NBER WORKING PAPER SERIES

THE INTERNET AND LOCAL WAGES: CONVERGENCE OR DIVERGENCE?

Chris Forman
Avi Goldfarb
Shane Greenstein

Working Paper 14750
http://www.nber.org/papers/w14750

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2009

We thank Mercedes Delgado and seminar participants at Arizona State University, Carnegie Mellon University, Hunter College, London School of Economics, Northwestern University, Temple University, University of Arizona, University of Colorado, the University of Toronto, University of Vermont, Washington 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. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Internet and Local Wages: Convergence or Divergence?
Chris Forman, Avi Goldfarb, and Shane Greenstein
NBER Working Paper No. 14750
February 2009, Revised July 2010
JEL No. O33,R11

ABSTRACT

How did the diffusion of the internet affect regional wage inequality? We examine the relationship between business use of advanced internet technology and local variation in US wage growth between 1995 and 2000. We find no evidence that the internet contributed to regional wage convergence. Advanced internet technology is associated with larger wage growth in places that were already well off. These are places with highly educated and large urban populations, and concentration of IT-intensive industry. Overall, advanced internet explains over half of the difference in wage growth between these counties and all others.

Chris Forman
Georgia Institute of Technology
College of Management
800 West Peachtree Street, NW
Atlanta
GA 30308
USA
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
The Elinor and Wendell Hobbs Professor
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2013
and NBER
greenstein@kellogg.northwestern.edu

1. Introduction

A growing body of evidence suggests that investment in information technology (IT) in the 1990s produced gains in US productivity and economic growth at the national, industry, and firm levels. Such findings raise questions concerning the distribution of income growth across all parts of the country. Little analysis addresses the links between the distribution of economic growth and the distribution of business investment in information technology. This study contributes new empirical evidence to this topic.

We analyze the relationship between internet investment and changes in regional wage inequality. Did the internet contribute to divergence or convergence in wages across US counties? That is, did the deployment of the internet help well-off counties get even wealthier or did it help the poorer counties catch up to the rich? Overall our evidence rejects convergence and suggests divergence. Specifically, business internet use is correlated with larger wage growth in well-off locations than elsewhere, and our analysis is suggestive of a causal relationship.

Our study contributes to an old debate about the relationship between economic growth, technological development, and cities. With few exceptions, that literature concluded that information technology (IT) is skill-biased and advanced IT leads to increased wage inequality across workers. A related literature has emphasized that better local infrastructure and thick labor markets for skilled workers generate large benefits from IT in urban areas, leading to increased inequality across locations.

The rise of the commercial internet in the 1990s reshaped the terms of debate. A common optimistic forecast predicted that the economic gains from the internet would not primarily accrue to skilled workers in urban locations. In this view information and *communications* technology would generate growth in low-density locations because easier electronic communication would provide poor isolated areas with access to suppliers and markets. While this view has received widespread attention—e.g. Cairncross's (1997) *The Death of Distance* and Friedman's (2005) *The World is Flat*— the hypothesis went untested.

We do not doubt the potential for the internet to lower the costs of engaging in economic activity in geographically isolated locations,[1] but that does not settle the open question about whether its diffusion contributed to convergence or divergence. Even in the presence of growth in all locations, convergence or divergence depends on where the internet caused more (and less) growth to occur, and whether pre-internet income levels correlated with the effect of the internet on growth.

Lack of local US data has prevented systematic testing of convergence or divergence. We present novel data measuring the impact of internet investment as of December 2000 and relate it to county-level wages paid by local businesses from 1995 to 2000. As in prior research (Forman, Goldfarb, and Greenstein 2005), we look beyond the diffusion of email and web browsing, focusing on the diffusion of advanced internet applications. Unlike basic internet applications, advanced internet enabled possible productivity advances due to lower costs of communicating with suppliers and customers over long distances. Also unlike basic internet, advanced internet necessarily required skilled labor to implement and operate.

We focus on the change from 1995 to 2000 because this was the period of most rapid growth in business internet connectivity. It was also the first period of business internet adoption. The internet was largely unknown at the beginning of 1995 (Bill Gates' memo "The Internet Tidal Wave" was written in May 1995), but many firms had undertaken large investments by 2000. This period of rapid early investment is particularly conducive to a test of the starkly different short run implications of the competing hypotheses: divergence leads to faster short-run wage growth in counties that were already well off; convergence leads to the opposite.

[1] Prior work (Forman, Goldfarb and Greenstein 2005) showed that basic internet use was disproportionately adopted by businesses in low-density areas. Furthermore, lower communication costs have enabled the delivery of a set of tradable services at distances from the point of final demand (Arora and Gambardella 2005; OECD 2006).

While almost all counties were growing in the boom of the late 1990s, we find that the internet helps explain why wages in already well-off counties grew faster. We develop this argument in steps. First, we find a statistically significant but economically small positive correlation between internet investment and wage growth across all counties. This positive correlation remains robust to numerous specifications and changes in controls. More interestingly, we find that that this relationship is more pronounced in the 163 counties that, as of 1990, had a population over 150,000 and were in the top quartile in income, education, *and* fraction of firms in IT-intensive industries. Overall, the internet explains more than half of the difference in wage growth between the 163 counties that were already doing well and all other counties.

We address the assumption that internet investment is exogenous. First, although we add many controls for factors known to shape investment decisions, the results do not change. Second, we instrument for advanced internet in several ways. One, the Bartik procedure, is familiar to the literature in labor economics. The others are tailored to our setting, taking advantage of features of the cost structure for internet and communications technology. Third, we directly address omitted variables bias by showing that the timing of divergence is strongly associated with the timing of the diffusion of the business internet. The strong association between internet adoption and growth for those 163 counties that were already doing well starts in 1996, after the diffusion of the internet.

A scatterplot of the raw data forecasts our core results. Figure 1a shows the relationship between advanced internet investment and local wage growth for all types of counties in the raw data. While the regression line is upward sloping (it is also significantly positive), advanced internet is clearly not an important explanation of wage growth. In contrast, Figure 1b compares the 163 counties that were already doing well with the other counties. For the 163 counties that were already doing well (i.e., counties with high income, population, education, and agglomeration of IT-intensive firms), advanced internet is strongly correlated with wage growth; for the other counties, there is no relationship between advanced internet and wage growth.[2] Of course, our analysis goes far beyond this scatter-plot, but the result continues to hold after a wide battery of corrections and tests.

Our study sits between several different literatures. While there is a large macroeconomic literature on the causes of regional convergence and divergence,[3] we are the first study to examine the internet's role. In doing this we follow in the spirit of research showing that IT-using industries, firms, and locations had exceptionally good economic performance in the 1990s.[4]

Our paper also complements the literature on skill-biased technical change.[5] Our findings show that internet technology follows the skill-biased pattern observed with previous generations of IT. However, our results suggest that rich human capital in the form of a highly educated labor force is insufficient for a location to realize wage gains from internet investments. For internet investments to generate wage gains, there also must be other factors that shape local labor markets, coincident with local population size, industry composition, and income. It is the combination of all of these factors that appears to shift labor demand sufficiently to generate wage gains.

---

[2] Figure 1 truncates the picture by removing 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.

[3] See Magrini (2004) for a recent survey on the causes of convergence/divergence in regional growth. Also see, e.g., Glaeser et al (1992), Barro and Sala-i-Martin (1991), and Higgins, Levy, and Young (2006). Related is Glaeser and Ponzetto (2007), who argue that low communication costs help rich, idea-producing areas more than poor, goods-producing areas. They do not empirically focus on IT, but show that an increase in the share of skilled occupations is associated with greater local wage growth.

[4] This holds whether performance is measured at the national (Jorgenson, Ho, and Stiroh 2005), city (Beaudry, Doms, and Lewis 2006; Kolko 2002), industry (Stiroh 2002), firm (Brynjolfsson and Hitt 2003), or establishment (Bloom, Sadun, and Van Reenen 2007) levels.

[5] There is an extensive literature on wage inequality and skilled-biased technical change (e.g., Katz and Autor 1999). Recent literature has begun to investigate how the demand for computing has affected wage inequality (e.g., Autor, Levy, and Murnane 2003; Autor, Katz, and Kearney 2006).

While our results suggest that it is the combination of these factors that leads to disproportionate wage growth in response to advanced internet investment, we do not have empirical evidence that fully links our findings to the literature on biased technical change through comparisons of skilled and unskilled wages, mainly due to data limitations. In particular, data from this period are not detailed enough to allow examination of whether wage gains are greatest for high- or low-skilled occupations within a local labor market, nor are there sufficient data to examine how internet use changes the wage distribution within a location.

Our findings also speak to the difference between the experience of IT-using and IT-producing industries in the 1990s that is emphasized by Jorgenson, Ho, and Stiroh (2005). In contrast to the perspective that focuses on agglomeration of IT-producing industries in a handful of locations such as Santa Clara and Boston, we stress the somewhat broader economic growth that arose from deploying, building, and using new technology in both IT-producing and IT-using industries.

Our results have important public policy implications. A wide array of policies subsidizing internet infrastructure have arisen since the diffusion of the internet. In contrast to the motivation frequently given for these subsidies, we find little economic impact from the internet on wages in low density areas. Furthermore, while our results show that the first wave of advanced internet technology investment primarily benefited areas with educated workers and IT-using industry, the most common policies for subsidizing infrastructure include little or no provision for developing the human capital required to employ advanced IT.

2. Motivation

In this section we present a simple framework focusing on shifts in labor demand to describe how the diffusion of the internet could have generated divergence in average wages across US counties. This contrasts with an alternative potential framework that advanced internet led to convergence by disproportionately raising the benefit of geographically isolated computing capital. We describe how the shift in labor demand due to advanced internet investment likely varied with the local presence of skilled labor, IT-using industry, and high population. We will focus on average wages for practical reasons, namely, the absence of detailed measures of skilled and unskilled wages in a county and their growth during the deployment of the internet. There is, however, data about the level of *average wages* for every county in the US and about changes to those averages during the time when the internet diffused.

Corresponding with the internet's history and our data, we focus on understanding the short run response in demand and wages and assume the local labor supply curve does not shift.[6] 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 1995 as the commercial internet privatized and began to acquire the advanced capabilities affiliated with the growing World Wide Web. This almost sudden realization by so many firms contributed to a non-gradual response in investment (Greenstein 2010), yielding the short run consequences on local economies that we examine in this paper (and, potentially, long run consequences after supply-side adjustments such as labor mobility and skill acquisition).

The introduction of internet technology lowered both the cost of communication and the cost of information processing (the latter by reducing the cost of shared computing resources such as databases on servers). If we let computing capital represent a bundle of information processing and communication capabilities,[7] we can interpret the availability and deployment of advanced internet for commercial purposes as a large one-time fall in the price of the computing capital. That enabled firms to explore development of a range of new applications (Forman and Goldfarb 2006). As shown in a number of existing models of skill-biased technical change (e.g.

[6] In our empirical analysis, we include controls for potential changes in labor supply such as migration.
[7] This is a common assumption. See, e.g. Garicano and Rossi-Hansberg (2006) and Bloom et al. (2009).

Autor, Katz, and Kearney 2006), a fall in the price of computing capital will shift the demand for skilled labor, thereby raising the price of skilled labor relative to unskilled labor. Consequently, places with high levels of investment in advanced internet are likely to have higher growth in wages for skilled workers. And, due to basic arithmetic, places with a high fraction of skilled workers were likely to experience particularly high growth in average wages in association with advanced internet investment.

We further posit that the rise in demand for skilled labor was *location-biased*. First, its rise will be linked to geographic variance in pre-existing installations of computing capital. Investment in advanced internet had two distinct effects. It raised the productivity of *all future* computing capital investments in a wide range of applications, and it raised the potential productivity of the *existing stock* of long-lived computing capital in the range of applications where inexpensive retrofitting investments could be made. In practice the demand for IT investments in retrofitting existing installations, and the associated demand for skilled labor, are linked closely to the location of existing installations.[8] Therefore, locations that already had large numbers of IT-intensive firms were particularly well-positioned to benefit from advanced internet, especially if there were skilled workers present to help implement the changes. In these places, we expect to see a larger shift in labor demand and a larger increase in wages.

In addition to accruing in places with a large existing stock of IT, we expect these benefits to accrue disproportionately in large cities, due to Marshallian externalities.[9]

[8] Other types of demand for IT investments may or may not be tied to existing installations, depending on the specific needs of the enterprise. While prior evidence of the analysis of demand for other types of information technology has suggested that future demand is frequently tied to existing installations (e.g., Forman and Goldfarb 2006), in motivating our empirical analysis it is sufficient to focus on the harder-to-dispute effects on demand for retrofitting existing installations.

