WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 09-17
TECHNOLOGICAL ADAPTATION, CITIES, AND NEW WORK

Jeffrey Lin
Federal Reserve Bank of Philadelphia

July 28, 2009

Supersedes Working Paper 07-25

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Technological Adaptation, Cities, and New Work*

Jeffrey Lin†

This version: July 28, 2009

Abstract

Where does adaptation to innovation take place? I present evidence on the role of agglomeration economies in the application of new knowledge to production. All else equal, workers are more likely to be observed in *new work* in locations that are initially dense in both college graduates and industry variety. This pattern is consistent with economies of density from the geographic concentration of factors and markets related to technological adaptation. A main contribution is to use a new measure, based on revisions to occupation classifications, to closely characterize cross-sectional differences across U.S. cities in adaptation to technological change. Worker-level results also provide new evidence on the skill bias of recent innovations.

Keywords: innovation, agglomeration economies, occupations, human capital, industrial diversity.
JEL codes: J21, J24, O33, R12, R23

*I thank Gordon Hanson and Hoyt Bleakley for extensive advice, as well as David Autor, Eli Berman, Julian Betts, Roger Bohn, Steve Erie, Roger Gordon, Bob Hunt, Leslie McGranahan, Leonard Nakamura, Jim Rauch, Albert Saiz, Ian Savage, Kei-Mu Yi, the editor, Michael Greenstone, and the anonymous referees for helpful comments, Dana Bradley and Marisa Tegler at the U.S. Census Bureau for documentation on new occupations, Cristine McCollum at the FRBP library for help with the *DOT*, and Patrick Ho and Evan Taylor for excellent research assistance.

†Economist, Research Department, Federal Reserve Bank of Philadelphia. Post: Ten Independence Mall, Philadelphia, PA, 19106. E-mail: jeff dot lin at phil dot frb dot org. Web: http://jlin.org/. The views expressed here are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at http://www.philadelphiafed.org/research-and-data/publications/working-papers/.

1 Introduction

Adaptation to technology is important for our understanding of economic growth and income differences across people, firms, and locations. For example, additions to the set of non-rivalrous "recipes" for combining raw inputs into useful product are central to endogenous growth models (Romer, 1990), and it is the application of new knowledge in ways that reshape production that is critical to the history of long-run economic growth (Mokyr, 1992).[1] A key goal, then, is identifying what determines the adaptation of production to new technologies.

There is a large literature relating human capital to adaptation. Knowledgeable people can more quickly adapt their activities to the changing incentives that result from the appearance of new technologies.[2] Further, because of the non-rival nature of technology, an important question is whether there are external benefits from human capital to adaptation. As one illustration of a potential externality, firms might implement novel techniques of production after having observed or combined other preexisting, locally available techniques, as in Jacobs (1969). In general, agglomeration economies from the geographic concentration of economic activity may help explain differences in adaptation to technology across locations. The basic idea is that, in the presence of scale economies and transport costs, workers and firms may not fully internalize the net local benefits to the production of new varieties of goods or the use of new production activities when deciding where to locate factors, production, or consumption.[3] This is a central reason why Lucas (1988) proposes cities as a natural setting for the "engine of growth": by concentrating knowledge and skills with production and consumption, cities provide better access to both factors and markets for the adaptation to new technologies.

In this paper, I investigate the role of agglomeration economies in adaptation to new technologies. The starting point is the observation that certain locations appear better at attracting *new work*.[4] By new work, I mean jobs requiring new combinations of activities or techniques that have emerged in the labor market in response to the application of new information, technologies, or "recipes" to production. These new activities follow innovation, but unlike other measures of innovation inputs or output, new work captures the appearance and implementation of new knowledge and subsequent changes to the organization of production

[1] Lags in technological adaptation can also help to explain income differences across firms (Griliches, 1957), or countries (Parente and Prescott, 1994, and Comin, Hobijn, and Rovito, 2008). I sometimes use adaptation and adoption interchangeably, but to be clear, by adoption, I mean to refer to the decision to implement knowledge into production; by adaptation, I refer more broadly to changes in the organization of production that accompany this implementation. Thus new work is part of adaptation but follows adoption.

[2] See Nelson and Phelps (1966) and Schultz (1975). Benhabib and Spiegel (2005) link human capital to total factor productivity growth, with adaptation to technology being a key mechanism.

[3] I use agglomeration economies in a broad sense: scale economies can come from any source, whether sharing, matching, or learning, and there can be costs in transporting people, goods, or ideas. The list of possible mechanisms includes classic descriptions by Marshall (1920) of input sharing, labor pooling, and knowledge spillovers, as well as pecuniary externalities in the presence of increasing returns, working through goods or factor prices, as in Krugman (1991b). Two recent surveys, by Duranton and Puga (2004) and Rosenthal and Strange (2004), provide excellent introductions to modern theory and evidence of various sources of agglomeration economies.

[4] I borrow the phrase "new work" from Jacobs (1969).

and labor markets. Using new work, I can more closely observe how workers and firms adapt to technological change, and I can more systematically characterize the impact of technological change on the structure of production, revealing information that may not be contained in other measures (e.g., patents, research and development spending, case studies of specific technologies, or estimates of total factor productivity).

To identify new work in data, I use occupation classifications, which describe the organization of tasks and activities across jobs in the economy. The observation at the heart of this paper is that new knowledge, once applied to production, requires novel activities or combinations of activities. Thus, changes to the census occupation classification system, the official catalog of activities, contain information about the extent of adaptation to technological change. (For example, the classification of data-coder operators appears earliest, followed by microcomputer support specialists, and then web administrators.) I collect new *occupation titles* from U.S. classification indexes in 1977, 1991, and 2000. Using U.S. census microdata from 1980, 1990, and 2000, I estimate worker selection into new occupations as a function of worker and initial local characteristics. I am then able to compare new work across a panel of U.S. cities.

The main evidence in favor of the role of agglomeration economies in adaptation is the result that, all else equal, workers in new occupations are more likely to be observed in locations that are initially dense in both college graduates and industry variety. I argue that this pattern is consistent with economies of density from the geographic concentration of factors that are complements to new technologies (e.g., skilled labor) or markets for goods or services that use new work (educated consumers or industry variety). The estimation results are not sensitive to controls for local industry or worker composition, nor are they due primarily to worker sorting across locations, the presence of fixed factors, or unobserved changes in local characteristics. The results are also robust to alternative measures of new work and the main explanatory variables.

Also, I find that more-educated workers are more likely to be observed in new work and that the educational attainment of new work has risen since 1980—consistent with the hypothesis of skill-biased technological change. Workers in new work also earn higher wages than observationally similar workers in preexisting work.

In using classification data, this paper is similar to some other recent papers that study technological change. For example, in order to examine the complementarity between technology and skill, Autor, Levy, and Murnane (2003) use changes in the task content of detailed occupations, and Xiang (2005) uses new products in industry classifications.[5] Their aims are to relate the skill bias of technology to shifts in labor demand, whereas I try to relate factors and adaptation. In addition, as a measure of broad changes in the organization of production, new work contains information that cannot be captured by changes in task content, which are available only for continuing occupation classifications, or by new products, which may not as fully cover innovations in processes or services. There is also a literature that identifies external

[5]Alexopoulos (2006) introduces a measure of technological change based on new books published in the field of technology. Relative to this measure, new work may be more directly tied to the changes in the labor market.

effects of human capital in cities, on wages (e.g., Rauch, 1993, and Glaeser and Saiz, 2003), or patents (Jaffe, Trajtenberg, and Henderson, 1993, and Carlino, Chatterjee, and Hunt, 2007). Relative to this literature, new work might provide further insight into the link between concentrations of human capital and rates of invention, the application of inventions to production, and productivity differences across locations.

Finally, there is a literature concerned with the role of factor supply in innovation (e.g., Acemoglu, 2009). In a related paper, Beaudry, Doms, and Lewis (2006) report that the ratio of personal computers to workers varies across U.S. cities, a pattern that they attribute to differences in relative factor supplies. In contrast, I emphasize that the geographic distribution of new work is partly an outcome of agglomeration economies, which can operate through the concentration of either factors *or* markets for new technologies. The main contribution of this paper is to use a new measure, based on new occupations, in order to better understand cross-sectional variation, across locations, in the ability to adapt to technological change.

The rest of the paper proceeds as follows: section 2 outlines the process of collecting data on new occupation titles and matching these titles to census microdata. I also compare new work to other cross-sectional measures of innovation output. In section 3, I discuss, informally, the results of a model in which the location of new work is explained by agglomeration economies. This model then informs the estimation described in section 4 and the interpretation of the results presented in section 5.

2 Data

Over time, changes in the occupation classification system form an important—if accidental—record of the changing organization of work in the United States. In this section, I describe the process of collecting information on new work. First, I identify new *occupation titles*, related to the emergence of new work, by comparing successive lists of titles and using documentation from U.S. statistical agencies. These title lists contain information on new occupation classifications related to new work, but they do not contain employment data. The second step therefore matches these new title lists to three cross-sectional censuses, which contain information on workers' occupations and locations. Using census microdata, I obtain information about new work employment, but I only observe occupation at a (three-digit) level of aggregation that is coarser than the underlying (five- or nine-digit) title information. I provide details on this matching process and try to assess the extent of possible bias from aggregating occupation titles.

2.1 Comparing lists of occupation titles

Each decennial census uses an occupation classification system to catalog the various types of work in the U.S. economy. This system is updated periodically, to reflect both the changing nature of work and the changing needs of data users, relying on previous versions of the classification system, field research, the *Dictionary* *of Occupational Titles*, and written descriptions from census respondents of the type of work they perform.[6] These reviews help ensure that jobs reflecting new bundles of tasks and activities are consistently captured by the classification system after they appear in the U.S. economy.

An *occupation title*—the atomistic unit of the classification system—describes a small number of individual jobs that require the use of a similar set of activities or techniques. This narrow scope means there are a large number of titles: between 1950 and 2000, the number of titles in the Census Bureau's *Classified Index of Industries and Occupations* expanded from about 25 to 31 thousand. Any broad category of work consists of hundreds of occupation titles. For example, in the 2000 *Classified Index*, there are over 500 titles that contain the word engineer, including at least nine computer engineering occupations (e.g., computer software applications engineer) and eleven aerospace engineering occupations (flight test engineer). There are also at least twenty varieties of economists, with descriptions spanning specialty (econometrics, finance, labor, trade), and function (teacher, research assistant, policy advisor).

