet (as of 2000) and the interaction was estimated for each year. Controls are the same as in table 2.

ONLINE APPENDIX TO

THE INTERNET AND LOCAL WAGES: CONVERGENCE OR DIVERGENCE?

NOT FOR PUBLICATION

Online Appendix Table 1: Continuous measures for income, education, IT-intensive industry, and population

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Advanced Internet	-0.0317	-0.0421	0.0022	-0.0134	-0.0500	0.0213	-0.0226
	(0.0463)	(0.0361)	(0.0141)	(0.0301)	(0.0531)	(0.0129)+	(0.0532)
Advanced Internet x	2.65e-06				-2.21e-06		-2.56e-06
county-level income	(2.07e-06)				(3.65e-06)		(3.63e-06)
Advanced Internet x		0.5662			0.5497		0.5250
county-level education		(0.3025)+			(0.4691)		(0.4693)
Advanced Internet x			1.07e-06		9.30e-07		4.27e-07
county-level population			(2.25e-07)**		(2.36e-07)**		(2.41e-07)+
Advanced Internet x				0.1684	0.1650		0.1274
county-level IT-intensity				(0.1022)+	(0.1056)		(0.1063)
Advanced Internet x income x						1.38e-10	1.03e-10
education x population x IT-intensity						(3.36e-11)**	(3.48e-11)**
Observations	5486	5486	5486	5486	5486	5486	5486
(Within) R^2	0.86	0.86	0.86	0.87	0.87	0.87	0.87
Fixed effects	County	County	County	County	County	County	County
Other controls	All	All	All	All	All	All	All

Dependent variable is logged wages. Controls are the same as in table 2. Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Online Appendix Table 2: Further Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Includes two-way interactions	No weighting by time of survey		Dependent variable is wages, not logged		High population defined as top quartile of counties (> 63,657)	
Advanced Internet	-0.0035	0.0119	-0.0020	11.23	-1.61	0.0227	0.0035
	(0.0154)	(0.0101)	(0.0096)	(8.35)	(10.79)	(0.0130)+	(0.0153)
Advanced Internet and High income, education, IT-intensity, and population county	0.1706	0.1933	0.1278	216.92	154.30	0.1609	0.1211
	(0.0990)+	(0.0576)**	(0.0630)*	(46.19)**	(47.12)**	(0.0521)**	(0.0549)*
Advanced Internet and High income county	0.0779		0.0595		23.24		0.0460
	(0.0501)		(0.0450)		(24.64)		(0.0505)
Advanced Internet and High education county	0.1145		0.0514		55.38		0.0837
	(0.0672)+		(0.0438)		(32.58)+		(0.0554)
Advanced Internet and High population county	-0.0974		0.0333		89.66		-0.0363
	(0.1026)		(0.0827)		(59.90)		(0.0439)
Advanced Internet and High IT-intensity county	0.0342		0.0269		5.13		0.0170
	(0.0233)		(0.0186)		(13.24)		(0.0240)
Advanced Internet and High IT-intensity and population county	0.0418						
	(0.0701)						
Advanced Internet and High education and IT-intensity county	-0.0936						
	(0.0574)						
Advanced Internet and High income and IT-intensity county	-0.0760						
	(0.0603)						
Advanced Internet and High income and population county	0.0637						
	(0.0717)						
Advanced Internet and High education and population county	0.1016						
	(0.0619)						
Advanced Internet and High income and education county	-0.0659						
	(0.0725)						
Observations	5486	5486	5486	5486	5486	5486	5486
(Within) R^2	0.87	0.86	0.87	0.70	0.71	0.86	0.87
Fixed effects	County	County	County	County	County	County	County
Other controls	All	All	All	All	All	All	All

Dependent variable is logged wages unless otherwise stated. Time periods are 1995 and 2000. Controls are the same as in table 2. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

Online Appendix Table 3: Instrumental Variables Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Instrument: Number of programmers residing in other establishment locations	Instrument: ARPANET connections	Instrument: BITNET connections	Instrument: Year when state adopted a telecom price cap or freeze	Instrument: Average cost per phone line by state	All five instruments combined	All five instruments combined
Advanced Internet	0.1592	3.193	2.653	1.105	3.533	0.3167	0.1979
	(0.1673)	(9.068)	(5.823)	(1.088)	(22.71)	(0.1705)+	(0.1622)
Advanced Internet and High All Factors							0.1767
							(0.0736)*
Observations	5486	5486	5486	5486	5486	5486	5486
Hausman test chi-sq	2.39	0.12	0.20	0.99	0.02	3.53	1.33
Hausman test p-value (H_0: IV=OLS)	1.00	1.00	1.00	1.00	1.00	0.999	1.00