[9] 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, particularly for skilled labor. We note that Jacobs-type urbanization effects would produce observationally equivalent predictions. In particular, Jacobs-type effects would cause the prices of IT to fall faster in large cities. This would not influence the shift in skilled labor for a given level of IT investment, but would influence the extent of IT investment across cities. It would lead to an association between advanced IT and high wage growth.

Historically, the presence of thicker labor markets, greater input sharing, and greater knowledge spillovers in cities lowered the costs of successful adoption of frontier information technologies (Henderson 2003; Forman, Goldfarb, and Greenstein 2008). Because cities with these features have also become home to skilled labor (long before the arrival of the business internet) and because these features facilitate the successful implementation of advanced internet investments, the relationship between advanced internet investment and average wages should be particularly strong in large cities, especially those with a high fraction of skilled workers and a large number of industries using prior generations of IT.

The prevalence of IT-intensive industries, high population, and educated labor are positively correlated with income across counties (perhaps because income is an alternative measure of skill). Hence, this discussion suggests that places that already had high levels of income were likely to gain disproportionately from advanced internet investment, leading to divergence.

While we expect demand for skilled labor to rise from the deployment of the internet, this framework suggests three plausible explanations why we might not observe large changes in wages in some places that did see some investment in advanced internet. First, if unskilled wages were static or fell, and they are a large proportion of the labor force, overall wages could remain unchanged. Second, an elastic supply of skilled labor (US unemployment rates for skilled labor are higher outside of cities) could lead to small wage gains, even in the face of large demand shifts. Third, competitive forces, such as international competition, could reduce the returns to investment in the short run, leading to advanced internet investment without wage gains.

This simple framework motivates our empirical analysis. We begin by examining whether advanced internet investment is correlated with local wage growth. Then, we examine three main predictions from this framework: (1) Advanced internet investment has a particularly large effect on wages in educated, high income, populated areas with a concentration of IT-intensive industries, (2) Advanced internet investment is associated with stronger employment growth in these well-off counties than in other counties, and (3) The effects of advanced internet investment on wages are strongest in places with tight labor markets, particularly for skilled workers.

3. Measuring the Localization of Growth

Our statistical approach proceeds in two broad steps. We first measure the average relationship between internet use and wage growth across all counties. Next, we examine whether advanced internet investment led to faster growth in areas that were already doing well.

Step 1, Advanced internet and local wage growth: We compare 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 local features that shaped labor markets and enterprises in 1995 had not changed by 2000. Our endogenous variable will be the log difference in wages between 1995 and 2000, yielding:

(1) $Log(Y_{i00})- Log(Y_{i95}) = \alpha X_i+\beta Internet_i+\varepsilon_i,$

Here, $Internet_i$ measures the extent of advanced internet investment by businesses in location i in 2000. We have assumed that ε_i is a normal i.i.d. variable. We include two kinds of controls in X_i: Controls for pre-existing initial conditions that may affect wage growth such as income, population, and education levels and 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 a complete list).

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 $\beta = 0$.

Step 2, Use of the internet and convergence/divergence: We test the predictions of our framework against the starkest alternative, that the internet would improve growth prospects in many locations, and especially in poor, economically isolated locations where internet services may be particularly valued for lowering communications costs.

In our first approach we test the simplest version of this hypothesis, namely, that the internet leads to divergence in local wages:

$$(2) \qquad Log(Y_{i00}) - Log(Y_{i95}) = \alpha X_i + \beta Internet_i + \phi(Internet_i \times HighIncome_i) + \varepsilon_i,$$

Here, ϕ measures the difference in the relationship between wages and advanced internet for higher and lower income counties. 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 our interpretation of the effect.

Focusing on income is suggestive, but our framework emphasizes how the presence of local county-level characteristics such as skills (measured by education), population, and IT-intensity shape the relationship between advanced internet investment and local wages. We focus in particular on the extreme position that locations with the combination of these factors and income will exhibit the strongest relationship between advanced internet and wage growth. We use this extreme position because it provides a way to simplify the underlying five-way interaction into a single variable. Those counties that score high on all factors are termed *HighAllFactors.* To investigate these comparative statics of our framework, we estimate the following:

$$\begin{aligned}(3) \qquad Log(Y_{i00}) - Log(Y_{i95}) = {} & \alpha_1 X_i + \beta Internet_i + \phi_1(Internet_i \times HighIncome_i) \\ & + \phi_2(Internet_i \times HighEducation_i) + \phi_3(Internet_i \times HighPopulation_i) \\ & + \phi_4(Internet_i \times HighITIntensity_i) + \phi_5(Internet_i \times HighAllFactors_i) + \varepsilon_i,\end{aligned}$$

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 locations 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_i and advanced internet investment, 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 a specific economic mechanism that produces such unobservables in just a limited number of places.

More generally, a potential concern in this framework 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 investment is strongly correlated with local wage growth. First, as noted above, we include many controls for the initial 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 distribution. 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 locations. If advanced internet investment is associated with wage growth controlling for these other margins of IT investment, then omitted variable bias must be specific to advanced internet.

Second, we present instrumental variables regressions that use measures of local telecommunications infrastructure costs, local industry, 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.

Finally, the internet's sudden deployment gives us an additional test for the role of location-specific omitted variables: It enables us to employ a useful falsification test. We should not see any affiliation between internet investment and the divergence in local economic activity before 1995.[10] 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 wage divergence in a period prior to 1995. If we find false positives, then it suggests that violations of our identification assumptions are artificially inducing a divergence between the counties that were already doing well and the other counties in the sample. If not, then it boosts confidence in our exogeneity assumptions.

4. Data

To measure how internet investment influenced growth in wages, we combine several data sources about medium and large establishments and about US counties.[11] Our IT data come from the Harte Hanks Market Intelligence Computer Intelligence Technology database (hereafter CI database). The database contains rich establishment- and firm-level data including 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.[12] Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 2000.

[10] Dating the rise of the commercial internet is not an exact science, but a few well-known events provide useful benchmark for understanding why investment began to boom in 1996 and not before. The first non-beta version of the Netscape browser became available in early 1995, followed by the firm's IPO in August 1995. Certainly no serious vendor in IT markets was ignoring the commercial internet by December 1995, after Microsoft's announcement of its change in strategy, neither was any large-scale investor in IT applications.

[11] This section contains an overview of our data. Further details—in particular on the construction of our measure of internet investment and of our controls—are available in the Data Appendix.

[12] 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.

Harte Hanks tracks over 300,000 establishments in the United States. We exclude government, military, and nonprofit establishments because the availability of advanced internet for these establishments and their relationship between adoption and labor demand is likely to be systematically different than for private establishments. For example, many military establishments had access to ARPANET as early as 1970. Furthermore, our framework emphasizing a link between the price of computing capital and wages is focused on for-profit enterprises. Our sample from the CI database contains commercial establishments with over 100 employees—in total 86,879 establishments.[13] While the sample only includes relatively large establishments, we do not view this as a problem because very few small establishments deployed advanced internet technology 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.[14]

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 applications, from basic access to software for internet-enabled ERP business applications software. Advanced internet involves frontier technologies and significant adaptation costs. We identify advanced internet from the presence of substantial investments in e-commerce or e-business applications.[15]

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

[13] 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.

[14] 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.

[15] 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, and, when necessary, we will contrast it with basic internet and personal computers.

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. 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 investment, 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 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 investment. We show robustness to alternative weighting schemes.

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 investment 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 county-level data about businesses on average weekly wages paid and total employment 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. Results are robust to using multiple imputation to deal with the missing data.

To examine divergence, we use our previously defined set of factors as variables to interact with our measure of advanced internet. 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 IT-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 the 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 provides 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, and (4) The number of patents granted in the 1980s in that county, as found in the NBER patent database.[16]

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

5. Empirical Results

We initially establish a link between advanced internet and wages, and show that advanced internet technology differs from basic internet and personal computers. We then present the main result that advanced internet investment is only associated with wage growth in counties with high levels of income, education, population, and IT-intensity industry. This is followed by robustness checks, instrumental variables analysis, and an analysis of the timing of the relationship between internet investment and wage growth in the counties that were already

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

doing well. Finally, we explore some additional implications of advanced internet investment.

5.1 Internet investment and average wages

In Table 2, we show the baseline results across counties. Column 1 shows the correlation between advanced internet investment and wage growth at the county level without any controls. As suggested by the scatterplots in Figure 1, the correlation is significant and positive. Column 2 provides what we view as our main specification: Namely, it includes controls for levels of presample demographics (such as county income and population in 1990) and presample innovativeness. It also includes controls for changes in non-income demographics (such as population from 1990 to 2000 and net migration from 1990 to 2000) and changes in home internet adoption (effectively zero in 1995). Column (3) uses an alternative weighting for the advanced internet variable, equal to *Total number of establishments in our data in location-NAICS* divided by *Total # of census establishments in locations-NAICS.* Column 4 includes the 370 counties which contained no firms surveyed by Harte Hanks (and therefore lacking internet data). We use multiple imputation (from the values for the other county-level covariates) to impute values for the missing data.

In the main specification the coefficient on advanced internet is 0.0278. That is, regions with an average level of advanced internet (8.9%) experienced wage growth 0.247 percentage points above that of regions with no internet use. A one standard deviation increase in the use of the internet is associated with 0.370 percentage point increase in wage growth. The data are skewed, so it is also interesting to look at the top decile of advanced internet, which is 21.6%. That leads to a 0.353 percentage point increase in wage growth above the mean. Consistent with Figure 1a, this suggests that advanced internet was not the primary force behind the 20% wage growth across all counties in our data from 1995 to 2000.

Even with such a small coefficient, omitted variable bias is an important concern in this analysis. Below, after presenting our main results on regional variation in the relationship between wage growth and internet investment, we use instruments and the timing of divergence to argue for a causal explanation of our results.

In Table 3, we examine whether this first set of findings are specific to advanced internet, or if advanced internet might proxy for other kinds of IT. 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.[17]

The results suggest that advanced internet is distinct from other measures of IT. While PCs per employee are positively correlated with wage growth, this relationship is not statistically significant. Furthermore, including PCs per employee and basic internet as controls does not substantially change the marginal relationship between advanced internet 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 is related to advanced internet technology in particular.

The lack of correlation between basic internet technologies (e.g., email and web browsing) and wage growth is surprising because levels of adoption 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 may simply be too little variation in the independent variable.

[17] Forman, Goldfarb, and Greenstein (2005) use the same measure of basic internet investment and show it was widely adopted by 2000. The measure of PCs per employee resembles that used by Beaudry, Doms, and Lewis (2006).

5.2 When Was Advanced Internet Investment Related to Local Wage Growth?

In this section, we provide evidence that advanced internet investment led to divergence in wages across counties. In particular, as suggested by our framework, advanced internet investment was especially 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 4 explore this pattern in several steps. Column 1 shows that advanced internet is significantly associated with wage growth in counties in the top quartile of median income as of 1990. In contrast, counties in other quartiles with high levels of advanced internet did not experience especially rapid wage growth. In short, advanced internet contributes to wage divergence.

Columns 2 through 4 show how the impact of advanced internet investment is influenced by variation in local education levels, IT-intensity, and population. Like Column 1, Column 2 shows that advanced internet use is associated with wage growth only for high education counties. The similarity of results is not surprising because 60% of the counties overlap. Column 3 shows that counties with over 150,000 people display a strong association between advanced internet use and wage growth.[18]

Column 4 examines counties in the top quartile in IT-intensity. This specification shows that there is no statistically significant incremental gain from advanced internet investment in 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 and therefore forms part of the framework presented in section 2. Second, the coefficient is positive and when added to the coefficient on the main effect in the first row, it

[18] We use 150,000 as our threshold though results are robust to a continuous specification and to using other thresholds such as 100,000 and 200,000.

is significantly different from zero with 95% confidence. Third, we tried several specifications and the coefficient was sometimes significantly positive and never negative.

Column 5 shows that when we include all four measures of pre-internet county strength (income, education, population, and IT-intensity), only population appears significant. This may not be surprising given that there is considerable overlap between the measures: Each measure contains roughly 680 counties (high population, which is not based on quartiles, contains 315), of which 163 are in the top group in all measures. Column 6 shows that in these 163 counties advanced internet is strongly correlated with wage growth. Column 7 shows that it is the combination of more than one factor that drives the relationship between advanced internet and wage growth.