The classification system assigns each occupation title to a unique census *detailed occupation*. (Titles use five digits in the census *Classified Index* and nine digits in the *DOT*, whereas census detailed occupations use three digits.) Unlike titles, detailed occupations are reported in census microdata; therefore, this assignment of titles to detailed occupations is important for matching title information to employment totals (a procedure that is described next). Each detailed occupation groups together titles according to the similarity of work performed and skills required. (In the 2000 census, the median number of titles in each detailed occupation is 33. For example, detailed occupation 110, *network and computer system administrators*, contains 30 occupation titles.)

In order to collect information on the changing nature of work, I use three classification revisions involving five title catalogs. The first comparison is between the *DOT*'s third (1965) and fourth (1977) editions, the second comparison is between the *DOT*'s fourth (1977) and revised fourth (1991) editions, and the third comparison is between the census *Classified Indexes* from 1990 and 2000.

I choose the 1965-1977 and 1977-1991 *DOT* revisions and the 1990-2000 census *Classified Index* revision because of the availability of both machine-readable title lists and supplemental documents on the sources of individual title changes. In particular, these supplemental documents are critical in order to rule out title additions that are unrelated to new work. For example, growing interest in a specific sector may increase the likelihood of non-relevant revisions: employment growth between 1960 and 1970 led the census to separate lawyers and judges into two separate detailed occupations, and the 2000 revision featured the creation of "job families" and thus shifts in large numbers of titles across detailed occupations, confounding the identification of new detailed occupations.

[6]Technical papers from the U.S. Census Bureau (Scopp, 2003) and the Bureau of Labor Statistics (Meyer and Osborne, 2005) provide overviews of this process. The *DOT* is a separately produced classification system from the U.S. Department of Labor. Its successor is known as the *Occupational Information Network* or O*NET.

Crucially, this problem is less severe for (five- or nine-digit) titles than it is at the level of (three-digit) detailed occupations. A census technical paper notes that titles, unlike detailed occupations, "provide information about the intended, or 'ideal' changes from each [...] occupation code of one classification into each [...] occupation code of the other classification" (Scopp, 2003, p. 9). In other words, detailed occupations may be combined or split apart according to the needs of the census or of a growing population, but titles remain anchored to a small number of jobs that use a similar set of tasks or activities. I rely on Scopp (2003) and other census documents to isolate genuinely new titles, reflecting actual new work, in the 2000 *Classified Index*. Similarly, two *Conversion Tables* (U.S. D.O.L. and U.S. E.S., 1979, and U.S. D.O.L. et al., 1991) are available that separate the 1977 and 1991 *DOT* into lists of titles re-coded or renamed and lists of actual new title definitions. For example, based on the 1991 *Conversion Table*, I am able to eliminate 310 seemingly new titles that are in fact re-coded or renamed 1977 titles.

I have relegated remaining details about the construction of new title lists to the data appendix. The supplemental documents separating new titles from other kinds of classification changes are clearest for the 1977-1991 *DOT* revision. (In contrast, matching the new title lists to three-digit detailed occupations, described next, is most straightforward in the case of the 1990-2000 *Classified Index* revision.) Finally, for two revisions (1977-1991 and 1990-2000), I create alternate new title lists, using somewhat independent algorithms, in order to check the robustness of the results to the method of identifying new titles.

The result of the procedures described here is three lists—in 1977, 1991, and 2000—of new occupation titles, where I have used statistical-agency documentation to eliminate classification changes unrelated to the actual appearance of new work. The 1977 list contains 1,152 new titles (out of 12,695 total titles), and the preferred lists in 1991 and 2000 contain 830 (12,741) and 840 (30,900) new titles, respectively. (The alternate lists contain 89 and 814 new titles.) Note that variation in the total number of titles and the proportion of new titles is in part because of source-specific effects, and I control for these effects in the presentation of main results.

2.2 Matching new occupation titles to census microdata

The second step matches the new title lists to worker-level data, and it requires collapsing information from the (five- or nine-digit) title level to the (three-digit) detailed occupation level observed in census microdata. This collapse is based on title counts: define ν as the number of new titles divided by the total number of titles, for each detailed occupation.[7]

Table 1 lists, by classification year, detailed occupations with the highest new title shares, with examples of matched new occupation titles. For example, in 2000, detailed occupation 111, *network systems and data communication analysts*, contains the highest share of new titles: 29 out of 30 ($\nu = 0.967$) titles (e.g., chat

[7] I crosswalk 1977 and 1991 *DOT* titles to census detailed occupations in 1980 and 1990; titles from the 2000 *Classified Index* are already mapped directly to census detailed occupations in 2000. Crosswalk details are in the data appendix.

room hosts) do not match any title from the 1990 *Classified Index*. In 1991 and 2000, detailed occupations related to information technology and medicine have the highest incidence of new titles, while the newest detailed occupations in 1977 reflect engineering activities.

A maintained assumption is that the distribution of new title *counts* and new title *employment* is similar across detailed occupations. There are a number of ways in which employment might be distributed differently than titles across detailed occupations, with different implications for the results. For example, if new titles systematically contain fewer jobs than continuing titles, then I will overstate total employment in new work—though without bias in cross-sectional comparisons of new work employment across locations or industries. More problematically, bias may be serious if a small set of new occupation titles contains much of the actual employment in new work: then, I will overestimate new work employment in locations with concentrations in the corresponding detailed occupations.

I attempt to gauge the extent of possible bias in a number of ways. First, note that the distribution of new titles across detailed occupations is extremely skewed, so that the majority of detailed occupations contain zero new titles, while only a handful of detailed occupations contain many new titles (see Figure 1). Thus, for the large number of detailed occupations with new title shares (ν) near zero and the small number of detailed occupations with ν near one, imputed values are likely to reflect actual new work employment.

Second, I measure new title employment using a special version of the April 1971 Current Population Survey (N.A.S., 1981). Here, workers' occupations are coded with (three-digit) 1960, 1970, 1977 and 1980 census detailed occupations, and (nine-digit) 1965 and 1977 *DOT* titles.[8] There are a number of important caveats in using these data: 11.5% of sampled workers are not assigned 1977 titles (and hence are excluded), and only 3,885 out of the 12,695 total 1977 *DOT* titles are actually observed (thus complicating the calculation of title shares). Also, most new titles have zero or near-zero employment. Of course, employment in new occupation titles is likely less in April 1971 than in 1977, when I collect the new titles in the *DOT*, or in 1980, when I observe workers in census microdata. On the other hand, favorably, none of the future new *DOT* titles from 1991 appear with positive employment in 1971.

I examine two correlations: first, I compare new title count shares of 1980 detailed occupations, computed using all 12,695 titles in the 1977 *DOT*, to new title employment shares of 1980 detailed occupations, computed using the 1971 CPS (Figure 2, left panel). Second, I compare new title count shares of 1977 detailed occupations, computed using only the 3,885 titles that appear in the 1971 CPS, to new title employment shares of 1977 detailed occupations, again computed using the 1971 CPS (Figure 2, right panel). The second comparison thus excludes occupation titles from the 1977 *DOT* that are not reported in the 1971 CPS.

In this exercise, smaller deviations from the fitted line indicate less potential for bias. That is, small

[8]To be precise, the study reports 1977 and 1980 standard occupational classification (SOC) detailed occupations instead of 1980 census detailed occupations; the lists of detailed occupations in the 1970 and 1980 SOC and the 1980 census are virtually identical (see U.S. Census, 1980a). To my knowledge, the extensive multiple-coding of occupation is unique, over this period, to this data set.

deviations indicate that employment is nearly uniform across new occupation titles. Relative differences in the preponderance of new titles across locations, then, should accurately reflect relative differences in new work employment. In both cases, new title count share explains much of the variation in new title employment share across detailed occupations: regression R^2 values are 0.64 and 0.87, respectively. Further, the slope of the fitted line indicates that, within detailed occupations, new occupation titles actually contain more jobs than existing occupation titles in 1971. At least for the 1977 list of new titles, estimates of overall new work employment based on a title imputation may understate true new work employment. (Given the data limitations noted earlier, it may be difficult to generalize from this example.) Taken together, these results suggest that unobserved variation in employment across new occupation titles is unlikely to be a significant source of bias when comparing differences in new work across locations.

As a final robustness check, later I examine the sensitivity of the main results to the number of titles, the new title share, and employment in detailed occupations, and again, I find little evidence that imputing new work employment using title shares produces misleading results.

Upon inspection, newer detailed occupations seem to reflect changes in labor demands that result from actual innovation. Consider again detailed occupation 111 in 2000, *network systems and data communication analysts*. Examples of new occupation titles within this detailed occupation are chat room host/monitor, computer networks consultant, network engineer, Internet developer, and web designer. According to the *Classified Index*, workers in this detailed occupation "analyze, design, test and evaluate network systems, such as local area networks (LAN), wide area networks (WAN), Internet, intranet, and other data communications systems." Clearly, these new jobs are tied to innovations in network and computer technology such as the creation of the first TCP/IP wide-area network in the mid-1980s, the launch of the World Wide Web in 1991, and the development of the first graphical web browser, Mosaic, in 1993. That the classification system catalogs these occupations in 2000 but not in 1990 or 1980 suggests a close, if somewhat delayed, relationship between new work and new occupation classifications.

I match new title shares ν for detailed occupations to 1980, 1990, and 2000 census microdata from the IPUMS (Ruggles et al., 2009). The data contain worker characteristics, such as detailed occupation and location of residence, for a 5% (1980) or 1% (1990 and 2000) sample of the U.S. population. My sample includes all respondents, age 16 to 70, in identified occupations, excluding Alaska and Hawaii. For location of residence, I construct county aggregates that can be consistently and universally identified over the period 1970-2000, using the consistent public-use microdata area and the 1970 county group variables in the IPUMS.[9] In addition, I combine county aggregates if they are part of the same metropolitan (core-based statistical) area (U.S. Census Bureau, 2003), as part of an effort to use locations that correspond to local labor markets. As a result, I can identify worker location, across four IPUMS extracts, in one of 363

[9]This is necessary because I use initial local characteristics to predict the subsequent distribution of new work (e.g., 1970 cities and 1980 workers), and I use the IPUMS to calculate some 1970 city characteristics.

consistent, county-based areas, exhausting the physical area of the contiguous U.S. For brevity, throughout the paper, I refer to these locations as "cities."

2.3 Characteristics of those employed in new work

I estimate that in 1980, 8.5% of workers were employed in work that was not cataloged in 1965; in 1990, 8.2% of workers were employed in work that was not cataloged in 1977; and in 2000, 4.4% of workers were employed in work that was not cataloged in 1990. (Again, this assumes that employment within detailed occupations is equally distributed across titles.) Note that direct time-series comparisons are problematic, due to changes in title sources and comparison methods. However, because I can control for census year effects in regressions, comparisons of successive cross-sectional distributions of new work can remain valid and informative.