Dependent variable is logged wages. Time periods are 1995 and 2000. Controls include county fixed effects, year dummy, population in 1990, median income in 1990, percentage African Americans in 1990, percentage university graduates in 1990, percentage high school graduates in 1990, percentage below poverty line in 1990, per capita number of people attending Carnegie Type 1 schools in 1990, net migration into the county in 1995, number of patents granted to inventors located in the county in the 1980s, per capita number of students in engineering program in 1990, fraction professional in 1995, percentage of persons over age 65 in 1990, change in total population between 1990 and 2000, change in percentage of university graduates between 1990 and 2000, change in percentage of high school graduates between 1990 and 2000, change in percentage of African American between 1990 and 2000, change in net migration into the county between 1995 and 2000, and change in percentage of persons over age 65 between 1990 and 2000. Columns 1, 6, and 7 also include controls for the fraction of establishments in multi-establishment firms. Standard errors in parentheses. Columns 6 and 7 are also shown in the main paper. + significant at 10%; * significant at 5%; ** significant at 1%

Instrument definitions and sources:

Number of programmers residing in other locations in the same multi-establishment firm: Based on the "Organizational Capabilities" variable in Forman, Goldfarb, and Greenstein (2008), this measures the total number of programmers at other establishments in other locations within the same firm, averaged across establishments in the county. Constructed from the Harte Hanks dataset.

ARPANET connections: Count of the number of ARPANET connections in the county. Compiled from Hobbes' Internet Timeline http://www.zakon.org/robert/internet/timeline/ accessed Dec. 2008) and the ARPANET map (http://som.csudh.edu/cis/lpress/history/arpamaps/ accessed Dec. 2008).

BITNET connections: Count of the number of BITNET connections in the county. Compiled from CyberGeography (http://www.cybergeography.org/atlas/bitnet_typology.txt accessed July 2005).

Year when state adopted a price cap or freeze: Year, in the wake of telecommunications deregulation, that the state froze (or capped) the prices incumbent carriers could change entrants (source: FCC).

Average cost per telephone line by county: The average estimated cost to a telecommunications provider of providing a phone line (source: FCC and Chen and Savage 2008).

Online Appendix Table 4: First Stage of Online Appendix Table 3 Instrumental Variables Analysis

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Number of programmers residing in other establishment locations	0.000136 (0.0000387)**					0.000138 (0.000039)**	0.000139 (0.000038)**
ARPANET connections		0.00473 (0.0133)				0.00328 (0.0137)	0.00348 (0.0137)
BITNET connections			0.00246 (0.00532)			0.00248 (0.00549)	0.00239 (0.00548)
Year when state adopted a telecom price cap or freeze				0.000963 (0.000857)		0.00102 (0.000856)	0.000945 (.000851)
Average cost per phone line by state					-0.000114 (0.000730)	-0.000210 (0.000730)	-0.000249 (0.000727)
Interacted with High All Factors							
Number of programmers residing in other establishment locations							0.000249 (0.000017)**
ARPANET connections							0.0101 (0.00116)**
BITNET connections							-0.0006114 (0.0004922)
Year when state adopted a telecom price cap or freeze							0.000655 (0.0000802)**
Average cost per phone line by state							0.00202 (0.000328)**
Observations	5486	5486	5486	5486	5486	5486	5486
(Within) R^2	0.33	0.32	0.32	0.32	0.32	0.33	0.025; 0.92

Instrumented variable is logged wages. Time periods are 1995 and 2000. Controls include county fixed effects, year dummy, population in 1990, median income in 1990, percentage African Americans in 1990, percentage university graduates in 1990, percentage high school graduates in 1990, percentage below poverty line in 1990, per capita number of people attending Carnegie Type 1 schools in 1990, net migration into the county in 1995, number of patents granted to inventors located in the county in the 1980s, per capita number of students in engineering program in 1990, fraction professional in 1995, percentage of persons over age 65 in 1990, change in total population between 1990 and 2000, change in percentage of university graduates between 1990 and 2000, change in percentage of high school graduates between 1990 and 2000, change in percentage of African American between 1990 and 2000, change in net migration into the county between 1995 and 2000, and change in percentage of persons over age 65 between 1990 and 2000. Columns 1, 6, and 7 also include controls for the fraction of establishments in multi-establishment firms. Standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%