What does this mean? Increases in advanced internet investment will lead to higher wage growth in the 163 counties than in the other 2580 counties in the sample. These results suggest that advanced internet is related to 22.7% (6.5 percentage points out of 28.6 on average) of the total wage growth in the 163 counties that were already doing well in 1990. For the other counties, advanced internet explains just 1% (0.21 percentage points out of 20.5 on average) of overall wage growth.[19] Using back of the envelope calculations, this means that advanced internet explains over half of the 8.1 percentage point difference in wage growth between the average for those 163 counties and the other 2580 counties in the sample.[20] These 163 counties represent 42% of the US population. While our results do not pertain to all of the population in these counties, we believe it is evidence that the internet explains a substantial part of faster wage growth for the skilled labor force in those areas.

[19] These calculations use the coefficient estimates in Table 4 column 6, the average internet use for the 163 counties, and the average internet use in all other counties.

[20] More precisely, for the approximately 40 counties out of the sample of 163 counties with low internet investment, the investment contributes little to explaining the difference in wage growth. Similarly, for the approximately 80 counties with mean values or higher, the internet explains as much as half and or more of the differences in wage growth. Indeed, at the max 0.253 (Arapahoe, CO) the internet can explain all the additional wage growth.

The core results of Table 4 are robust to using continuous measures of income, education, population, and IT-intensity. Income loses statistical significance and IT-intensity gains significance, but the significance and importance of the interaction term remains. Furthermore, adding all two-way interactions to Column 6 (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 163 counties that were already doing well on all four measures.[21]

We stress these results reflect a general experience found in a special set of urban counties. It does not arise solely from the inordinate influence of canonical outliers. For example, removing Santa Clara or San Francisco from the data set does not change the qualitative results. In part, this should not be surprising; no single variable, not even advanced internet investment, could possibly explain the anomalous experience in Santa Clara in this time period (i.e., over 80% wage growth in five years). More to the point, these counties shared similar demographic and industrial traits prior to the internet's diffusion and reacted to the adoption and use of the internet in business with similar economic experiences, and heterogeniety in experience is consistent with the general explanation. 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.[22]

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

[22] Counties among the top 163 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 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). Only Hudson NJ (New York-Northern New Jersey-Long Island MSA), has high wage growth (one standard deviation above mean) but relatively low advanced internet use (below mean).

5.3: Justifying a causal interpretation: Instrumental variables and the timing of divergence

This section provides the results of a variety of additional tests we run to address omitted variable bias and simultaneity. We first discuss the results of a series of instrumental variables estimates. Three of our instruments are correlated with local costs of internet investment. First, we instrument using variance in the costs of internet deployment among 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. Forman, Goldfarb, and Greenstein (2008) show that establishments that are part of firms with many programmers in other locations adopt faster (even if there are few programmers at the focal establishment). They argue for a causal interpretation, partly because these programmers would have been hired for reasons other than internet investment. In other words, programmers elsewhere in the firm make internet investment at the focal establishment more likely. We use the average across establishments within a county as an instrument. In these regressions, we also include a control for multi-establishment firms, since the variable is defined only for such firms.

Our second instrument is the number of local county connections to ARPANET—a wide area data communications network that was a predecessor of the internet—which will capture local data communications infrastructure and expertise.

Our third instrument proxies for local deployment costs: the year in which the local state capped the prices that ILECs could charge entrants, including data communications companies.[23] By influencing local costs of internet deployment through additional regulation, this variable should be correlated with local county internet adoption.

All three of these variables are unlikely to be correlated with unobservables influencing local wage growth. Our programmers variable reflects the presence of IT skills in linked counties; ARPANET reflects historical decisions (from the 1970s) about connectivity to

[23] Based on data from Abel and Clements (1998).

Department of Defense or US university networks; and price caps reflect exogenous state government decisions.

Our last instrument is an industry-level proxy of the demand for advanced internet investment outside the focal county, which is sometimes called a Bartik index.[24] For each county, we compute the mean propensity to adopt internet by industry. This is average industry adoption excluding the establishments in the focal county. We then sum these industry propensities up, using as weights the percentage of establishments in each industry in the local county.[25] To the extent that it reflects industry-level propensities to adopt advanced internet and variance in industry composition across counties this variable should be correlated with adoption; however it excludes local and establishment-specific features of the county and so should be uncorrelated with local wage growth. This instrument therefore links industry to wage growth, and assumes advanced internet as the mechanism.

Table 5 presents the results of LIML instrumental variable estimates of Table 2 Column 2. We present the results of just-identified median-unbiased results using only the programmers and Bartik index instruments in Columns 1 and 2, a combination of these two instruments in Column 3, and the combination of all four of our instruments in Column 4. We focus on the programmers and Bartik instruments because they are stronger in the first stage. The first stage results suggest that advanced internet investment is increasing in the number of linked programmers found elsewhere in county establishments, in industry propensity to adopt advanced internet, (weakly) in the number of historical ARPANET nodes, and (weakly) when the state adopted price cap regulations earlier. The F-statistic for the first stage instruments ranges from 14.16 for our just-identified estimates using programmers to a weaker 2.84 for our

[24] Our index shares similarities with similar indexes used by Bartik (1991) and Blanchard and Katz (1992).

[25] Formally, for each county i, and industry j, compute $\hat{\theta}_{ij}$, the average adoption rate for industry j excluding the establishments in county i. The instrument is equal to $\hat{\rho}_i = \sum_j \gamma_{ij}\hat{\theta}_{ij}$, where γ_{ij} is the share of establishments in industry j in county i.

just-identified estimates using the Bartik measure of industry propensity to adopt advanced internet. The results of these regressions remain qualitatively similar to our main specification in Column 2 of Table 2, with the coefficient on advanced internet remaining statistically significant at the 5% level in all results except for those in Column 2. A Hausman test retains the null that the coefficients in Table 5 and Column 2 of Table 2 are the same with p- ranging between 0.9982 and 1.0000. For the overidentified regressions, the p-value of the overidentification test statistic ranges between 0.2110 and 0.2229. While the results are somewhat noisy, consistent with our OLS estimates in Table 2 these IV estimates do suggest a statistically significant but economically weak link between advanced internet investment and wage growth.

Next, we turn to instrumental variables analysis of our stronger divergence results. Table 6 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 8 instruments for 2 potentially endogenous variables. The F-statistics for the first stage estimates for advanced internet range from 10.26 to 11.04 and for the first stage of advanced internet and *HighAllFactors* they range from 161.08 to 175.35. Therefore, as is clear from the significance of the interactions in the first stage, the instruments for the *HighAllFactors* and advanced internet interaction are quite strong. The second stage estimates support the results of Column 6 of Table 4 that the marginal effect of advanced internet on local wages is stronger in *HighAllFactors* counties than in other counties. Moreover, with the exception of Column 1, advanced internet's interaction with *HighAllFactors* is positive and statistically significant, and of similar magnitude to the related estimate in Column 6 of Table 4. Further, a Hausman test retains the null that the coefficients in Table 6 and Column 6 of Table 4 are the same (with p-values ranging between 0.9998 and 1.0000). Further, in Columns 3 and 4, the p-value of the overidentification test statistic ranges between 0.1738 and 0.3611. Because of the strength of the interacted instruments in the first stage and the robustness of the results in the second stage, we view Table 6 as suggestive of a causal relationship from advanced internet investment to wage growth.

Next, we examine the timing of divergence. As noted above, advanced internet investment should only contribute to divergence in wages in the latter half of the 1990s. So, as a falsification test, we examine whether our measure of internet investment contributes to divergence prior to 1995.

Figure 2 provides a graphical representation of the results of this falsification test. It shows a replication of the results in Column 6 of Table 4 using a panel of all years from 1990 to 2000. The controls are the same as in Column 6 and the dependent variable is logged wages. We interact year dummies from 1991 to 2000 with the measure of advanced internet (as of 2000) and the interaction of advanced internet with *HighAllFactors*. This generates a measure of the association between advanced internet (measured as of 2000) and wage divergence over the period. We expect no relationship between the advanced internet measure and divergence prior to the actual diffusion of the internet. Figure 2 clearly shows this pattern: advanced internet is not correlated with divergence until 1996 (when the internet began to diffuse widely). Between 1991 and 1995 the coefficients on both variables are statistically indistinguishable from zero in every year. In other words, there is no evidence of a correlation between advanced internet investment and divergence between 1990 and 1995. Starting in 1996, we begin to see divergence associated with advanced internet investment. In these latter years, the association between advanced internet and local wage growth in well-off counties is larger than that in other counties, and this difference is statistically significant. Further, all of the coefficients for the interaction between advanced internet and *HighAllFactors* counties over 1996-2000 are greater than the coefficients for the same interaction over 1991-1994 (and these differences are also statistically significant).

5.4 Additional Implications of Advanced Internet Investment

In this section we discuss several additional implications of our labor demand framework and extensions of our main results. In particular, we first examine the implications of advanced internet investment for growth in employment. Next, we study the implications for wages in counties where the elasticity of supply is likely to be high. We also examine whether the effects persisted after the dot-com crash.

In Table 7 we examine the implications of advanced internet investment for employment growth. The results are consistent with the prediction that advanced internet increased labor demand in counties with high skill, income, population, and levels of IT-intensive firms. Column 2 shows that employment growth is significantly correlated with advanced internet in *HighAllFactors* counties (the coefficient is 0.2025). At average values of advanced internet within *HighAllFactors* counties (13.5%), this suggests that *HighAllFactors* counties experienced employment growth 2.7 percentage points larger than all other counties as a result of investment in advanced internet. In addition, our findings show no measurable relationship between advanced internet and employment for other counties.

One implication of our framework is that the effects of advanced internet on local wages will be strongest in places with tight labor markets, particularly for skilled workers. Table 8 shows the results of two approaches to explore this hypothesis. In Columns 1 through 4 we provide estimates of the association between advanced internet and local wages for subsamples of counties where the overall unemployment rate and unemployment rate among skilled workers (college graduates and above) was above and below the median. We use 1990 Census data from the 532 metropolitan area counties for which skilled unemployment data are available.[26] Re-estimating the model in Column 2 of Table 2 over each of these subsamples demonstrates a clear pattern. While advanced internet has an economically strong and statistically significant association with local wages for counties where the overall unemployment and skilled unemployment rate are below the median, there is no such connection between advanced internet and local wages in other counties.

In Columns 5 and 6 of Table 8 we examine the association between advanced internet and wages for counties inside and outside the rust belt.[27] Rust belt counties will have, on average, both more elastic labor markets and weak output markets relative to other counties. Outside of the rust belt we find the results continue to hold: advanced internet is related to local wage growth. However, we find no connection between advanced internet and wage growth within the rust belt. In sum, we interpret the results of Table 8 as providing evidence in support of our framework's prediction about the implications for advanced internet in locations where local labor supply is elastic: while the labor demand curve may shift in response to advanced internet investment, wages do not adjust accordingly.

Table 9 examines whether the counties that gained disproportionately from the internet lost these gains after the dot-com crash. Specifically, it repeats the regressions in Column 2 of Table 2 and Column 6 of Table 4, but uses wage growth between 1999 and 2005 as the dependent variable. The results suggest that these counties maintained their new position in absolute terms. They did not grow faster, but their gains were not reversed either.

[26] Our source of unemployment data is the 5% microdata sample from the Integrated Public Use Microdata Series (IPUMS) data. We focus on metropolitan area counties because non-metro areas have small cell sizes and because converting from public use microdata areas (in the IPUMS data) to counties gives rise to additional errors in variables in our sample.
[27] The rust belt is defined as Illinois, Indiana, Ohio, Wisconsin, Michigan, New York outside of the New York City CMSA, Pennsylvania outside of the Philadelphia CMSA, Maryland, and West Virginia.

5.5 Open Issues about Biased Technical Change

Our study emphasizes the role of regional variation in the economic impact of information technology: advanced internet investments benefited locations that were already doing well. Furthermore, the benefits were not simply accruing to the IT-producing firms in Silicon Valley and Route 128 outside Boston. Combined, this raises a number of questions that are beyond the scope of currently available data.

Specifically, our findings motivate further research on the mechanisms connecting these results to the extant literature on skill-biased technical change. Two broad hypotheses are consistent with our results: (1) It is primarily workers in the information technology sector whose wages rise with advanced internet, and their wages rise sufficiently in some places to increase average wages. (2) The gains from advanced internet investment are experienced throughout the skilled workforce. Unfortunately, outside of a handful of major cities, consistent wage data at the county level for programmers and other IT-intensive occupations were not kept earlier than 1999. This means that our current framework is unable to distinguish between these hypotheses.

A second open question relates to the long run impact of advanced internet investments (and information and communication technology more generally) on geographic variation in wages. With time, labor mobility might dampen the impact of further investments in advanced internet on wage disparity, or perhaps that the wage divergence may disappear over time. While the results of Table 9 suggest that there was no reversion by 2005, understanding the longer run effects requires a more recent census of advanced internet investment. Following the dot-com crash, Harte Hanks scaled back the scope of their survey and, to our knowledge, no private or public sector survey has arisen to take its place. In the absence of a new survey, it will be difficult to assess the impact of the internet's diffusion.