Maps in Figure 3 compare the share of employment in new work across U.S. locations, the county-based aggregates described earlier. (Darker shades correspond to higher quantiles in the new work share distribution.) In 1980, locations in the southern and western U.S. are most concentrated in new work; in 1990 and 2000, urban concentrations are more pronounced. For example, in 2000, greater Washington, the San Francisco Bay area, Raleigh-Durham, and Austin had the highest new work shares. Other cross-sectional patterns are displayed in Table 2. In general, women and younger, more educated, and more urbanized workers are more likely to be employed in new work. Professional services and management occupations also have higher new work shares.[10]

In 1980, high school graduates are somewhat more likely to be employed in new work, while in 1990 and 2000, college graduates are much more likely to be employed in new work. The types of new work vary across educational attainment groups—college graduates in new work might be computer software engineers while high school graduates might be telecommunications line installers. While, on average, new workers are more educated, the presence of new work at all levels of the skill distribution highlights the interpretation of new work as reflecting changes in the organization of production in response to new technologies, rather than only the most exclusive and specialized activities requiring cutting-edge skills.

New work can also help to explain variation in wages. In each IPUMS extract, I regress log hourly wage on flexible indicator controls for sex, marital status, race, ethnicity, nativity, education, age, relationship to household head, city, and ν, the new work share of the worker's detailed occupation. Table 3 displays the regression results; reported standard error estimates are robust to clustering on detailed occupation. The estimated wage premium of new work is positive, and, in 1990 and 2000, large and significant–upwards of 30% (logarithmic) over observationally similar workers who are not in new work. One interpretation of the wage

[10] In 1980, the agriculture, construction, and mining sector has a high new work share, which contributes to the spatial pattern observed in Figure 3 and comes from the high new title share of agricultural engineers, observed in Table 1. There may be special concern in the 1980 case that the distribution of new titles does not closely match the distribution of new work employment. Later, I check the sensitivity of the results to excluding the 1980 data.

premium is that new workers are more productive and, at the same time, especially well-suited for these new activities. A second interpretation is that entry into one of these new occupations is inherently risky—who knows if this activity is an effective way to use a new technology?—and the wage premium reflects this risk. Future research may explore different explanations for the wage premium, but for now, I only emphasize that workers employed in new work appear different from other workers across many dimensions.

2.4 New work compared to other measures of innovation and adaptation

Consider new work and patents, a common, and complementary, measure of innovation. Both are outputs related to the invisible production of new knowledge. Patent applications are governed by patentability rules and affected by strategic considerations; new work instead reflects both market acceptance of new knowledge and subsequent effects on production, labor demand, and labor supply. These differences drive their respective strengths: patents are readily available, can identify incremental advances, and track the birth of ideas, while new work is broader in industrial scope, is less sensitive to firm strategy, and measures adaptation to new ideas.

To demonstrate some differences in the informational content of different measures of innovation and adaptation, I compare patents to new work for industries and cities, 1980-2000, in Figure 4.[11] Each panel compares new work employment share in industries or cities, in one census year, to accumulated patent counts in industries or cities over the previous decade. For industries in 1990 and 2000, and cities in 2000, patents and new work are positively correlated. However, holding patent counts fixed, there is still substantial variation in new work, which could reflect differences in the application of inventions to production. Finally, I also compare imputed city-manufacturing total factor productivity growth over the previous decade to new work employment shares in each census year (Figure 5).[12] Again, holding constant new work, there is still substantial variation in measured total factor productivity growth.

3 Theoretical framework

This section describes a model in which initial city characteristics—in particular, aggregate educational attainment and industry variety—matter for the subsequent location of new work. The starting point is to imagine innovation as a shock to the economy. How do workers and firms across locations adapt to this

[11] Specifically, I use utility patents, or patents for invention. The industry-level data are from the NBER U.S. Patent Citations Data File (Hall, Jaffe, and Tratjenberg, 2001), and I have used the U.S. P.T.O.'s concordance (2007) between patent classes and Standard Industrial Codes. The concordance focuses on manufacturing industries, which limits detail in non-manufacturing industries. City-level data are aggregated from county-level statistics from the U.S. P.T.O. (2000). New work shares are computed using IPUMS extracts.

[12] I use estimated TFP in 1970, 1980, 1990, and 1996 (the latest year available), for two-digit SIC industries, from the NBER-CES database (Bartelsman, Becker, and Gray, 2000). Then, by city, I take a weighted average of industry TFP, where the weights are initial industry employment shares within each city, from the IPUMS. This procedure yields an imputed measure of log TFP growth by city.

shock? The model helps to interpret the resulting spatial distribution of new work as coming from a kind of agglomeration economy, as neither workers nor firms fully internalize benefits associated with adaptation, realized by others nearby, that follow from their location decisions. The main purpose here is to guide the estimation strategy and inform the interpretation of the estimation results; the discussion here is informal while technical details are relegated to the theory appendix.

Consider a static, general equilibrium, economic geography model, as in Helpman (1998). There are a variety of goods that are traded among locations, with (iceberg) transportation costs, subject to plant-level scale economies in production. This generates incentives to agglomerate: assuming traded goods are demanded at every location, firms want to locate plants where demand is greatest, thus achieving scale and minimizing transport costs (cf. "backward linkages" in Hirschman, 1958). Add to this a non-traded good that is fixed in supply across locations. Increasing local demand for the non-traded good raises its price, thus providing incentive for households to disperse; these congestion costs prevent economic activity from concentrating in a single location. In equilibrium, the marginal mobile household is indifferent across locations; typically, one location may have high productivity, a large number of households, a large variety of goods produced, yet high offsetting congestion costs, while another location may have low productivity, a small number of households, a small variety of goods produced, and low congestion costs.

In such a model, I introduce innovation via new varieties of traded goods and new types of work; assume that new-good plants employ more new work as intermediate inputs than existing-good plants and that new goods are normal, so that educated workers demand more new goods. In the new (static) equilibrium, the same (net) agglomeration economies operate on location decisions; as before, firms producing new goods, using new work, will want to locate where demand for new goods is greatest, in this case the location with more educated workers. Similarly, industry variety will ensure greater local demand for new work inputs. The agglomeration economy arises because households or firms do not fully internalize the price effects of their location decisions, via the sharing of plant-level fixed costs across greater local demand for products or intermediate inputs that use new work intensively.

The key empirical prediction is that, following innovation, new work will most likely appear in locations that, initially, have a variety of industries and educated workers. In the model, this happens because educated workers (or industries) use more goods (inputs) that use new work intensively, without generating congestion costs that dissuade new work from locating near demand.

Some caveats are in order. First, the model lacks fully specified dynamics. However, if factor rewards across locations are persistent, slow adjustment processes are unlikely to reverse the main prediction (cf. Krugman, 1991a). Second, other micro-foundations for agglomeration economies can generate similar results. In this discussion, I have emphasized pecuniary externalities and the demand for new work as a source of economies of density, as in Krugman (1991b). However, alternative mechanisms are available. For example, new ideas may diffuse faster in locations with educated workers (Glaeser, 1999), matching between firms and processes may be less costly in industrially diverse areas (Duranton and Puga, 2001), or factors that are production complements to new technologies (e.g., college-educated workers) may be relatively abundant and thus cheap (Beaudry, Doms, and Lewis, 2006). (These models emphasize externalities related to the local supply of inputs to new work.) Less formally, any location contains agents that might use new ideas, embodied in either intermediate inputs or final products. Cities then operate like thick-market matching machines between ideas and work: a variety of firms or educated workers make it more likely in that location that a new idea will be recognized as productive or useful, and thus implemented into production or consumed.

In practice, identical reduced-form predictions make it difficult to distinguish between specific theoretical mechanisms. However, the key lessons are that (1) adaptation to new technologies is a potentially distinct form of agglomeration economies, whether these economies of density are related to demand or supply, and (2) we can evaluate these economies of density by examining the spatial distribution of new work.

4 Estimation

In this section, I describe a test for economies of density in adaptation to technological change. The main evidence in favor of these kinds of economies is the result that, all else equal, workers are more likely to be observed in a *new occupation* when they live in cities that were initially dense in both college graduates and industry variety.

To show this, I use worker-level microdata, pooled from the 1980, 1990, and 2000 U.S. censuses. The basic model, for $m = 1, ..., M_{gt}$, $g = 1, ..., 363$, and $t = 1, 2, 3$, is

$$\nu_{mgt} = \alpha + \mathbf{X}_{g,t-1}\beta + \mathbf{Z}_{mgt}\gamma_{gt} + \mu_{mgt} \tag{1}$$

where ν_{mgt} is the new work variable, a scalar, measuring the likelihood that a worker m, observed (once) living in city g and census year t, has selected into a new activity. As described earlier, these new activities are detailed occupations, with new titles, that have emerged in the national labor market since census year $t-1$. $\mathbf{X}_{g,t-1}$ is a vector of time-varying initial city characteristics measured in the census year *prior* to the worker observation (since the outcome of interest dates from $t-1$). To characterize cities' initial education distributions, it includes the college- and high school-graduate share of the population age 25 years and older; it also includes a measure of industry variety (a Herfindahl-Hirschman index based on employment) and, to control for city size, the logarithm of population density.[13] $\mathbf{Z}_{mgt}$ is a vector of worker characteristics,

[13] I check if the results are robust to alternate measures of city education and industry variety in Table 9. Those mentioned here win in a horse-race regression.

including flexible, dummy-variable controls for sex, race, ethnicity, nativity, educational attainment, marital status, industry, and potential experience. Of course, using census microdata, rather than city-level averages, I can better separate the location of new work due solely to the skill- or industry-bias of new occupations from location outcomes due to agglomeration economies.

The vector of interest is β, which I interpret as identifying the reduced-form effect of initial local characteristics on the subsequent appearance of new production activities. I allow γ, the coefficient vector on worker characteristics, to vary over t; among other reasons, the skill bias of new technologies has changed over time. Similarly, in some specifications, I allow γ to vary across locations, since, as documented by Bleakley and Lin (2007), the returns to occupation-specific skills can vary jointly across cities and levels of potential experience.

Suppose that the error term μ_{mgt} has city and census-year components that can be separated additively: $\mu_{mgt} = c_g + d_t + e_{gt} + u_{mgt}$. Here, c_g and d_t are (unobserved) city and census-year effects, e_{gt} is an unobserved time-varying city effect, and u_{mgt} is the idiosyncratic error.

At the worker level, the identifying restriction is that u_{mgt} be uncorrelated with observed worker characteristics and both observed and unobserved city characteristics. Note that this restriction is plausibly violated if workers' location choices are related to their unobserved characteristics. For example, if entrepreneurial workers are increasingly attracted to hip cities over time, and entrepreneurship and hipness are both unobserved and correlated with the included regressors, then the estimated $\hat{\beta}$ in part could reflect changes in sorting patterns, rather than agglomeration economies. I attempt to address this identification issue in three ways: first, by examining groups of workers where unobserved heterogeneity is less severe; second, by using geographic aggregates larger than cities (U.S. states) across which moving may be more costly; and third, by using an instrument to estimate $\hat{\beta}$ for those workers whose location choice is affected by their place of birth.

A second potential concern in interpreting $\hat{\beta}$ is the presence of c_g, the city error component. Note that since c_g is unobserved, a cross-section regression of ν_{mgt} on $\mathbf{X}_{g,t-1}$ leads to estimates of β that reflect any arbitrary correlation between $\mathbf{X}_{g,t-1}$ and unobserved, fixed city factors captured in c_g. I correct this problem using a fixed-effects specification, made possible by pooling three census-year cross-sections. In addition, d_t absorbs unobserved changes over time, say, in the census procedures used to add new occupation categories. This is especially important because the two earlier revisions use the *DOT*, while the later revision uses a different source, the census *Classified Index*.

I could estimate Equation (1) in a single step, but for the small (and in my experiments, negligible) bias that comes from not observing e_{gt}. More relevantly, computational challenges arise when $\mathbf{X}_{g,t-1}$ and $\mathbf{Z}_{mgt}$ contain large sets of dummy variables and $\sum_g \sum_t M_{gt}$ contains over eight million worker observations. Therefore, as in Loeb and Bound (1996), I perform the estimation in two steps. In the first step, at the worker level, I estimate

$$\nu_{mgt} = \delta_{gt} + \mathbf{Z}_{mgt}\gamma_{gt} + u_{mgt}, m = 1, ..., M_{gt} \tag{2}$$

using ordinary least squares, to obtain estimates and robust standard errors of the city-year effects on new work captured by δ_{gt}. (The first-step estimates of the variance matrix are made cluster-robust to three-digit detailed occupation to account for the grouping of the dependent variable ν_{mgt}. This procedure also helpfully simplifies the identifying restriction noted earlier by combining both observed and unobserved city characteristics in δ_{gt}.) I then use the minimum distance estimator (Wooldridge, 2003) to obtain $\hat{\beta}$, the coefficient of interest, from a second-step regression, at the city-year level.

$$\hat{\delta}_{gt} = \alpha + \mathbf{X}_{g,t-1}\beta + c_g + d_t + e_{gt}, g = 1, ..., 363, t = 1, 2, 3 \tag{3}$$

In practice, the second step is estimated using weighted least squares, where the weights are $1/\widehat{Avar}(\hat{\delta}_g)$. Of course, with included city and time effects, $\hat{\beta}$ is identified by variation in within-city changes in initial industry variety and college share.

At the city-year level, the identifying restriction is that e_{gt} is uncorrelated with the observed changes in city characteristics. Thus a final potential concern in interpreting $\hat{\beta}$ is the presence of unobserved, endogenous time-varying city characteristics (including changes in the value of city fixed factors). I attempt to address this concern in a number of ways: first, by including a host of additional controls in $\mathbf{X}_{g,t-1}$, as well as state$\times$year or census region$\times$year effects; and second, by using an instrumental-variables strategy involving lagged city-industry structure and lagged city-age structure to generate variation in initial city characteristics that is perhaps exogenous to contemporaneous, unobserved city changes.

5 Results

Table 4 displays typical worker-level results from the first-step estimation, specified in Equation (2). In each census year, I include estimates with and without indicator variables for major (one-digit) industry.[14] (For presentation, I have multiplied the coefficients by 100 so that the dependent variable, new work, is expressed in percentage-point units.) Each regression includes city dummies, which I then use in the second-step regression.[15]

There is a consistent correlation between new work and workers' educational attainment. While coefficient

[14] Since occupation and industry choice are often closely related, I am hesitant to include regressors that are clearly endogenous. Still, I would like to control for the fact that technological change may be biased toward particular industries. It turns out that coefficient estimates are mostly stable across specifications.

[15] In general, I report conservative standard error estimates, relying on cluster-robust variance-matrix estimators in the fixed effects regressions. The large number of cities ($G = 363$) and the large number of worker observations allow me to rely on asymptotic properties of robust variance-matrix estimators (Wooldridge, 2003), and the short panel ($T = 3$) helps to avoid misleading standard-error estimates due to serial correlation (Kézdi, 2004).

estimates for other control variables are either close to zero or changing over time, workers with less than a high school diploma are significantly less likely to be observed in new work across all census years. In 1980, workers with high school diplomas or some post-secondary education are most likely to be in new work; in both 1990 and 2000, college graduates are most likely to be in new work. Notably, increasing educational attainment of new work over this period is coincident with shifts in labor demand attributed to skill-biased technological change by Berman, Bound and Griliches (1994).

Also, after controlling for observed worker characteristics, location remains an important explanation for selection into new work: across all census years and specifications, the reported F-statistics suggest that I can reject the hypothesis that the included city-year dummies are jointly zero.

Each numbered column of Table 5 reports second-step estimates from a separate regression of estimated city-year dummies on initial city-year variables, city dummies, and census year dummies, as in Equation (3). The first four regressions use different sets of worker-level controls in the first step—no controls in column (1), basic controls, as in the first column of Table 4, in column (2), basic controls including industry dummies in column (3), and, in column (4), coefficient estimates on all first-step controls are allowed to vary without restriction across cities and years. Finally, in column (5), I drop the 1980 data in response to general concerns about the validity of the 1977 title data (see footnote 10).

Estimates of the coefficient on cities' initial college graduate share suggest that workers are more likely to be observed in new work when they live in a city with a larger initial stock of college graduates, holding initial population constant. Taking one estimate, from column (2), a worker is about 0.6 percentage point more likely to be in new work relative to an observationally similar worker when they live in cities separated by one standard deviation in initial college share. Workers are less likely to be in new work when they live in cities with larger initial stocks of high school graduates, controlling for both city size and the college share. Since high school dropouts are the omitted group, the interpretation here is that of comparing cities of the same population and area, replacing dropouts with high school (or college) graduates. Again taking the estimate from column (2), a worker is about 0.3 percentage point less likely to be in new work than a similar worker when they live in cities separated by one standard deviation in initial high school share.

Taken together, the estimates of the coefficients on cities' initial college and high school shares are consistent with a model featuring agglomeration economies in adapting to new technologies. In the context of such a model, the positive college share coefficient estimate is easy to interpret: educated workers, in some way, generate local externalities to new work that, net of the congestion costs they create, are positive. Interpreting the negative high school share coefficient estimate is perhaps trickier. One natural interpretation is that, unlike college-educated workers or high school dropouts, these workers generate congestion costs greater than any local external benefits to new work for which they might be responsible. In other words, the omitted category of workers, high school dropouts, may not attract new work, but they probably do not compete for the local factors that do attract new work. A more specific alternative interpretation is that there are pecuniary externalities from density in the factor market for high-school-educated labor that discourage adaptation to new technologies; this is similar to the interpretation that Beaudry, Doms, and Lewis (2006) attach to the distribution of PCs across U.S. cities.[16]

Workers are more likely to be observed in new work when they live in cities with greater initial industry variety, as measured by a Herfindahl-Hirschman index, controlling for city size and city education.[17] (I multiply this index by -1, so that higher values indicate more industry variety.) Holding city population and area constant, the interpretation is that of replacing a worker in a well-represented industry with a worker in a minimally represented industry. Taking the estimate from column (2), a worker is about 0.2 percentage point more likely to be in new work than an observationally similar worker in another city separated by one standard deviation in initial industry variety. In contrast, controlling for other initial city characteristics, there is no clear correlation between initial city population density and new work.[18]

Allowing first-step coefficients to vary across cities and census years, as in column (4), yields estimates that are much larger than when allowing coefficients to vary only across census years, as in columns (2) and (3). For comparison, refer to column (1), which essentially reports raw partial correlations between initial city characteristics and new work. In columns (2) and (3), the estimates attenuate toward zero because the location of new work is partly explained by the location of industries and skilled workers favored by technological change. If the bias of technological adaptation varies across locations—for example, if the propensity of a college graduate to select into new work is *less* in locations with a lot of college graduates—then, as in column (4), there is even more unexplained variation in the location of new work attributed to the city dummies. This may be an appropriately flexible specification for some of the first-step coefficients. For example, Bleakley and Lin (2007) show that a worker's propensity to change occupations varies jointly over experience and location: in dense cities, younger workers are more likely, while older workers are less likely, to churn through occupations. However, the restriction that, say, Asians should select into new work at the same rate across locations (within the same census year) does not seem unreasonable. For this reason, I prefer the estimates from columns (2) and (3) using time- (but not city-) varying first-step coefficients, with the caveat that relaxing this restriction tends to increase the magnitude of the estimates.

The high R^2 values reflect the contribution of the city and census-year dummies; however, the four reported city-year characteristics are still important explanatory variables for city-year variation in new work. The reported F-statistic is for the test that the four reported coefficients are jointly zero; for each

[16] In this interpretation, new technologies appearing in the 1970s, 1980s, and 1990s are production substitutes for high-school-educated workers, but complements to high school dropouts and college graduates. See Autor, Levy, and Murnane (2003) for some evidence on this complementarity between skill and recent technological change.

[17] I use county-level data from U.S. Census Summary File 4, aggregated to "cities," from 1970, 1980, and 1990, to calculate this index, based on employment in two-digit industries. Alternate measures are discussed in Table 9.

[18] Population density and industry variety are correlated; including them in this way may be a misspecification. In Table 8, I estimate separately by city size, in order to check the robustness of the results to the specification of population density.

specification I can reject the hypothesis that these variables are unimportant.

To summarize, results displayed in Tables 4 and 5 suggest that, controlling for worker characteristics and fixed city factors, cross-sectional variation in new work has a geographic component that can be explained by initial city education and industry variety. These results are consistent with agglomeration economies from the spatial concentration of factors and markets related to technological adaptation. In the following sections, I consider the robustness of this interpretation to plausible alternative explanations.

5.1 Unobserved worker characteristics

Differences in estimates across columns in Table 5 suggest that it is important to account for observed worker characteristics when trying to explain the location of new work. By extension, the interpretation of these estimates as external effects may be confounded by the presence of unobserved worker characteristics—in particular, those unobserved characteristics that might influence location choice (i.e., the city dummies). (Unobserved worker characteristics that are orthogonal to the city dummies will not bias second-step city-year estimates.)