A third open question is the impact of advanced internet investment on real wages. We do not attempt to measure whether nominal wage increases become permanent changes in worker income or eventually become a transfer to landowners through higher rents. The distinction is inessential to our goal of measuring geographic variance in wage growth across the country.

6. Discussion

In this study, we find evidence of an association between investment in advanced internet technology and local wage growth. We also find that wage gains associated with advanced internet investment 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. In addition, we find evidence that use of advanced internet was associated with employment growth in these top locations, but find little evidence of a connection between internet and employment growth elsewhere. Last, we find that the association between internet and wages holds primarily where labor markets were tight.

Thus, despite recent assertions that internet use may lower the costs of geographically isolated economic activity, there is no evidence in our data that advanced internet contributed to convergence in wages. In particular, our results suggest the existence of a considerable divide in the benefits of advanced internet investment across urban and rural areas.

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.

Our finding that the internet does not contribute to convergence 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. In the absence of considerable investment in human (and computing) capital, our results suggest that efforts to subsidize internet development in low density settings would have little economic impact.

References

Abel, Jaison and Michael E. Clements. 1998. A time series and cross sectional classification of state regulatory policy adopted for local exchange carriers," NRRI report 98-25, Columbus Ohio, The National Regulatory Research Institute. December.

Arora, Ashish 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.

Autor, David, Lawrence F. Katz, and Melissa S. Kearney. 2006. The Polarization of the U.S. Labor Market. *American Economic Review* 96(2): 189-194.

———, 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.

Bartik, Timothy. 1991. Who Benefits from State and Local Economic Development Policies? Kalamazoo, MI: W.E. Upjohn Institute.

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.

Blanchard, Olivier and Lawrence Katz. 1992. Regional Evolutions. *Brookings Papers on Economic Activity* 1992: 1-61.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. 2009. The distinct effects of Information Technology and Communication Technology on Firm Organization. NBER Working Paper #14975.

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

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.

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 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.

Garicano, Luis and Esteban Rossi-Hansberg. 2006. Organization and Inequality in a Knowledge Economy. *Quarterly Journal of Economics* 121(4): 1383-1485.

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

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

Greenstein, Shane, 2010, "Innovative Conduct in Computing and Internet Markets," (Eds) Bronwyn Hall and Nathan Rosenberg, *Handbook of Economics of Technical Change.*

Henderson, J. Vernon. 2003. Marshall's Scale Economies. *Journal of Urban Economics* 53: 1-28.

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.

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

Katz, Lawrence F. and David H. Autor. 1999. Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics, Volume 3A,* eds. Orley Ashenfelter and David Card, 1463-1558.

Kolko, Jed. 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.

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

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.

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
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
Bartik index	0.1126	0.0216	0	0.2664	2743
ARPANET connections	0.0215	0.2383	0	7	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
Home internet use	Percentage of households with internet at home (2000)	Current Population Survey (CPS) internet Use Supplement (Census)	0.444
Home internet data missing	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)	US Census	89173
% African American	% Population African American as of Decennial Census (1990)	US Census	0.0908
% University Graduates	% Population university graduates as of Decennial Census (1990)	US Census	0.1379
% High School Graduates	% Population high school graduates as of Decennial Census (1990)	US Census	0.6996
% Below Poverty Line	% Population below poverty line as of Decennial Census (1990)	US Census	0.1622
Median Household Income	Median county household income as of Decennial Census (1990)	US 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 County 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	US Census	0.3258
Net Migration	Net migration to county	US Census	123.5
% Population over Age 65	% of county population over 65 as of Decennial Census	US Census	0.1452

Table 2: Wages increase with internet use

	(1)	*(2)*	*(3)*	*(4)*
	No controls	Full set of controls	Alternative weighting	Multiple imputation of missing data on advanced internet
Advanced internet	0.0372	0.0278	0.0245	0.0243
	(0.0132)***	(0.0126)**	(0.0111)**	(0.0143)*
Home internet use		0.0823	0.0832	0.0873
		(0.0379)**	(0.0377)**	(0.0393)**
Home internet data missing		0.0281	0.0288	0.0337
		(0.0170)*	(0.0170)*	(0.1778)*
Log population in 1990		-0.0065	-0.0066	-0.0037
		(0.0019)***	(0.0019)***	(0.0021)*
Percentage African Americans in 1990		0.0133	0.0134	0.0232
		(0.0118)	(0.0118)	(0.0130)*
Percentage university graduates in 1990		0.5720	0.5731	0.4219
		(0.0789)***	(0.0785)***	(0.0912)***
Percentage high school graduates in 1990		-0.1555	-0.1539	-0.0590
		(0.0520)***	(0.0521)***	(0.0656)
Percentage below poverty line in 1990		-0.1615	-0.1605	-0.0646
		(0.0464)***	(0.0464)***	(0.0500)
Median income in 1990 ($000)		-0.0006	-0.0006	0.0002
		(0.0006)	(0.0006)	(0.0007)
Percentage population attending Carnegie Type 1 schools in 1990		0.0320	0.0311	0.0322
		(0.0475)	(0.0467)	(0.0463)
Percentage population enrolled in engineering program in 1990		-0.2202	-0.2127	-0.1197
		(0.3630)	(0.3588)	(0.3447)
# of patents granted to inventors in the county in the 1980s (000)		0.0165	0.0165	0.0153
		(0.0043)***	(0.0043)***	(0.0043)***
Percentage professional in 1995		-0.0102	-0.0124	0.448
		(0.0535)	(0.0536)	(0.0674)
Percentage of persons over age 65 in 1990		0.0443	0.0449	0.1485
		(0.0513)	(0.0512)	(0.0650)**
Net migration into the county in 1995 (000)		0.0033	0.0033	0.0022
		(0.0032)	(0.0032)	(0.0032)
Change in log total population between 1990 and 2000		0.0527	0.0522	0.0820
		(0.0152)***	(0.0152)***	(0.0176)***
Change in percentage of African American 1990 to 2000		0.0265	0.0319	-0.0422
		(0.0756)	(0.0754)	(0.0814)
Change in percentage of university graduates 1990 to 2000		0.8219	0.8235	0.6358
		(0.1604)***	(0.1605)***	(0.1799)***
Change in percentage of high school graduates 1990 to 2000		-0.0224	-0.0214	0.07760
		(0.0947)	(0.0947)	(0.1200)
Change in percentage of persons over age 65 1990 to 2000		-0.5628	-0.5602	-0.4853
		(0.1192)***	(0.1192)***	(0.1326)***
Change in net migration into the county 1990 to 2000 (000)		0.0020	0.0020	0.0011
		(0.0037)	(0.0037)	(0.0037)
Constant	0.1848	0.2995	0.2983	0.1597
	(0.0017)***	(0.0458)***	(0.0458)***	(0.0524)***
Observations	2743	2743	2743	3133
R-squared	0.004	0.131	0.131	N/A (F=12.4)

Dependent variable is change in logged wages from 1995 to 2000. Heteroskedasticity-robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Is advanced internet different from other measures of IT use?

	(1)	*(2)*	*(3)*
	Compare all three measures of IT use	Compare advanced internet and basic internet	Compare advanced internet and PCs per employee
Advanced internet	0.0247	0.0261	0.0248
	(0.0135)*	(0.0132)**	(0.0133)*
Basic internet	0.0007	0.0050	
	(0.0078)	(0.0075)	
PCs per employee	0.0152		0.0156
	(0.0108)		(0.0104)
Home internet use	0.0822	0.0823	0.0822
	(0.0379)**	(0.0378)**	(0.0379)**
Home internet data missing	0.0282	0.0279	0.0282
	(0.0170)*	(0.0170)	(0.0170)*
Log population in 1990	-0.0068	-0.0067	-0.0068
	(0.0019)***	(0.0019)***	(0.0019)***
Percentage African Americans in 1990	0.0124	0.0134	0.0123
	(0.0119)	(0.0119)	(0.0118)
Percentage university graduates in 1990	0.5590	0.5671	0.5594
	(0.0807)***	(0.0802)***	(0.0802)***
Percentage high school graduates in 1990	-0.1589	-0.1569	-0.1587
	(0.0522)***	(0.0522)***	(0.0521)***
Percentage below poverty line in 1990	-0.1598	-0.1600	-0.1600
	(0.0463)***	(0.0464)***	(0.0464)***
Median income in 1990 ($000)	-0.0006	-0.0005	-0.0006
	(0.0006)	(0.0006)	(0.0006)
Percentage population attending Carnegie Type 1 schools in 1990	0.0338	0.0323	0.0338
	(0.0480)	(0.0476)	(0.0480)
Percentage population enrolled in engineering program in 1990	-0.2357	-0.2162	-0.2367
	(0.3647)	(0.3648)	(0.3644)
# of patents granted to inventors in the county in the 1980s (000)	0.0160	0.0164	0.0160
	(0.0043)***	(0.0043)***	(0.0043)***
Percentage professional in 1995	-0.0089	-0.0078	-0.0093
	(0.0543)	(0.0543)	(0.0537)
Percentage of persons over age 65 in 1990	0.0470	0.0435	0.0472
	(0.0513)	(0.0512)	(0.0513)
Net migration into the county in 1995 (000)	0.0034	0.0033	0.0034
	(0.0032)	(0.0032)	(0.0032)
Change in log total population between 1990 and 2000	0.0539	0.0528	0.0539
	(0.0153)***	(0.0152)***	(0.0153)***
Change in percentage of African American 1990 to 2000	0.0251	0.0248	0.0253
	(0.0759)	(0.0759)	(0.0758)
Change in percentage of university graduates 1990 to 2000	0.8161	0.8169	0.8167
	(0.1613)***	(0.1617)***	(0.1605)***
Change in percentage of high school graduates 1990 to 2000	-0.0259	-0.0227	-0.0260
	(0.0947)	(0.0948)	(0.0947)
Change in percentage of persons over age 65 1990 to 2000	-0.5621	-0.5629	-0.5620
	(0.1190)***	(0.1191)***	(0.1190)***
Change in net migration into the county 1990 to 2000 (000)	0.0022	0.0020	0.0022
	(0.0037)	(0.0037)	(0.0037)
Constant	0.3006	0.2978	0.3009
	(0.0460)***	(0.0460)***	(0.0458)***
Observations	2743	2743	2743
R-squared	0.13	0.13	0.13

Dependent variable is change in logged wages from 1995 to 2000. 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: Effect primarily occurs in places that are already high income, education, IT-intensity, *AND* population

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Advanced internet	0.0168	0.0120	0.0246	0.0214	0.0049	0.0239	0.0067
	(0.0137)	(0.0125)	(0.0127)*	(0.0159)	(0.0149)	(0.0128)*	(0.0150)
Advanced internet and	0.0960				0.0442		0.0377
High income county	(0.0389)**				(0.0492)		(0.0496)
Advanced internet and		0.1101			0.0770		0.0757
High education county		(0.0455)**			(0.0548)		(0.0547)
Advanced internet and			0.3631		0.2378		0.0182
High population county			(0.0934)***		(0.1018)**		(0.1027)
Advanced internet and				0.0206	0.0134		0.0102
High IT-intensity county				(0.0228)	(0.0235)		(0.0237)
Advanced internet and High income, education,						0.4588	0.3393
IT-intensity, and population county						(0.1585)***	(0.1904)*
Observations	2743	2743	2743	2743	2743	2743	2743
R^2	0.13	0.13	0.13	0.13	0.14	0.14	0.14

Dependent variable is change in logged wages from 1995 to 2000. In addition to the controls in table 2, regressions include dummies for the main effects of the interactions where appropriate (high income, high eduction, high IT-intensity, high population, and high all factors). Internet at home is not included because Internet home data missing is collinear with high population. Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%.