One way to gauge the possible extent of bias is to examine workers' recent migration history using information in the IPUMS. Workers who lived in a different U.S. state (or country) 5 years before the survey date might be more likely to have confounding unobserved characteristics. If the results seem stronger using this sample of movers, that would raise concerns about worker sorting based on unobservables. However, as shown in columns (1) and (2) of Table 6, estimated coefficients are similar across samples. Similarly, unobserved characteristics related to location choice may be correlated with educational attainment—as Bound and Holzer (2000) document, more-educated workers are also more mobile. Columns (3)–(5) report results using three separate samples composed of high school dropouts, high school graduates, and college graduates. Estimates are similar across groups, or at least it is difficult to attach a sorting explanation to the results. Finally, for results reported in column (6), the first-step regression is limited to men in their 40s, a group where problems of unobserved heterogeneity may be less severe than the sample as a whole; these estimates are similar to the benchmark results.

Another strategy to address sorting on unobservables is to consider larger geographic units, under the assumption that moving is more costly, and less frequent, across states than it is across sub-state areas. I recalculate city-level variables at the level of 48 U.S. states; state-year results are similar to city-year results (column 7).[19] I can also match these state-year characteristics to workers' U.S. state of birth (thus excluding foreign-born workers). To the extent that state-of-birth characteristics help explain selection into new work (column 8), I assume it is because many people continue to live in their state of birth. This suggests an instrumental variables strategy, reported in column (9) (Evans, Oates, and Schwab, 1992, use

[19] I group the Washington, D.C. metro area with the state of Maryland.

a similar strategy). Here, I instrument the 48 state-of-residence dummies in the worker-level regression using 65 place-of-birth dummies (including the foreign-born) interacted with two age-group dummies. In the state-year regression, I regress the instrumented estimates of the dummies against state-year characteristics. The results using IV in the first step are similar, though less precise, than the results using OLS in the first step, though these results should be interpreted with a great deal of caution. Diagnostic tests report that some of the instruments are weak (some of the partial–R^2 values are as low as 0.05; see Shea, 1997) and the instruments fail the overidentification test. In addition, in this case I am estimating a local average treatment effect for those workers who were induced to live in a particular state because it was their state of birth, and these workers may have different (unobserved) characteristics than the sample as a whole.

Results reported in Table 6 suggest that there may be unobserved worker characteristics that are correlated with location choice. However, based on these results, it seems unlikely that these unobserved worker characteristics are responsible for generating seriously misleading estimates.

5.2 Unobserved city-year characteristics

The second-step results presented already control for city fixed effects, including the presence of any fixed factors that might contribute to local adaptation. However, there may be unobserved endogenous, time-varying city characteristics, including changes in the value of fixed factors, that could explain remaining variation in the location of new work. For example, increasing demand over time for good weather could affect the location of new work, without leaving any explanatory role for agglomeration economies. I attempt to address factors like this in a number of ways. In column (1) of Table 7, I report estimates including census year-U.S. state effects (3×48 additional regressors). These additional regressors should control for changes that are common to cities that are within the same state. The results remain similar, though an F-test indicates that the year-state effects are important explanatory variables.

As a second strategy, I directly include other city-year characteristics. I calculate a set of three regressors, each an index based on cities' initial industry composition. These indexes are similar to those used in Bartik (1991) and Blanchard and Katz (1992); in these applications, the original motivation is to capture employment growth in a location, as predicted by initial industry composition and the growth of its industries nationally. For each city-year, I calculate $\hat{\eta}_{g,t} = \sum_k \xi_{gk,t-1}\eta_{k,t}$, where $\xi_{gk,t-1}$ is the share of employment of (three-digit) industry k in city g in the previous census year, and $\eta_{k,t}$ is either (1) the logarithm of national employment growth in industry k between census years $t-1$ and t, (2) the logarithm of national new work employment in industry k in census year t, or (3) the logarithm of total accumulated patents in industry k, nationally, between census years $t-1$ and t. The indexes thus measure predicted city-year values of employment growth, new work employment, and patenting, based on industry mix in the previous census year. I find, as reported in column (2), that initial city education and industry variety still explain variation in the location of new work, even when controlling for predicted new work employment based on initial industry composition.

The regression reported in column (3) includes additional city-year regressors, including cities' initial workforce shares in major (one-digit) occupation and industry categories and the logarithm of average establishment size in four major industries. Changes in the industry, occupation, or establishment size composition of cities do not appear to significantly bias the estimated contribution of initial city education or industry variety to the location of new work. Taken together, reported estimates in the first three columns of Table 7 suggest that unobserved city-year characteristics are unlikely to seriously bias the results. In addition, the series of reported F-statistics suggest that remaining city-year variation in the location of new work is successively more difficult to explain.

A third strategy is to attempt to find a source of exogenous variation in city-year education and industry variety. Potential candidates are the lagged characteristics of cities; assuming that past city characteristics are the result of historical processes that have dissipated, these instruments may be valid. I use as instruments lagged values (i.e., 20 years before the census year in which the worker is observed) of the three indexes based on initial industry composition. For additional instruments, I use lagged variables measuring city population shares in three age groups, motivated by Moretti's (2004) argument that secular changes in educational attainment may exogenously drive differences in college share across cities. Finally, I consider using lagged values of the four reported city-year characteristics themselves. For these lagged values to be valid as instruments, I have to make a strong assumption that current changes in cities are unrelated to processes that generated historic location patterns.

Columns (4) and (5) report city-year two-stage least-squares estimates, for different sets of instruments. In column (4), using lagged age and industry structure as instruments, the point estimates are similar (though the estimated coefficient on initial industry variety is imprecise); however, the instruments are rather weak—partial R^2 values are low. In column (5), using lagged age structure, city education, industry variety, and population density, the estimates are more precise, the partial–R^2 values are relatively high, and the instruments pass the overidentification test. Overall, results reported in Table 7 suggest that, even when controlling for other city characteristics that are changing over time, initial city education and industry variety still matter for explaining the location of new work.

5.3 Other robustness checks

Columns (1a), (1b), and (1c) of Table 8 report estimates from a city-year regression when I allow for time-varying coefficients. The estimates are less precise, but the main correlations hold.[20]

In columns (2)–(5) of Table 8, I report estimates for samples separated by (1970) city size. In these

[20]Of course, it would be difficult to interpret cross-sectional estimates, since panel estimates include controls for city fixed factors. An earlier version of this paper does report similar cross-sectional patterns.

regressions, I therefore no longer restrict the coefficients on initial city education or industry variety to be the same for every size of city, and I control for initial population density in a more flexible way. Across all sizes of cities, initial city education is an important explanatory variable for new work. However, only in smaller cities does the coefficient on initial industry variety appear large and significantly different from zero. Part of this result is due to greater variation in initial industry variety in smaller cities; big cities are more likely to have high employment across many sectors, and thus differences in industry variety across big cities are small.

An alternative explanation to agglomeration economies is that the location of new work is driven by demand for non-traded goods and services in rapidly growing cities in the southern and western U.S. This explanation concerns catch-up to a new long-run equilibrium in response to a secular trend (e.g., increasing demand for good weather), rather than an equilibrium path generated by agglomeration economies. If this is the case, new work in non-traded sectors should be particularly sensitive to bias. Put another way, in the absence of trade, agglomeration economies should have little impact on location decisions. If the main results were mostly due to the geographic distribution of non-traded goods production, it would be harder to infer the presence of agglomeration economies. In fact, the main estimates seem driven by new work in traded-goods industries such as manufacturing (column 6). Controlling for fixed city factors, new work in non-traded services shows no tendency to agglomerate with initial city education and city variety (column 9).

Finally, because "city" definitions are county-based, it may be that the large sizes of counties, especially in the western U.S., are responsible for misleading comparisons across city-years. Splitting the sample of cities by census divisions in columns (10)–(12), I find little evidence that this is the case.

5.4 Measurement

Table 9 reports experiments with alternative measures of industry variety, city education, and new work. Column (1) repeats benchmark estimates reported in the second column of Table 5. In column (2), I use as the dependent variable new work based on alternate new title lists in 1991 and 2000 (described in the data appendix). In columns (3) and (4), I report estimates using alternate measures of initial city education and industry variety. Here, instead of the college- and high school-graduate shares, I use the logarithm of total college and high school graduates; holding population density constant, the interpretation of the coefficients is that of replacing a high-school dropout with another worker having higher educational attainment. Instead of a Herfindahl-Hirschman index of industry concentration, I use the logarithm of the total number of two-digit industries observed initially in the city. In addition, since most cities contain almost all possible two-digit industries, in columns (5) and (6) I only count the number of city-industries with initial employment above 1,000. To summarize the results reported in Table 9, the correlations between new work and initial city education and industry variety do not appear to be especially sensitive to choices made in measuring any of these variables.

In Table 10, I return to questions first raised in the data section related to the distribution of title counts and title employment across detailed occupations. Each of the first seven columns reports estimates using a subsample of workers, separated based on detailed occupation. The idea is to check to see whether the results are due primarily to detailed occupations where suspected bias in the imputation of new work employment is greatest. Favorably, in columns (1) and (2), the main results appear strongest for those detailed occupations where the share of new titles is probably closest to actual new work employment. In these regressions I have separated workers based on the new title share of their detailed occupation; the regression reported in column (2) uses workers with detailed occupations with new title shares closest to 1.

A related concern is that bias in a few high-employment detailed occupations may be responsible for the correlation between new work and initial city characteristics. Results in columns (3) and (4) suggest that this is not the case. The top 23 largest detailed occupations by employment, employing about one-fifth of the overall number of workers, do not appear to be distributed across locations in the same way as other occupations. If the new title share does not accurately measure new work employment in these 23 detailed occupations, then this bias is working against the main result. In addition, comparing means and standard deviations of the dependent variable ν shows that detailed occupations with the smallest employment levels have the highest new title shares.

The total number of titles in each detailed occupation may be correlated with the degree of bias coming from imputing new work employment based on new titles. If a detailed occupation contains few titles overall, then the new title share is likely to be a poor approximation of employment. On the other hand, if a detailed occupation contains many titles, then small differences in the distribution of employment across titles may introduce greater bias. As reported in columns (5)–(7), I find that neither type of detailed occupation appears responsible for the main pattern of results. Instead, it is workers in detailed occupations that have a medium number of titles, with high new title shares, that locate in cities with high initial education and industry variety.

Finally, a small percentage of workers in each census report detailed occupations that are allocated by the census based on other observed characteristics; dropping these observations, as in column (8), does not affect the results.

6 Conclusions

In this paper, I find evidence for the role of agglomeration economies in adaptation to new technologies. Cities with high levels of initial education and industry variety are better able to attract new work, even after accounting for a variety of fixed and mobile factors, suggesting that geographic concentrations of factors and markets are important for speeding the application of new knowledge in production. A main contribution of this paper is to create a measure of adaptation to technological change, based on changes in occupation classifications, that can more closely and systematically characterize the impact of new discoveries and inventions on the organization of production.

New work may have further value as a way to investigate other cross-sectional questions related to technological adaptation. Much of the literature on the direction of technological progress focuses on relative factor supplies across locations or time periods, but there is still a sizable amount that is not well understood. For example, there is less evidence relating adaptation on the level of the individual to, say, residual wage inequality. New work may also help to understand organizational differences across firms. Finally, to the extent that international occupation crosswalks are available, new work may be useful as a systematic measure of the breadth of technological adaptation across countries.

References

Acemoglu, Daron (2009). "When does labor scarcity encourage innovation?" NBER working paper 14809.

Alexopoulos, Michelle (2006). "Read all about it! What happens following a technology shock?" University of Toronto working paper.

Autor, David H., Frank Levy, and Richard J. Murnane (2003). "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics.* 118 (4), 1279–1333.

Bartelsman, Eric J., Randy A. Becker, and Wayne B. Gray (2000). *NBER-CES Manufacturing Industry Database*, [computer file], National Bureau of Economic Research [distributor].

Bartik, Timothy J. (1991). *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute for Employment Research.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2006). "Endogenous skill bias in technology adoption: City-level evidence from the IT revolution." NBER working paper 12521.

Benhabib, Jess and Mark M. Spiegel (2005). "Human capital and technology diffusion." in P. Aghion and S. N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, Elsevier, chapter 13, pp. 965–966.

Berman, Eli, John Bound, and Zvi Griliches (1994). "Changes in the demand for skilled labor within U.S. manufacturing: Evidence from the annual survey of manufacturers." *The Quarterly Journal of Economics.* 109 (2), 367–397.

Bils, Mark and Peter J. Klenow (2001). "The acceleration in variety growth." *The American Economic Review.* 91 (2), 274–280.

Blanchard, Oliver Jean and Lawrence F. Katz (1992). "Regional evolutions." *Brookings Papers on Economic Activity.* 1992 (1), 1–75.

Bleakley, Hoyt and Jeffrey Lin (2007). "Thick-market effects and churning in the labor market: Evidence from U.S. cities." Federal Reserve Bank of Philadelphia working paper 07-23.

Bound, John and Harry J. Holzer (2000). "Demand shifts, population adjustments, and labor market outcomes during the 1980s." *Journal of Labor Economics.* 18 (1), 20–54.

Carlino, Gerald A., Satyajit Chatterjee, and Robert M. Hunt (2007). "Urban density and the rate of invention." *Journal of Urban Economics.* 61, 389–419.

Comin, Diego, Bart Hobijn, and Emilie Rovito (2008). "Technology usage lags." *Journal of Economic Growth.* 13 (4), 237–256.

Duranton, Gilles and Diego Puga (2001). "Nursery cities: Urban diversity, process innovation, and the life-cycle of products." *The American Economic Review.* 91 (5), 1454–1477.

——— and ——— (2004). "Micro-foundations of urban agglomeration economies." in J. V. Henderson and J.-F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier, chapter 48, pp. 2063–2117.

Evans, Wiliam N., Wallace E. Oates, and Robert M. Schwab (1992). "Measuring peer group effects: A study of teenage behavior." *The Journal of Political Economy.* 100 (5), 966–991.

Glaeser, Edward L. (1999). "Learning in cities." *Journal of Urban Economics.* 46, 254–277.

——— and Albert Saiz (2003). "The rise of the skilled city." NBER working paper 10191.

Griliches, Zvi (1957). "Hybrid corn: An exploration in the economics of technological change." *Econometrica.* 25 (4), 501–522.

Haines, Michael R. (2004). *Historical, Demographic, Economic, and Social Data: The United States, 1790-2000*, [computer file], Inter-university Consortium for Political and Social Research [distributor], study number 2896.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER working paper 8498.

Helpman, Elhanan (1998). "The size of regions." in D. Pines, E. Sadka, and Y. Zilcha, eds., *Topics in Public Economics*, Cambridge University Press, pp. 33–54.

Hirschman, Albert O. (1958). *The Strategy of Economic Development*, Yale University Press.

Jacobs, Jane (1969). *The Economy of Cities*, Random House.

Jaffe, Adam B., Manual Trajtenberg, and Rebecca Henderson (1993). "Geographic localization of knowledge spillovers as evidenced by patent citations." *The Quarterly Journal of Economics.* 108 (3), 577–598.

Kézdi, Gábor (2004). "Robust standard error estimation in fixed-effects panel models." *Hungarian Statistical Review.* Special (9), 95–116.

Krugman, Paul (1991a). "History versus expectations." *The Quarterly Journal of Economics.* 106 (2), 651–667.

——— (1991b). "Increasing returns and economic geography." *The Journal of Political Economy.* 99 (3), 483–499.

Lin, Jeffrey (2007). "Innovation, cities, and new work." Federal Reserve Bank of Philadelphia working paper 07-25.

Loeb, Susanna and John Bound (1996). "The effect of measured school inputs on academic achievement: Evidence from the 1920s, 1930s, and 1940s birth cohorts." *The Review of Economics and Statistics.* 78 (4), 653–664.

Lucas, Robert E. (1988). "On the mechanics of economic development." *Journal of Monetary Economics.* 22, 3–42.

Marshall, Alfred (1920). *Principles of Economics*, Macmillan.

Meyer, Peter B. and Anastasiya M. Osborne (2005). "Proposed category system for 1960-2000 census occupations." U.S. Bureau of Labor Statistics working paper 383.

Mokyr, Joel (1992). *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford University Press.

Moretti, Enrico (2004). "Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data." *Journal of Econometrics.* 121 (1-2), 175–212.

National Academy of Sciences, Committee on Occupational Classification and Analysis (1981). *Dictionary of Occupational Titles (DOT): Part I - Current Population Survey, April 1971, Augmented With DOT Characteristics*, [computer file], U.S. Department of Commerce [producer], Inter-university Consortium for Political and Social Research [distributor], study number 7845.

Nelson, Richard R. and Edmund S. Phelps (1966). "Investment in humans, technological diffusion, and economic growth." *The American Economic Review.* March 56 (1/2), 69–75.

Parente, Stephen L. and Edward C. Prescott (1994). "Barriers to technology adoption and development." *The Journal of Political Economy.* 102 (2), 298–321.

Rauch, James E. (1993). "Productivity gains from geographic concentration of human capital: Evidence from the cities." *Journal of Urban Economics.* 34, 380–400.

Redding, Stephen J. and Daniel M. Sturm (2008). "The costs of remoteness: Evidence from German division and reunification." *The American Economic Review.* 98 (5), 1766–1797.

Romer, Paul M. (1990). "Endogenous technological change." *The Journal of Political Economy.* 98 (5), S71–S102.

Rosenthal, Stuart S. and William C. Strange (2004). "Evidence on the nature and sources of agglomeration economies." in J. V. Henderson and J.-F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier, chapter 49, pp. 2119–2171.

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2009). *Integrated Public Use Microdata Series: Version 4.0*, [computer file], Minnesota Population Center [producer and distributor].

Schultz, Theodore W. (1975). "The value of the ability to deal with disequilibria." *Journal of Economic Literature.* 13 (3), 827–846.

Scopp, Thomas S. (2003). "The relationship between the 1990 census and census 2000 industry and occupation classification systems." U.S. Census Bureau technical paper 65.

Shea, John (1997). "Instrument relevance in multivariate linear models: A simple measure." *The Review of Economics and Statistics.* 79 (2), 348–352.

U.S. Census Bureau (1950). *Classified Index of Industries and Occupations*, U.S. Department of Commerce.

——— (1968). "Changes Between the 1950 and 1960 Occupation and Industry Classifications." U.S. Census Bureau technical paper 18.

——— (1970). *Census of Population and Housing, 1970: Summary Statistic File 4C—Popuulation [Fourth Count]*, [computer file], U.S. Department of Commerce [producer], Inter-university Consortium for Political and Social Research [distributor], study number 8107.

——— (1972). "1970 Occupation and Industry Classification Systems in Terms of Their 1960 Occupation and Industry Elements." U.S. Census Bureau technical paper 26.

——— (1980a). *Alphabetical Index of Occupations and Industries: Technical Documentation*, [computer file], U.S. Department of Commerce.

——— (1980b). *Census of Population and Housing, 1980: Summary Tape File 4B Extract*, [computer file], U.S. Department of Commerce [producer], Inter-university Consortium for Political and Social Research [distributor], study number 8229.

——— (1989). "The Relationship Between the 1970 and 1980 Industry and Occupation Classification Systems." U.S. Census Bureau technical paper 59.

——— (1990a). *Census of Population and Housing, 1990: Summary Tape File 4B*, [computer file], U.S. Department of Commerce [producer], Inter-university Consortium for Political and Social Research [distributor], study number 6271.

——— (1990b). *Classified Index of Industries and Occupations*, [computer file], U.S. Department of Commerce.

——— (2000). *Classified Index of Industries and Occupations*, [computer file], U.S. Department of Commerce.

——— (2003). *Metropolitan and micropolitan statistical areas and components, with codes*, [computer file], http://www.census.gov/population/www/metroareas/lists/2003/03mfips.txt.

U.S. Department of Labor and U.S. Employment Service (1977). *Dictionary of Occupational Titles, Fourth Edition.*

——— and ——— (1979). *Conversion Table of Code and Title Changes, Third to Fourth Edition, Dictionary of Occupational Titles.*

———, ———, and North Carolina Occupational Analysis Field Center (1991). *Dictionary of Occupational Titles, Revised Fourth Edition*, [computer file], Inter-university Consortium for Political and Social Research [distributor], study number 6100.

U.S. Patent and Trademark Office (2000). *United States Patent Grants by State, County, and Metropolitan Area: Utility Patents, 1990-1999*, [computer file], U.S. Department of Commerce.

——— (2007). *Concordance Between the U.S. Patent Classification System and the 1972 Standard Industrial Classification System*, [computer file], Electronic Information Products Division, Patent Technology Monitoring Team [distributor].