Table 5: Instrumental variables analysis of Table 2 column 2

	(1)	*(2)*	*(3)*	*(4)*
Instrument→	Programmers in other establishments within the same firm	Bartik index	Both main instruments	Two main instruments plus ARPANET nodes and year when state adopted a price cap
FIRST STAGE: Dependent variable is advanced internet				
Programmers in other establishments within the same firm	0.0002		0.0002	0.0002
	(0.0000)***		(0.0001)***	(0.0001)***
Bartik index		0.2990	0.2584	0.2554
		(0.1774)*	(0.1788)	(0.1791)
ARPANET nodes				0.0055
				(0.0046)
Year when state adopted a price cap				0.0013
				(0.0008)
Partial R-squared	0.0082	0.0022	0.0098	0.0106
F-Statistic	14.16	2.84	8.39	5.26
Observations	2743	2743	2743	2743
R-squared	0.0282	0.0223	0.0298	0.0217
SECOND STAGE: Dependent variable is logged wages				
Advanced internet	0.3691	0.0156	0.3174	0.4014
	(0.1450)**	(0.2374)	(0.1356)**	(0.1549)***
Overidentification test (p-value)	N/A	N/A	0.2110	0.2229
Hausman test (p-value)	0.9997	1.0000	0.9999	0.9982
Observations	2743	2743	2743	2743
R-squared	0.131	0.133	0.133	0.133

Dependent variable is change in logged wages from 1995 to 2000. Heteroskedasticity-robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Instrumental variables analysis of Table 4 column 6

	(1)	*(2)*	*(3)*	*(4)*
	Programmers in other establishments within the same firm	Bartik index	Both main instruments	Two main instruments plus ARPANET nodes and year when state adopted a price cap
FIRST STAGE: Dependent variable is advanced internet				
Programmers in other establishments within the same firm	0.0001		0.0001	0.0001
	(0.0000)***		(0.0000)***	(0.0000)***
Bartik index		0.2681	0.2227	0.2296
		(0.1809)	(0.1813)	(0.1822)
ARPANET nodes				0.0089
				(0.0173)
Year when state adopted a price cap				0.0012
				(0.0008)
Programmers in other establishments within the same firm and high all factors	0.00005		0.00005	0.00005
	(0.0001)		(0.0001)	(0.0001)
Bartik index and high all factors		0.4084	0.1282	0.1988
		(0.7077)	(0.7325)	(0.7160)
ARPANET nodes and high all factors				-0.0051
				(0.0175)
Year when state adopted a price cap and high all factors				0.0007
				(0.0012)
Partial R-squared	0.0076	0.0017	0.0073	0.0097
F-Statistic	10.98	10.48	11.04	10.26
FIRST STAGE: Dependent variable is advanced internet and high all factors				
Programmers in other establishments within the same firm	-4.44e-07		-3.93e-07	-2.96e-07
	(4.36e-07)		(4.09e-07)	(3.86e-07)
Bartik index		0.0009	0.0009	0.0010
		(0.0015)	(0.0014)	(0.0011)
ARPANET nodes				-0.0002
				(0.0006)
Year when state adopted a price cap				-0.00003
				(0.00001)**
Programmers in other establishments within the same firm and high all factors	0.0002		0.0002	0.0002
	(0.0001)**		(0.0001)*	(0.0001)*
Bartik index and high all factors		1.3382	0.9366	0.9976
		(0.5054)***	(0.5180)*	(0.4966)**
ARPANET nodes and high all factors				0.0052
				(0.0030)*
Year when state adopted a price cap and high all factors				0.0020
				(0.0007)***
Partial R-squared	0.0656	0.0443	0.0856	0.1260
F-Statistic	169.80	175.35	174.30	161.08
SECOND STAGE: Dependent variable is logged wages				
Advanced internet	0.3273	-0.1218	0.1660	0.3690
	(0.1452)**	(0.2892)	(0.1565)	(0.1782)**
Advanced internet and High income, education, IT-intensity, and population county	0.6139	1.9206	1.1108	0.9331
	(0.5595)	(0.8584)**	(0.5531)**	(0.4692)**
Overidentification test (p-value)	N/A	N/A	0.1738	0.3611
Hausman test (p-value)	1.000	1.000	1.000	0.9998
Observations	2743	2743	2743	2743
R^2	0.07	0.06	0.08	0.08

Dependent variable is change in logged wages from 1995 to 2000. In addition to the controls in table 2, regressions include dummies for the main effects of the interactions where appropriate (high income, high education, high IT-intensity, high population, and high all factors. Internet at home is not included because it is collinear with high all factors. Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 7: Employment

	(1)	(2)
	Full set of controls	High all factors interaction and full set of controls
Advanced internet	-0.0190	-0.0206
	(0.0164)	(0.0166)
Advanced internet and High income, education, IT-intensity, population county		0.2025
		(0.1096)*
Observations	2743	2743
R^2	0.32	0.32

Dependent variable is change in logged employment from 1995 to 2000. In column (1) controls are the same as table 2. In column (2) controls are the same as table 4. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%.

Table 8: Additional implications of advanced internet investment

	Unemployment rate		*Unemployment rate among college graduates*		*The Rust Belt*	
	(1)	(2)	(3)	(4)	(5)	(6)
	Below median	At or above median	Below median	At or above median	In rust belt	Outside rust belt
Advanced internet	0.1027	0.0787	0.1665	-0.0203	0.0030	0.0311
	(0.0599)*	(0.0682)	(0.0507)***	(0.0730)	(0.0327)	(0.0138)**
Observations	266	266	263	269	642	2101
R^2	0.31	0.29	0.33	0.30	0.10	0.16

Dependent variable is change in logged wage growth from 1995 to 2000. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%.

Columns (1) through (4) show estimates of the association between advanced internet and local wages for subsamples of counties where the overall unemployment rate and unemployment rate among skilled workers was above and below the median. Dependent variable is logged wages. Controls are the same as in Table 2. We found reliable county-level unemployment numbers for 532 counties.

Columns (5) and (6) show estimates of the association between advanced internet and local wages for subsamples of counties inside and outside the rust belt. Dependent variable is logged wages. Controls are the same as table 2. The rust belt is defined as Illinois, Indiana, Ohio, Wisconsin, Michigan, New York outside of the New York City CMSA, Pennsylvania outside of the Philadelphia CMSA, Maryland, and West Virginia.

Table 9: Wage growth from 1999 to 2005 is not related to early use of advanced internet

	(1)	*(2)*
Advanced internet	-0.0057	-0.0053
	(0.0132)	(0.0134)
Advanced internet and High income, education, IT-intensity, and population county		0.0007
		(0.1023)
Observations	2743	2743
R^2	0.06	0.06

Dependent variable is change in logged wage growth from 1999 to 2005. In column (1) controls are the same as table 2. In column (2) controls are the same as table 4.
Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%.

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

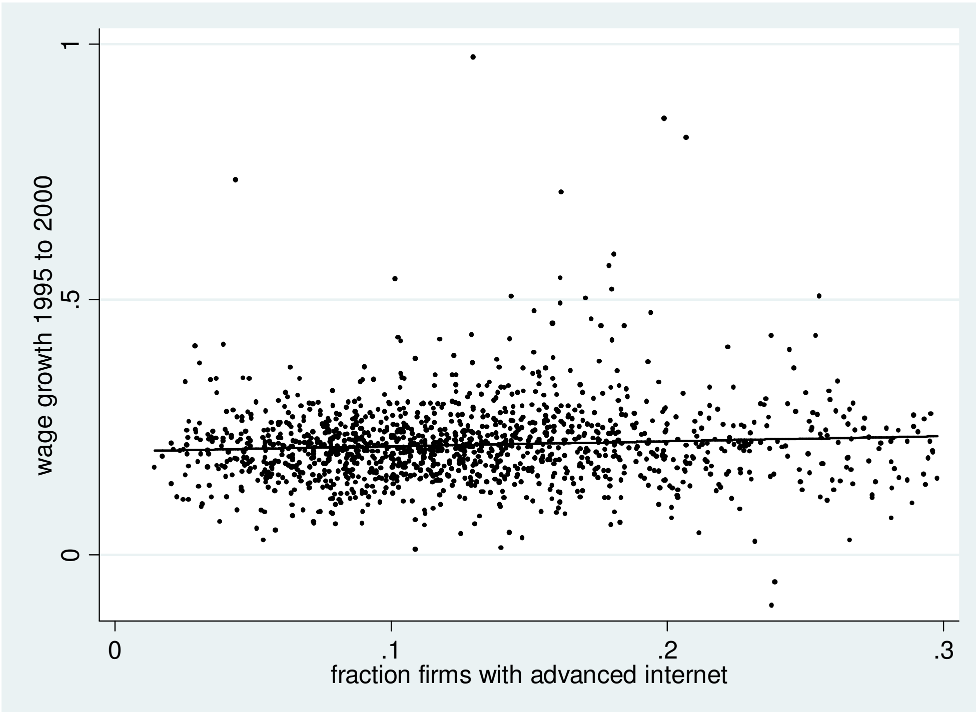


Figure 1b: Advanced Internet Investment and Wage Growth by County Type

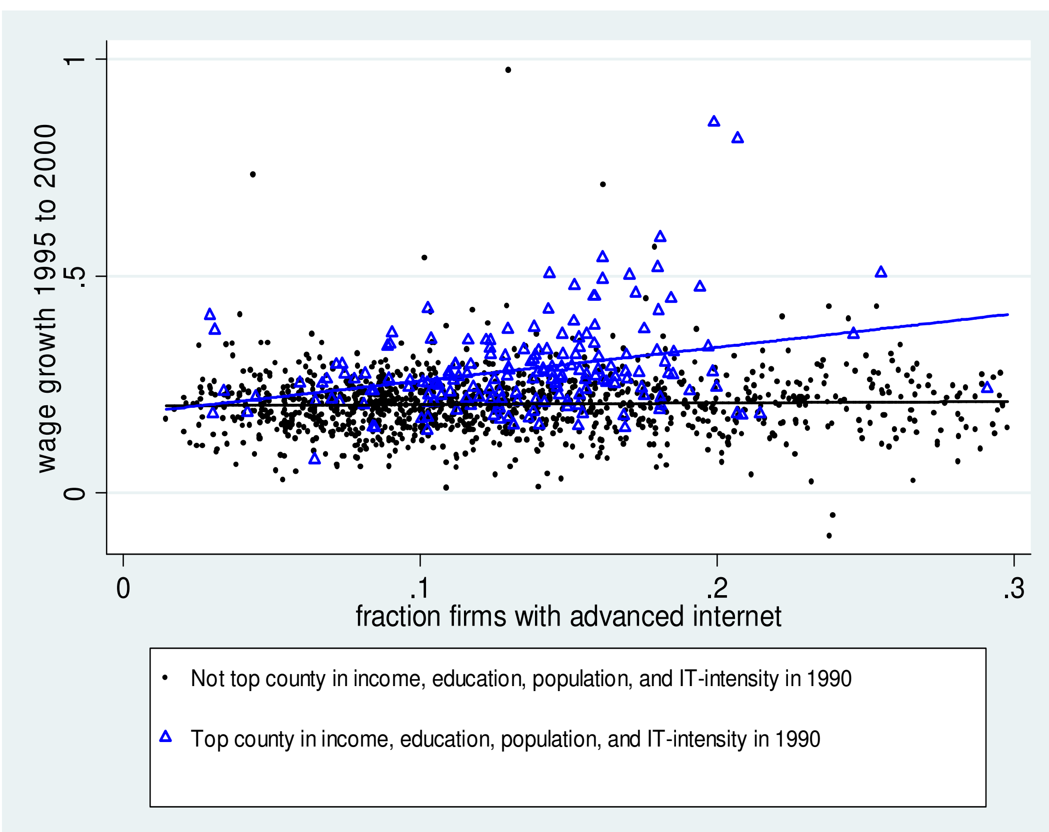
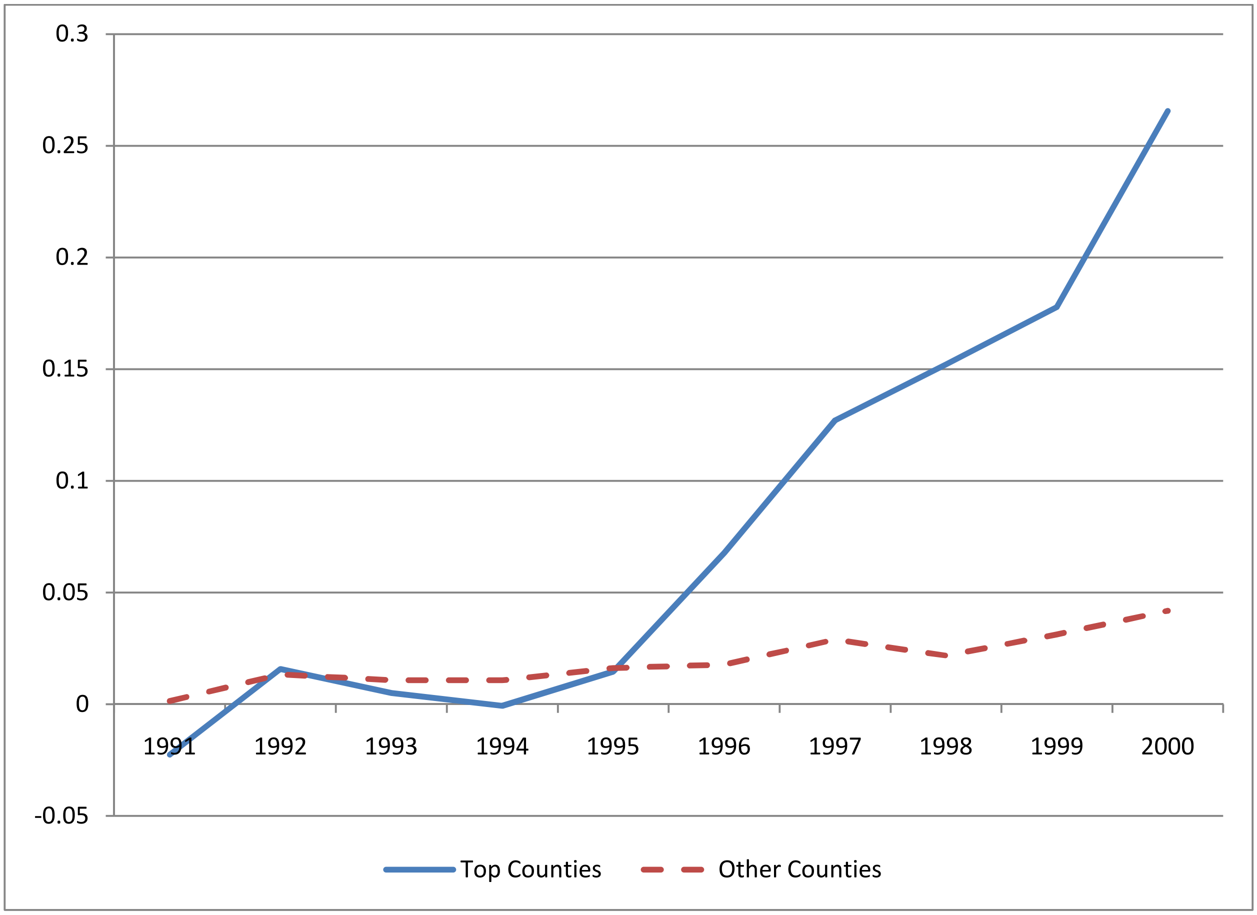