Venables, Anthony J. (1996). "Equilibrium locations of vertically linked industries." *International Economic Review.* 37 (2), 341–359.

Wooldridge, Jeffrey M. (2003). "Cluster-sample methods in applied econometrics." *The American Economic Review, Papers and Proceedings.* 93 (2), 133–138.

Xiang, Chong (2005). "New goods and the relative demand for skilled labor." *The Review of Economics and Statistics.* 87 (2), 285–298.

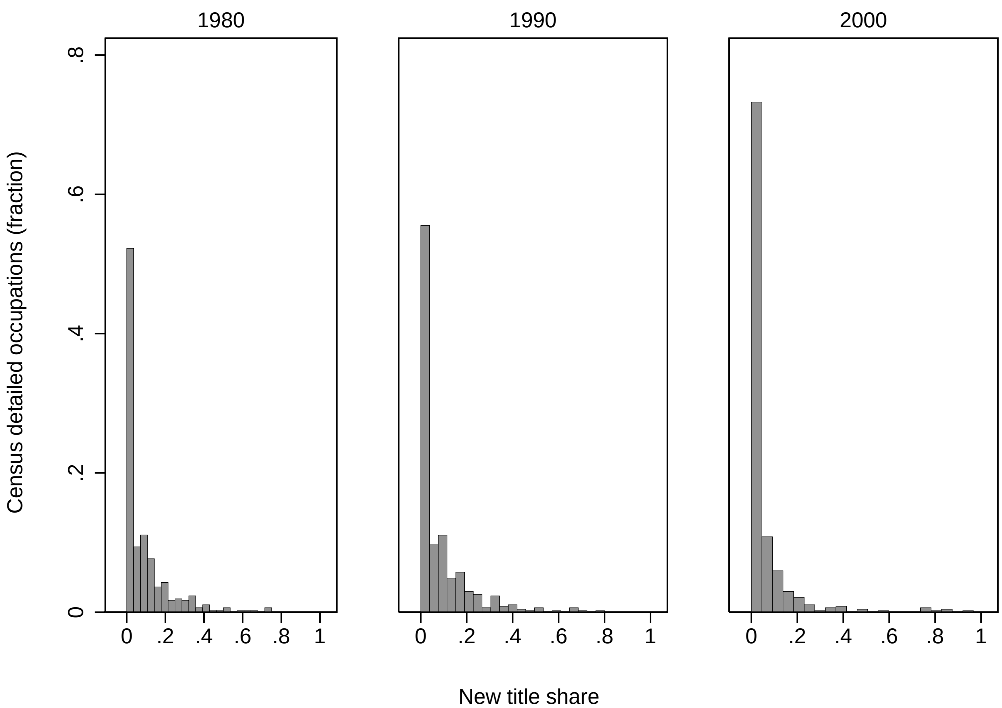


Figure 1: Distribution of census detailed occupations, by new title share

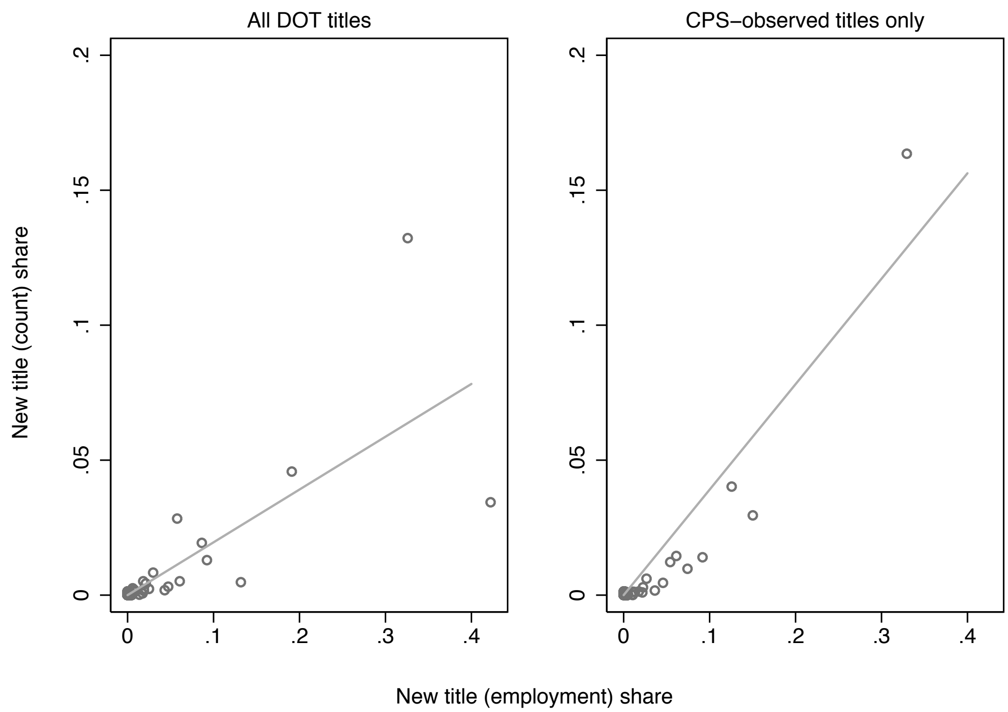
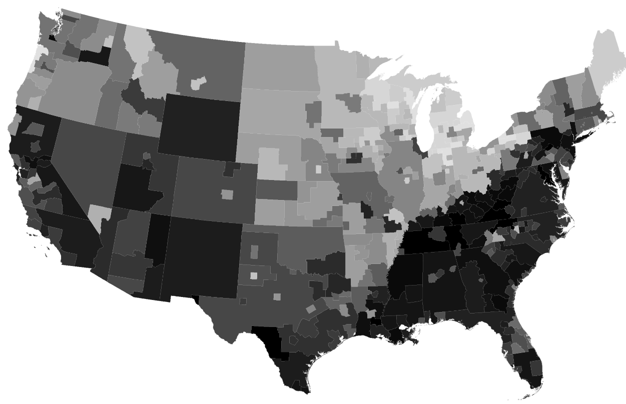
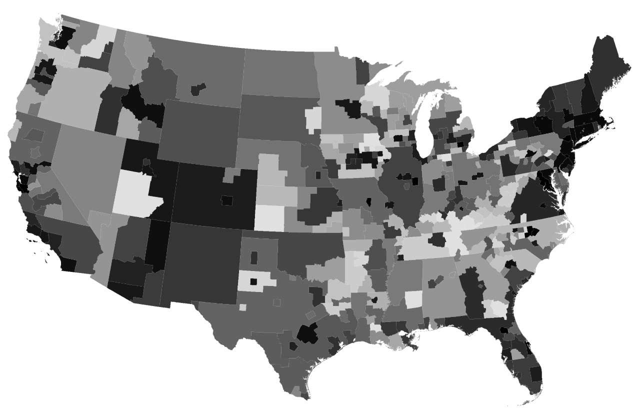


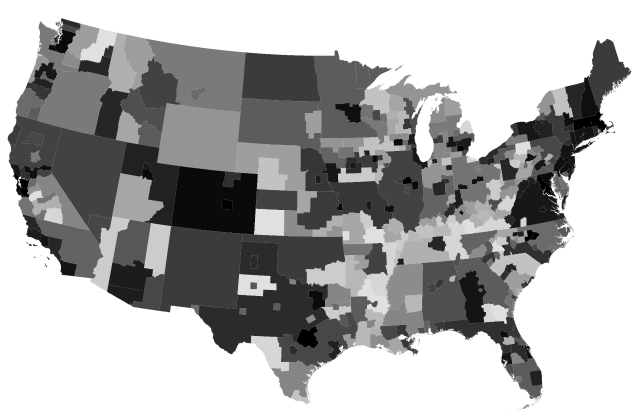
Figure 2: New title shares for detailed occupations, by count and employment, April 1971 CPS



1980 (New work: 1965-1977)



1990 (New work: 1977-1991)



2000 (New work: 1990-2000)