Figure 2: Marginal effect of advanced internet year-by-year in top counties



This is based on a panel version of the regression model is table 4 column 6 where 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 4.

ONLINE APPENDIX TO

THE INTERNET AND LOCAL WAGES:

CONVERGENCE OR DIVERGENCE?

A: Data Appendix

B: Appendix Tables

NOT FOR PUBLICATION

A. Data Appendix

In this section we discuss the construction of our data set. We first describe the construction of our measures of internet investment, and then briefly describe our measures of county characteristics that we interact with advanced internet. Last we describe the construction of our instruments.

A.1 Construction of measures of advanced internet investment.

As noted in the paper, our IT data come from the Harte Hanks Market Intelligence Computer Intelligence Technology database (hereafter CI database). Harte Hanks tracks over 300,000 establishments in the United States. For the reasons described in the paper, 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.[1]

The CI database contains several measures of internet usage that we use to construct measures of advanced internet. Advanced internet is the type of investment that has been historically described in books on electronic commerce. Typically this involves altering sales, manufacturing, production, or distribution systems within the firm. We aggregate many applications under this umbrella. Business-to-business or business-to-consumer e-commerce fall in this category, so too does TCP/IP versions of software such as enterprise resource planning or customer relationship management. Our measure of advanced internet assumes nothing about the *intensity* of use, nor about complexity.

An establishment is counted as making an investment in advanced internet when two or more of the following conditions hold: (1) the establishment uses two or more languages common in web applications, such as Active-X, Java, CGI, Perl, VB Script, or XML; (2) the

[1] Parts of this section draw from an earlier paper on the dispersion of internet investment, Forman, Goldfarb, and Greenstein (2002).

establishment has over five internet developers; (3) the establishment has two or more "e-business" applications such as customer service, education, extranet, publications, purchasing, or technical support; (4) the establishment reports LAN software that performs one of several functions: e-commerce, enterprise resource planning, web development, or web server; (5) the establishment has an internet server that is a UNIX workstation or server, mainframe, or minicomputer, or has 5 or more PC servers, or has internet storage greater than 20 gigabytes (this was a lot of storage in 2000); (6) the establishment indicates the use of three or more applications related to internet server software, internet/web software, or intranet applications.

We tested this definition and found that it generated many false positives. These false positives arose more frequently when an establishment was experimenting with, but not actually regularly using, advanced internet applications. In other words, they were trying something small or contracting out for a test. To correct for this, we exclude establishments if: (a) They indicate they have outsourced hosting of their internet/web servers; (b) These experimenters responded affirmatively to exactly two of (1) through (6) but not any question about e-commerce. Such establishments typically had not yet done very much advanced internet as of the time of our sample (but might someday). Previous work compared the baseline measure of advanced internet with one that includes such "experimenters" (Forman, Goldfarb, and Greenstein 2002). While that latter measure shows higher levels of internet penetration (23.2% v. 12.6%), the quantitative difference between the two measures remains similar across geographic regions.

Our measure of basic internet (used in Table 4) is constructed similarly. To be counted as investing in basic internet, an establishment must engage in two or more of the following activities: (1) have an internet service provider; (2) indicate it has basic access; (3) use commerce, customer service, education, extranet, homepage, publications, purchasing or technical support; (4) indicate that it has an intranet or email based on TCP/IP protocols; (5) indicate there are internet users or internet developers on site; or (6) outsource some internet activities. We looked for two or more activities to guard against "false positives". As it was, this was a minor issue. Most respondents responded affirmatively to many of these criteria.

The CI database also contains information on the number of personal computers and number of employees at the establishment. We divide the number of establishment personal computers by the number of employees to obtain our PCs per employee measure.

Timing bias and sampling bias are two concerns with these measures. We first discuss timing bias. Establishments in our sample were surveyed between July 1998 and August 2000. Because advanced internet diffused between 1998 and 2000, earlier respondents are likely to have a lower adoption rate. 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. Specifically, we divide our sample into six semi-annual periods between 1998 and 2000. For establishments who are surveyed in some semi-annual period *t* prior to the end of 2000, we reweight the adoption rate by (*average adoption rate in county at end of 2000*) / (*average adoption rate in county in semi-annual period t*).

To obtain location-level measures of the extent of advanced internet investment, 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 to correct for sampling bias. 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 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

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

In other words, the weights are the proportion of establishments in a location that are a given NAICS code, divided by the proportion of times it is in our database. This means that if our data undersamples a given two-digit NAICS at a location, then each observation in that NAICS-location is given more importance. We divide establishment adoption by the above weights and then 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. As a robustness check, we also show results with the following weights

$$\frac{\text{Total \# of establishments in our data in location}}{\text{Total \# of census establishments in location}}.$$

A.2 Construction of variables measuring county characteristics

The construction of our controls for county characteristics is described in Table 1b; here we describe the computation of the variables that we interact with advanced internet. As noted in the paper, we focus on the roles of *income*, *education*, *population*, and *IT-intensity*; these measures are equal to one when the corresponding continuous variable is in the highest quartile of the distribution. Income, education, and population data are from the 1990 Census: High Income is based upon median household income, High Education is based upon the fraction of the population that is university graduates, and High population is based upon 1990 county population estimates. 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).[2]

A.3 Construction of the instruments

Here we discuss the computation of our four instruments: the number of programmers in other establishments within the same firm, the "Bartik index" of expected local internet demand, the number of local ARPANET nodes, and the year when a state adopted a price cap.

To compute the number of programmers in other establishments, we use information on establishment programmers from the CI database. For each establishment that is part of a multi-establishment firm, we compute the number of programmers that reside within the same firm but in other counties. This variable is based on the "organizational capabilities" measure used in Forman, Goldfarb, and Greenstein (2008). We then compute the weighted average number of programmers for the county, using the weights described in section A.1.

Our measure of the Bartik index for local propensity of internet adoption uses our establishment-level internet data. For each county in our sample, we compute average industry adoption of advanced internet excluding the contribution of establishments in the industry-county, for industry j in county i we label this $\widehat{Internet}_{ij}$.[3] This is equal to the average propensity for an establishment in industry j to have advanced internet, excluding the contribution of establishments in county i. We then compute an index of advanced internet

---

[2] Specifically, they include the following industries: Communications (SIC 48), Business Services (Including Computer Services; SIC 73), Wholesale Trade (SIC 50-51), Finance (SIC 60-62, 67), Printing and Publishing (SIC 27), Legal Services (SIC 81), Instruments and Miscellaneous Manufacturing (SIC 38-39), Insurance (SIC 63-64), Machinery (Including Computers and Office Equipment; SIC 35), Gas utilities (SIC 492, parts of 493, 496), Professional and Social Services (SIC 832-839), Other Transportation Equipment (SIC 372-379), Other Electronic Machinery (including Communications Equipment and Electronic Components; SIC 36).

[3] Again, we use weighted adoption, where the weights are analogous to our location weights in section A.1. That is, we weight establishment adoption by ((total Census establishments in a location-industry)/(total Census establishments in an industry))×((total establishments in our data in an industry)/(total establishments in our data in a location-industry)).

investment in county i by weighting these industry propensities according to the fraction of establishments in county i that are in industry j, ρ_{ij}. That is, $Bartik_{ij} = \sum_j \rho_{ij} \widehat{Internet}_{ij}$. This is a generalization of similar share-weighted local demand proxies used in earlier work such as Bartik (1991) and Moretti (2009).

Our measure of number of local ARPANET nodes is simply a count of the number of local nodes, 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).

Our measure indicating the year when a state adopted a price cap or freeze, is the year that the state froze (or capped) the prices incumbent carriers could change entrants (source: Abel and Clements 1998). These regulatory caps were attempts to facilitate entry by competitive local exchange carriers.

References

Abel, Jaison and Michael E. Clements. 1998. A time series and cross sectional classification of state regulatory policy adopted for local exchange carriers," NRRI report 98-25, Columbus Ohio, The National Regulatory Research Institute. December.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2002. Digital Dispersion: An Industrial and Geographic Census of Commercial internet Use. NBER Working Paper #9287.

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

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

Moretti, Enrico. 2009. Real Wage Inequality. Working Paper, University of California, Berkeley.

B. Appendix Tables

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

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Advanced internet	-0.0371	-0.0424	0.0251	-0.0169	-0.0853	0.0242	-0.0411
	(0.0472)	(0.0360)	(0.0128)*	(0.0299)	(0.0570)	(0.0127)*	(0.0548)
Advanced internet x	3.02e-06				-9.06e-07		-1.83e-06
county-level income	(2.13e-06)				(3.60e-06)		(3.61e-06)
Advanced internet x		0.5897			0.6510		0.5850
county-level education		(0.3050)*			(0.4413)		(0.4526)
Advanced internet x			1.52e-07		1.09e-07		-1.79e-07
county-level population			(9.84e-08)		(9.34e-08)		(1.16e-07)
Advanced internet x				0.1947	0.2332		0.1738
county-level IT-intensity				(0.1006)*	(0.1007)**		(0.1011)*
Advanced internet x income x						9.88e-11	1.12e-10
education x population x IT-intensity						(2.66e-11)***	(3.66e-11)***
Observations	2743	2743	2743	2742	2742	2742	2742
R^2	0.13	0.13	0.13	0.13	0.14	0.14	0.15

Dependent variable is logged wages. Controls are the same as in table 5. 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)*
	Includes two-way interactions	MSA only		No controls	
Advanced internet	0.0001	0.0860	0.0528	0.0372	0.0232
	(0.0152)	(0.0362)**	(0.0392)	(0.0132)***	(0.0133)*
Advanced internet and High income, education, IT-intensity, and population county	0.4166		0.3949		0.6702
	(0.1996)**		(0.1566)**		(0.1864)***
Advanced internet and High income county	0.0832				
	(0.0469)*				
Advanced internet and High education county	0.1130				
	(0.0653)*				
Advanced internet and High IT-intensity county	0.0269				
	(0.0231)				
Advanced internet and High population county	-0.0654				
	(0.1152)				
Advanced internet and High IT-intensity and population county	0.0164				
	(0.0670)				
Advanced internet and High education and IT-intensity county	-0.0814				
	(0.0579)				
Advanced internet and High income and IT-intensity county	-0.0708				
	(0.0570)				
Advanced internet and High income and population county	0.0483				
	(0.0579)				
Advanced internet and High education and population county	0.0910				
	(0.0585)				
Advanced internet and High income and education county	-0.0918				
	(0.0712)				
Observations	2743	843	843	2744	2743
R^2	0.14	0.33	0.34	0.004	0.05

Dependent variable is logged wages. . In columns (2) and (4) controls are the same as table 2. In columns (1), (3), and (5) controls are the same as table 4. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

Online Appendix Table 3: Robustness to alternative weighting in constructing the advanced internet variable

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
	Weight is # observations in our data over # observations in census		No weights		No weights for county differences		No weights for time differences	
Advanced internet	0.0245	0.0217	0.0227	0.0195	0.0236	0.0205	0.0264	0.0229
	(0.0111)**	(0.0112)*	(0.0130)*	(0.0131)	(0.0119)**	(0.0120)*	(0.0142)*	(0.0144)
Advanced internet and High income,		0.2220		0.5847		0.5243		0.3163
education, IT-intensity, and population county		(0.0907)**		(0.1792)***		(0.1627)***		(0.1355)**
Observations	2743	2743	2743	2743	2743	2743	2743	2743
R^2	0.13	0.14	0.13	0.14	0.13	0.14	0.13	0.14

Dependent variable is logged wages. . In columns (1), (3), (5), and (7) controls are the same as table 2. In columns (2), (4), (6), and (8) controls are the same as table 4. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

Online Appendix Table 4: Robustness to multiple imputation of missing data

	(1)	*(2)*
Advanced internet	0.0243 (0.0143)*	0.0239 (0.0127)*
Advanced internet and High income, education, IT-intensity, and population county		0.4587 (0.1585)***
Observations	3133	2743
F statistic	12.43	15.49

Dependent variable is logged wages. In column (1) controls are the same as table 2. In column (2) controls are the same as table 4.
Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

Appendix Table 5A: Full set of coefficients for Table 5 (First stage)

	(1)	*(2)*	*(3)*	*(4)*
Instrument→	Programmers in other establishments within the same firm	Bartik index	Both main instruments	Two main instruments plus ARPANET nodes and year when state adopted a price cap
Programmers in other establishments within the same firm	0.0002		0.0002	0.0002
	(0.0000)***		(0.0001)***	(0.0001)***
Bartik index		0.2990	0.2584	0.2554
		(0.1774)*	(0.1788)	(0.1791)
ARPANET nodes				0.0055
				(0.0046)
Year when state adopted a price cap				0.0013
				(0.0008)
Average number of establishments in Harte Hanks firms	-0.0252	-0.0116	-0.0260	-0.0264
	(0.0141)*	(0.0141)	(0.0141)*	(0.0141)*
Home internet use	0.0350	0.0345	0.0319	0.0355
	(0.0325)	(0.0327)	(0.0327)	(0.0326)
Home internet data missing	0.0190	0.0191	0.0176	0.0178
	(0.0150)	(0.0151)	(0.0151)	(0.0151)
Log population in 1990	0.0097	0.0107	0.0009	0.0092
	(0.0033)***	(0.0033)***	(0.0033)***	(0.0033)***
Percentage African Americans in 1990	0.0266	-0.0325	-0.0356	-0.0348
	(0.0153)	(0.0267)	(0.0266)	(0.0267)
Percentage university graduates in 1990	0.5340	0.1010	0.0752	0.0620
	(0.0864)	(0.1223)	(0.1225)	(0.1233)
Percentage high school graduates in 1990	-0.1865	0.0624	0.0626	0.0600
	(0.0670)	(0.1240)	(0.1234)	(0.1235)
Percentage below poverty line in 1990	-0.1874	0.0559	0.0513	0.0526
	(0.0560)	(0.0996)	(0.0988)	(0.0990)
Median income in 1990 ($000)	-0.0003	-0.0004	-0.0003	-0.0002
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Percentage population attending Carnegie Type 1 schools in 1990	0.0324	0.0315	0.0327	0.0323
	(0.0594)	(0.0578)	(0.0593)	(0.0597)
Percentage population enrolled in engineering program in 1990	-0.3791	-0.4261	-0.3662	-0.3242
	(0.5477)	(0.5531)	(0.5466)	(0.5497)
# of patents granted to inventors in the county in the 1980s (000)	0.0004	0.0035	0.0033	0.0021
	(0.0026)	(0.0027)	(0.0026)	(0.0027)
Percentage professional in 1995	-0.0615	-0.0438	-0.0457	-0.0411
	(0.0870)	(0.0873)	(0.0874)	(0.0877)
Percentage of persons over age 65 in 1990	0.0387	0.0527	0.0458	0.0656
	(0.0927)	(0.0933)	(0.0926)	(0.0935)
Net migration into the county in 1995 (000)	-0.0018	-0.0019	-0.0018	-0.0015
	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Change in log total population between 1990 and 2000	-0.0057	-0.0058	-0.0015	-0.0040
	(0.0266)	(0.0268)	(0.0265)	(0.0265)
Change in percentage of African American 1990 to 2000	0.2834	0.2829	0.2879	0.2889
	(0.1465)*	(0.1465)*	(0.1464)**	(0.1474)*
Change in percentage of university graduates 1990 to 2000	-0.1989	-0.2016	-0.2126	-0.2141
	(0.2781)	(0.2781)	(0.2888)	(0.2780)
Change in percentage of high school graduates 1990 to 2000	0.0885	0.0719	0.0798	0.0737
	(0.2190)	(0.2197)	(0.2187)	(0.2190)
Change in percentage of persons over age 65 1990 to 2000	-0.0895	-0.0407	-0.0596	-0.0526
	(0.2261)	(0.2292)	(0.2277)	(0.2279)
Change in net migration into the county 1990 to 2000 (000)	-0.0023	-0.0024	-0.0022	-0.0019
	(0.0017)	(0.0017)	(0.0016)	(0.0016)
Constant	-0.0501	-0.1081	-0.0812	-0.2098
	(0.0804)	(0.0826)	(0.0829)	(0.1208)*
Partial R-squared	0.0082	0.0022	0.0098	0.0106
F-Statistic	14.16	2.84	8.39	5.26
Observations	2743	2743	2743	2743
R-squared	0.0282	0.0223	0.0298	0.0217

Heteroskedasticity-robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 5A: Full set of coefficients for Table 5 (Second stage)

	(1)	*(2)*	*(3)*	*(4)*
Instrument→	Programmers in other establishments within the same firm	Bartik index	Both main instruments	Two main instruments plus ARPANET nodes and year when state adopted a price cap
Advanced internet	0.3691 (0.1450)**	0.0156 (0.2374)	0.3174 (0.1356)**	0.4014 (0.1549)***
Home internet use	0.0730 (0.0380)*	0.0865 (0.0387)**	0.0750 (0.0378)**	0.0718 (0.0382)*
Home internet data missing	0.0232 (0.0174)	0.0305 (0.0175)*	0.0242 (0.0173)	0.0225 (0.0175)
Average number of establishments in Harte Hanks firms	-0.0159 (0.0080)**	-0.0195 (0.0067)***	-0.0164 (0.0076)**	-0.0156 (0.0083)*
Log population in 1990	-0.0096 (0.0029)***	-0.0055 (0.0031)*	-0.0090 (0.0027)***	-0.0100 (0.0030)***
Percentage African Americans in 1990	0.0266 (0.0153)*	0.0158 (0.0135)	0.0250 (0.0144)*	0.0275 (0.0158)*
Percentage university graduates in 1990	0.5340 (0.0864)***	0.5767 (0.0835)***	0.5402 (0.0837)***	0.5301 (0.0879)***
Percentage high school graduates in 1990	-0.1865 (0.0670)***	-0.1661 (0.0537)***	-0.1835 (0.0631)***	-0.1884 (0.0694)***
Percentage below poverty line in 1990	-0.1873 (0.0560)***	-0.1734 (0.0468)***	-0.1853 (0.0533)***	-0.1886 (0.0576)***
Median income in 1990 ($000)	-0.0005 (0.0007)	-0.0007 (0.0006)	-0.0005 (0.0007)	-0.0005 (0.0007)
Percentage population attending Carnegie Type 1 schools in 1990	0.0183 (0.0336)	0.0293 (0.0490)	0.0199 (0.0354)	0.0173 (0.0326)
Percentage population enrolled in engineering program in 1990	-0.0447 (0.3243)	-0.2015 (0.3832)	-0.0676 (0.3225)	-0.0304 (0.3269)
# of patents granted to inventors in the county in the 1980s (000)	0.0149 (0.0042)***	0.0162 (0.0044)***	0.0151 (0.0042)***	0.0148 (0.0042)***
Percentage professional in 1995	0.0052 (0.0620)	-0.0168 (0.0551)	0.0020 (0.0597)	0.0072 (0.0631)
Percentage of persons over age 65 in 1990	0.0155 (0.0582)	0.0313 (0.0515)	0.0178 (0.0559)	0.0141 (0.0595)
Net migration into the county in 1995 (000)	0.0039 (0.0032)	0.0032 (0.0032)	0.0038 (0.0032)	0.0040 (0.0032)
Change in log total population between 1990 and 2000	0.0600 (0.0176)***	0.0562 (0.0153)***	0.0595 (0.0170)***	0.0604 (0.0181)***
Change in percentage of African American 1990 to 2000	-0.0648 (0.0996)	0.0333 (0.0993)	-0.0504 (0.0958)	-0.0737 (0.1046)
Change in percentage of university graduates 1990 to 2000	0.8811 (0.1902)***	0.8157 (0.1665)***	0.8715 (0.1832)***	0.8871 (0.1960)***
Change in percentage of high school graduates 1990 to 2000	-0.0643 (0.1196)	-0.0354 (0.0960)	-0.0601 (0.1129)	-0.0670 (0.1241)
Change in percentage of persons over age 65 1990 to 2000	-0.5470 (0.1411)***	-0.5735 (0.1198)***	-0.5508 (0.1353)***	-0.5446 (0.1454)***
Change in net migration into the county 1990 to 2000 (000)	0.0028 (0.0037)	0.0019 (0.0038)	0.0027 (0.0037)	0.0029 (0.0037)
Constant	0.3330 (0.0550)***	0.3069 (0.0478)***	0.3292 (0.0526)***	0.3354 (0.0560)***
Overidentification test (p-value)	N/A	N/A	0.2110	0.2229
Hausman test (p-value)	0.9997	1.0000	0.9999	0.9982
Observations	2743	2743	2743	2743
R-squared	0.131	0.133	0.133	0.133

Dependent variable is logged wages. Heteroskedasticity-robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 6: Other instrumental variables applied to table 2 column 2

	(1)	*(2)*	*(3)*	*(4)*
Instrument→	ARPANET nodes		Year when state adopted a price cap	
	First stage	Second stage	First stage	Second stage
Advanced internet		2.4859		0.7180
		(2.1301)		(0.5994)
ARPANET nodes	0.0058			
	(0.0048)			
Year when state adopted a price cap			0.0011	
			(0.0008)	
Home internet use	0.0380	-0.0078	0.0417	0.0597
	(0.0323)	(0.1065)	(0.0325)	(0.0476)
Home internet data missing	0.0207	-0.0206	0.0207	0.0159
	(0.0149)	(0.0553)	(0.0149)	(0.0228)
Average number of establishments in Harte Hanks firms	-0.0101	0.0055	-0.0104	-0.0124
	(0.0140)	(0.0403)	(0.0140)	(0.0129)
Log population in 1990	0.0116	-0.0342	0.0118	-0.0137
	(0.0031)***	(0.0258)	(0.0031)***	(0.0075)*
Percentage African Americans in 1990	-0.0304	0.0909	-0.0295	0.0372
	(0.0267)	(0.0920)	(0.0268)	(0.0286)
Percentage university graduates in 1990	0.1178	0.2777	0.1114	0.4917
	(0.1229)	(0.4091)	(0.1231)	(0.1313)***
Percentage high school graduates in 1990	0.0586	-0.3090	0.0547	-0.2067
	(0.1240)	(0.3308)	(0.1240)	(0.1038)**
Percentage below poverty line in 1990	0.0400	-0.2710	0.0405	-0.2011
	(0.0990)	(0.2583)	(0.0990)	(0.0838)**
Median income in 1990 ($000)	-0.0003	0.0003	-0.0002	-0.0004
	(0.0009)	(0.0026)	(0.0009)	(0.0009)
Percentage population attending Carnegie Type 1 schools in 1990	0.0308	-0.0474	0.0308	0.0075
	(0.0579)	(0.1290)	(0.0583)	(0.0322)
Percentage population enrolled in engineering program in 1990	-0.4383	0.8941	-0.4088	0.1101
	(0.5553)	(1.5685)	(0.5564)	(0.4595)
# of patents granted to inventors in the county in the 1980s (000)	0.0029	0.0069	0.0034	0.0136
	(0.0027)	(0.0113)	(0.0026)	(0.0049)***
Percentage professional in 1995	-0.0619	0.1367	-0.0579	0.0269
	(0.0869)	(0.2568)	(0.0871)	(0.0877)
Percentage of persons over age 65 in 1990	0.0452	-0.0791	0.0629	-0.0001
	(0.0934)	(0.2458)	(0.0941)	(0.0830)
Net migration into the county in 1995 (000)	-0.0020	0.0080	-0.0015	0.0046
	(0.0015)	(0.0060)	(0.0015)	(0.0034)
Change in log total population between 1990 and 2000	-0.0105	0.0828	-0.0134	0.0638
	(0.0268)	(0.0725)	(0.0269)	(0.0243)***
Change in percentage of African American 1990 to 2000	0.2819	-0.6522	0.2746	-0.1616
	(0.1481)*	(0.6487)	(0.1468)*	(0.2168)
Change in percentage of university graduates 1990 to 2000	-0.1869	1.2731	-0.1852	0.9457
	(0.2786)	(0.7923)	(0.2786)	(0.2771)***
Change in percentage of high school graduates 1990 to 2000	0.0820	-0.2375	0.0757	-0.0929
	(0.2200)	(0.5711)	(0.2201)	(0.1832)
Change in percentage of persons over age 65 1990 to 2000	-0.0750	-0.3884	-0.0680	-0.5208
	(0.2278)	(0.5775)	(0.2279)	(0.1976)***
Change in net migration into the county 1990 to 2000 (000)	-0.0027	0.0080	-0.0020	0.0037
	(0.0016)	(0.0072)	(0.0016)	(0.0040)
Constant	-0.0748	0.4890	-0.1939	0.3587
	(0.0804)	(0.2503)*	(0.1161)*	(0.0833)***
Partial R-squared	0.0001		0.0007	
F-Statistic	1.4690		2.0315	
Hausman Test	1.0000		1.0000	
Observations	2743	2743	2743	2743
R-squared	0.0203	0.13	0.0209	0.13

Dependent variable is logged wages. Heteroskedasticity-robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 7: Other instrumental variables applied to Table 4 column 6

	(1)	(2)
	ARPANET nodes	Year when state adopted a price cap
FIRST STAGE—Advanced internet		
ARPANET nodes	0.0116 (0.0172)	
Year when state adopted a price cap		0.0011 (0.0008)
ARPANET nodes and high all factors	-0.0077 (0.0174)	
Year when state adopted a price cap and high all factors		0.0004 (0.0012)
Partial R-squared	0.0001	0.0007
F-Statistic	10.42	10.03
FIRST STAGE—Advanced internet and high all factors		
ARPANET nodes	-0.0011 (0.0008)	
Year when state adopted a price cap		-0.00005 (0.00002)***
ARPANET nodes and high all factors	0.0072 (0.0032)**	
Year when state adopted a price cap and high all factors		0.0016 (0.0008)*
Partial R-squared	0.0098	0.0195
F-Statistic	171.35	265.90
SECOND STAGE		
Advanced internet	-0.0724 (0.7287)	0.7416 (0.6297)
Advanced internet and High income, education, IT-intensity, and population county	3.0077 (1.2898)**	0.3859 (1.0988)
Overidentification test (p-value)		
Hausman test (p-value)	0.9415	1.0000
Observations	2743	2743
R^2	0.04	0.04

Dependent variable is logged wages. Controls same as table 4. Heteroskedasticity-robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 8: Full set of coefficients for Table 4

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Advanced internet	0.0168	0.0120	0.0214	0.0246	0.0049	0.0239	0.0067
	(0.0137)	(0.0125)	(0.0159)	(0.0127)*	(0.0149)	(0.0128)*	(0.0150)
Advanced internet and	0.0960				0.0442		0.0377
High income county	(0.0389)**				(0.0492)		(0.0496)
Advanced internet and		0.1101			0.0770		0.0757
High education county		(0.0455)**			(0.0548)		(0.0547)
Advanced internet and			0.0206		0.0134		0.0102
High IT-intensity county			(0.0228)		(0.0235)		(0.0237)
Advanced internet and				0.3631	0.2378		0.0182
High population county				(0.0934)***	(0.1018)**		(0.1027)
Advanced internet and High income,						0.4588	0.3393
education, IT-intensity, & population county						(0.1585)***	(0.1904)*
High income county	-0.0136				-0.0093	-0.0067	-0.0102
	(0.0062)**				(0.0066)	(0.0050)	(0.0067)
High education county		-0.0158			-0.0134	-0.0067	-0.0132
		(0.0061)***			(0.0066)**	(0.0048)	(0.0066)**
High IT-intensity county			0.0101		0.0086	0.0094	0.0083
			(0.0039)***		(0.0040)**	(0.0034)***	(0.0041)**
High population county				-0.0231	-0.0119	0.0078	0.0058
				(0.0117)**	(0.0125)	(0.0061)	(0.0121)
High income, education, IT-intensity, and						-0.0370	-0.0250
population county						(0.0214)*	(0.0239)
Log population in 1990	-0.0062	-0.0060	-0.0068	-0.0084	-0.0083	-0.0080	-0.0080
	(0.0017)***	(0.0018)***	(0.0017)***	(0.0022)***	(0.0022)***	(0.0022)***	(0.0022)***
Percentage African Americans in 1990	0.0128	0.0122	0.0125	0.0100	0.0098	0.0098	0.0100
	(0.0118)	(0.0118)	(0.0118)	(0.0118)	(0.0118)	(0.0118)	(0.0118)
Percentage university graduates in 1990	0.5712	0.6022	0.5693	0.5576	0.5855	0.5821	0.5774
	(0.0784)***	(0.0845)***	(0.0787)***	(0.0785)***	(0.0848)***	(0.0855)***	(0.0848)***
Percentage high school graduates in 1990	-0.1491	-0.1537	-0.1466	-0.1407	-0.1351	-0.1301	-0.1309
	(0.0519)***	(0.0517)***	(0.0514)***	(0.0513)***	(0.0509)***	(0.0513)**	(0.0510)**
Percentage below poverty line in 1990	-0.1449	-0.1483	-0.1398	-0.1490	-0.1326	-0.1327	-0.1309
	(0.0473)***	(0.0462)***	(0.0460)***	(0.0465)***	(0.0471)***	(0.0470)***	(0.0470)***
Median income in 1990 ($000)	-0.0000	-0.0003	-0.0003	-0.0004	-0.0000	-0.0000	0.0000
	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0007)
Percentage population attending Carnegie	0.0311	0.0237	0.0297	0.0310	0.0251	0.0289	0.0257
Type 1 schools in 1990	(0.0478)	(0.0427)	(0.0469)	(0.0480)	(0.0448)	(0.0475)	(0.0450)
Percentage population enrolled in	-0.1939	-0.1418	-0.1582	-0.1318	-0.0542	-0.0884	-0.0549
engineering program in 1990	(0.3652)	(0.3390)	(0.3629)	(0.3612)	(0.3440)	(0.3595)	(0.3436)
# of patents granted to inventors in the	0.0161	0.0161	0.0160	0.0130	0.0121	0.0104	0.0102
county in the 1980s (000)	(0.0043)***	(0.0043)***	(0.0044)***	(0.0041)***	(0.0040)***	(0.0040)***	(0.0041)**
Percentage professional in 1995	-0.0118	-0.0112	-0.0322	-0.0244	-0.0427	-0.0443	-0.0431

	(0.0535)	(0.0533)	(0.0525)	(0.0538)	(0.0529)	(0.0532)	(0.0530)
Percentage of persons over age 65 in 1990	0.0564	0.0605	0.0577	0.0566	0.0549	0.0507	0.0536
	(0.0508)	(0.0508)	(0.0513)	(0.0515)	(0.0510)	(0.0507)	(0.0506)
Net migration into the county in 1995	0.0038	0.0037	0.0037	0.0040	0.0040	0.0037	0.0038
(000)	(0.0031)	(0.0032)	(0.0031)	(0.0030)	(0.0030)	(0.0029)	(0.0029)
Change in log total population between	0.0530	0.0531	0.0589	0.0570	0.0636	0.0647	0.0647
1990 and 2000	(0.0151)***	(0.0152)***	(0.0158)***	(0.0152)***	(0.0156)***	(0.0154)***	(0.0154)***
Change in percentage of African American	0.0208	0.0275	0.0068	0.0048	-0.0051	0.0001	-0.0014
1990 to 2000	(0.0736)	(0.0738)	(0.0746)	(0.0734)	(0.0736)	(0.0739)	(0.0738)
Change in percentage of university graduates	0.8301	0.8466	0.8044	0.8190	0.8400	0.8228	0.8394
1990 to 2000	(0.1618)***	(0.1598)***	(0.1596)***	(0.1590)***	(0.1594)***	(0.1601)***	(0.1592)***
Change in percentage of high school	-0.0106	-0.0198	-0.0154	-0.0045	-0.0025	0.0006	-0.0017
graduates 1990 to 2000	(0.0941)	(0.0937)	(0.0939)	(0.0936)	(0.0928)	(0.0934)	(0.0928)
Change in percentage of persons over age	-0.5471	-0.5380	-0.5198	-0.5210	-0.4755	-0.4881	-0.4745
65 1990 to 2000	(0.1191)***	(0.1196)***	(0.1202)***	(0.1198)***	(0.1203)***	(0.1205)***	(0.1202)***
Change in net migration into the county	0.0025	0.0024	0.0023	0.0030	0.0031	0.0030	0.0031
1990 to 2000 (000)	(0.0037)	(0.0037)	(0.0037)	(0.0035)	(0.0035)	(0.0034)	(0.0034)
Constant	0.3075	0.3107	0.3190	0.3372	0.3270	0.3205	0.3216
	(0.0407)***	(0.0394)***	(0.0390)***	(0.0428)***	(0.0445)***	(0.0447)***	(0.0447)***
Observations	2743	2743	2743	2743	2743	2743	2743
R-squared	0.13	0.13	0.13	0.13	0.14	0.14	0.14

Dependent variable is logged wages. Heteroskedasticity-robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 9: 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	-0.0118 (0.0300)	-0.0008 (0.0309)
Advanced internet	0.0354 (0.0293)	0.0075 (0.0314)
Advanced internet and High income county		0.0377 (0.0497)
Advanced internet and High education county		0.0757 (0.0548)
Advanced internet and High population county		0.0178 (0.1045)
Advanced internet and High IT-intensity county		0.0101 (0.0236)
Advanced internet and High income, education, IT-intensity, and population county	0.4478 (0.1606)***	0.3390 (0.1903)*
Observations	2743	2743
R^2	0.14	0.14

Dependent variable is logged wages. Controls are the same as in Table 4.
Heteroskedasticity-robust standard errors in parentheses.
+ significant at 10%; * significant at 5%; ** significant at 1%.

Appendix Table 10: Coefficients used to generate Figure 2

	Advanced internet	Advanced internet and HighAllFactors
Interacted with 1991	0.0014	-0.0226
	(0.0068)	(0.0263)
Interacted with 1992	0.0134	0.0157
	(0.0089)	(0.0257)
Interacted with 1993	0.0108	0.0050
	(0.0099)	(0.0246)
Interacted with 1994	0.0108	-0.0007
	(0.0098)	(0.0255)
Interacted with 1995	0.0162	0.0147
	(0.0115)	(0.0278)
Interacted with 1996	0.0178	0.0677
	(0.0114)	(0.0312)**
Interacted with 1997	0.0289	0.1270
	(0.0131)**	(0.0386)***
Interacted with 1998	0.0218	0.1521
	(0.0128)*	(0.0373)***
Interacted with 1999	0.0313	0.1777
	(0.0163)*	(0.0441)***
Interacted with 2000	0.0418	0.2656
	(0.0144)***	(0.0587)***
Observations	30173	
R^2	0.87	

Dependent variable is logged wages. This table presents the results of one panel regression of 11 years (1990-2000) and 2743 counties. Base year in 1990. Controls are the same as in Table 4, but separately interacted with each year (e.g. 1991×MedianIncome1990, 1992×MedianIncome1990, 1993×MedianIncome1990, etc.). Year dummies also included. County fixed effects are differenced out. Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%.

Appendix Table 11: Advanced internet and number of establishments

	NUMBER OF ESTABLISHMENTS	
	(1)	*(2)*
	Full set of controls	High all factors interaction and full set of controls
Advanced internet	-0.0023	-0.0033
	(0.0123)	(0.0124)
Advanced internet and High income, education, IT-intensity, population county		0.0637
		(0.1248)
Observations	2743	2743
R^2	0.46	0.46

In columns (1) and (3) controls are the same as table 2. In columns (2) and (4) controls are the same as table 4.
Heteroskedasticity-robust standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%.