Figure 3: New work employment share across U.S. cities and locations

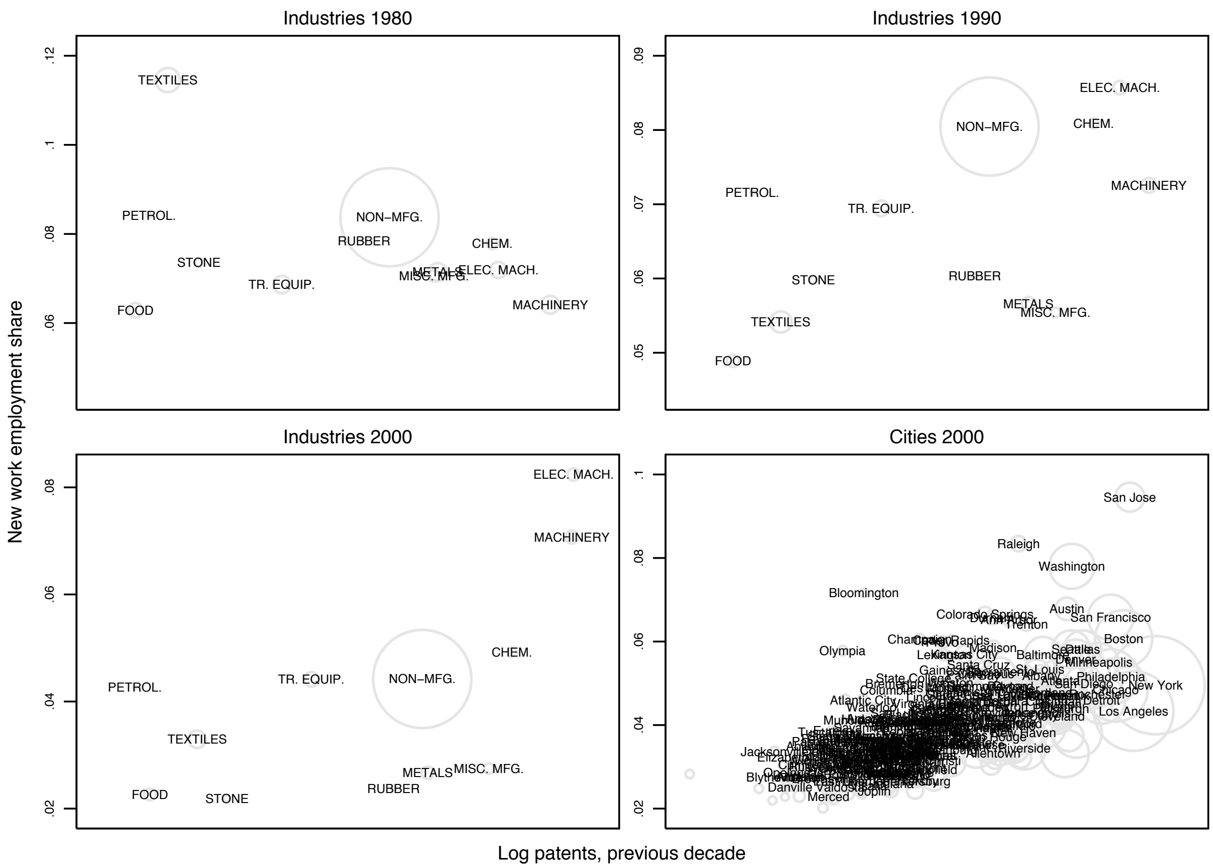


Figure 4: New work employment share and patents over previous decade

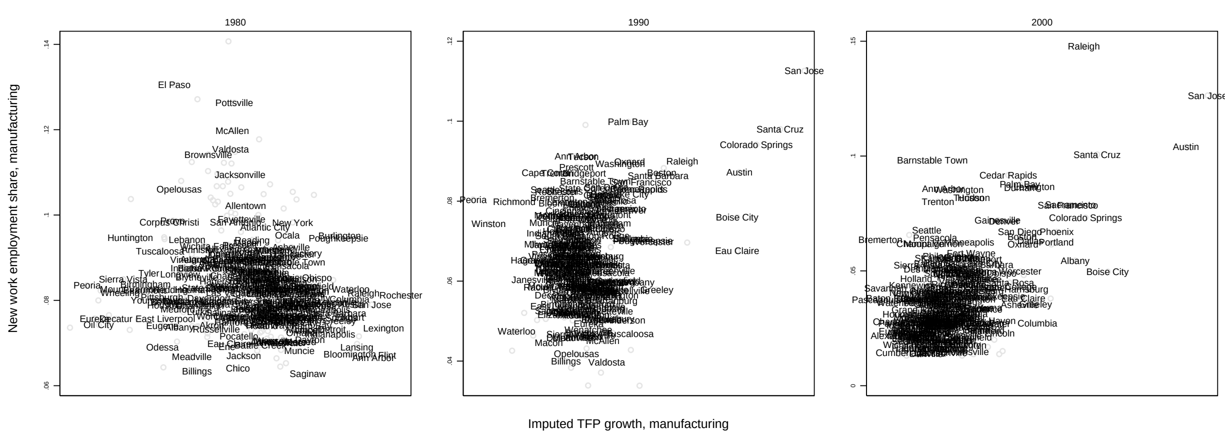


Figure 5: New work employment share and imputed city manufacturing TFP growth over previous decade

Table 1: New title examples and new title shares of corresponding census detailed occupations

Year	*Example new occupation titles*	*Corresponding census detailed occupation*	ν
2000	Chat room host/monitor	Network systems and data communication analysts	0.967
	Web developer		
	Database suppport	Computer support specialists	0.864
	Help desk specialist		
	Information systems security officer	Network and computer systems administrators	0.833
	Web administrator		
	Quality assurance specialist, applications	Computer software engineers	0.800
	Data recovery planner	Database administrators	0.769
	Manager, internet technology	Computer and information systems managers	0.765
	Dosimetrist	Radiation therapists	0.750
	Client-server programmer	Computer programmers	0.591
	Engineer, bio-mechanical	Biomedical engineers	0.571
1991	Database administrator	Computer systems analysts and scientists	0.800
	Programmer–analyst		
	Special procedures technologist, CT scan	Radiologic technicians	0.700
	Special procedures technologist, MRI		
	Radiopharmacist	Pharmacists	0.667
	Nesting operator, numerical control	Tool programmers, numerical control	0.667
1977	Agricultural-research engineer	Agricultural engineers	0.750
	Test engineer, agricultural equipment		
	Design engineer, nuclear equipment	Nuclear engineers	0.750
	Research engineer, nuclear equipment		
	Supervisor, animal cruelty investigation	Supervisors, guards	0.750
	Reports analyst	Management analysts	0.667

Notes: ν is the new occupation title share for each census detailed occupation. Constructed from U.S. Census Bureau and U.S. Department of Labor publications, including the *Classified Index of Occupation Titles* (U.S. Census, 1990b and 2000), *Dictionary of Occupational Titles, Revised 4th edition* (U.S. D.O.L., 1991), and *DOT4* (1977).

Table 2: New work employment shares

New work employment share (%)	1980 *base*	1990 *base*	*alt.*	2000 *base*	*alt.*
U.S.	**8.5**	**8.2**	**1.9**	**4.4**	**4.9**
Women	9.2	9.6	2.8	4.2	4.9
Age $<$ 26	8.2	6.7	1.4	3.4	4.4
26–41	8.5	8.6	2.2	5.1	5.3
42–55	8.7	8.7	1.9	4.5	5.0
56 $<$ Age	8.5	8.0	1.5	3.5	4.4
Less than high school	7.7	4.9	0.3	1.7	3.0
High school graduate	9.0	6.8	1.4	2.4	3.4
Some college	8.9	9.1	2.5	4.7	5.0
College graduate	8.1	11.6	2.9	8.0	7.7
Non-white	8.3	7.4	1.4	4.2	4.7
Non-Hispanic	8.5	8.3	1.9	4.6	5.1
City size $<$ 125,000	8.4	7.4	1.5	3.2	4.0
750,000 $<$ City size	8.5	8.3	2.0	4.7	5.1
Agriculture, construction, and mining	13.3	6.1	0.6	1.5	2.6
Manufacturing	8.4	7.0	1.3	4.2	3.4
Transp., communication, and utilities	7.7	6.3	1.2	4.0	4.5
Wholesale and retail trade	6.6	4.4	0.6	2.4	4.7
Personal services	3.2	7.2	0.7	3.9	4.2
Professional and other services	8.2	11.7	3.4	6.2	6.1
Managerial and professional occupations	7.1	14.3	3.7	8.3	8.9
Technical, sales, and support occupations	11.3	8.5	2.9	4.7	4.9
Service occupations	4.3	4.6	0.0	2.8	4.2
Production, craft, and repair occupations	7.7	5.0	0.0	1.4	1.0
Operators, fabricators, and laborers	8.5	3.7	0.2	0.7	1.0

Notes: Sample from IPUMS 1980-2000, workers age 16-70 with identified industries and occupations, excluding Alaska and Hawaii.

Table 3: Wages and new work

	1980			1990				2000			
Log hourly wage	(1)	(2)		(3)		(4)		(5)		(6)	
New work	0.101	0.103		0.793		0.303		0.846		0.339	
	(0.227)	(0.081)		(0.225)	**	(0.086)	**	(0.135)	**	(0.068)	**
Other controls	–	X		–		X		–		X	
R^2	0.000	0.261		0.014		0.325		0.022		0.308	

Notes: **–Significant at the 99% level of confidence. Robust standard errors, adjusted for clustering on detailed occupation, in parentheses. Each column is a separate OLS regression. M=1,324,406 in 2000, 1,112,143 in 1990, 4,905,016 in 1980. Sample is IPUMS 1980-2000, workers age 16-70 with calculated hourly wages between 0.01 and 100. Dependent variable is log hourly wage. Other controls are indicator variables for sex, marital status, race, ethnicity, nativity, educational attainment, age, relationship to household head, industry, and city.

Table 4: New work and worker characteristics (first-step estimates)

		1980			*1990*			*2000*	
New work (x100)	*Means*	(1)	(2)	*Means*	(3)	(4)	*Means*	(5)	(6)
H.S. graduate	0.38	1.28	1.35	0.34	1.45	1.16	0.31	0.67	0.45
(< h.s. omitted)		(0.65) *	(0.54) *		(0.37) **	(0.34) **		(0.18) **	(0.16) **
Some college	0.18	1.21	1.43	0.29	3.85	3.04	0.31	2.73	2.21
		(0.87)	(0.67) *		(0.64) **	(0.58) **		(0.62) **	(0.54) **
College graduate	0.17	0.45	0.99	0.21	6.31	4.62	0.25	5.66	4.63
		(1.08)	(0.88)		(1.18) **	(1.17) **		(1.41) **	(1.28) **
Pot. exp.	19.2	0.026	0.004	19.4	0.067	0.019	20.3	0.004	-0.009
	(15.1)	(0.038)	(0.033)	(13.9)	(0.032) *	(0.027)	(13.3)	(0.025)	(0.021)
Pot. exp.^{2} (/1000)		-0.31	0.04		-0.96	-0.34		-0.68	-0.57
		(0.63)	(0.56)		(0.54) †	(0.44)		(0.44)	(0.35)
Male	0.54	-1.22	-2.47	0.52	-2.83	-2.18	0.51	0.55	1.10
		(1.31)	(1.10) *		(0.81) **	(0.66) **		(0.61)	(0.63) †
Married	0.63	0.26	0.18	0.60	0.42	0.39	0.57	0.03	0.00
		(0.24)	(0.22)		(0.16) **	(0.15) **		(0.12)	(0.12)
Black	0.10	-0.44	-0.50	0.11	-0.34	-0.80	0.11	-0.35	-0.61
		(0.56)	(0.50)		(0.47)	(0.43) †		(0.30)	(0.30) *
Asian	0.01	0.09	0.22	0.03	-0.11	0.02	0.04	2.04	1.96
		(0.45)	(0.39)		(0.43)	(0.40)		(1.11) †	(1.08) †
Other race	0.01	0.12	-0.22	0.04	-0.13	-0.24	0.08	-0.14	-0.18
		(0.26)	(0.22)		(0.13)	(0.12) *		(0.08) †	(0.08) *
Hispanic	0.06	0.40	0.22	0.08	-0.06	-0.18	0.11	-0.79	-0.81
		(0.38)	(0.32)		(0.22)	(0.20)		(0.31) †	(0.30) **
Foreign-born	0.07	-0.02	0.04	0.10	-0.74	-0.61	0.14	0.06	0.15
		(0.43)	(0.36)		(0.23) **	(0.20) **		(0.30)	(0.30)
Industry		–	X		–	X		–	X
F (city dummies)		15.56 **	20.23 **		28.37 **	14.48 **		6.60 **	4.69 **
R^2		0.009	0.060		0.071	0.116		0.048	0.058
M (worker observations)		5,909,772			1,329,710			1,562,904	

Notes: **–Significant at the 99% level of confidence; *–95% level; †–90% level. Robust standard errors in parentheses, adjusted for clustering on census detailed occupation (O=469 in 1980, 462 in 1990, and 471 in 2000). Each numbered column is a separate OLS regression. Sample is IPUMS 1980-2000, workers age 16-70 with identified occupation and industry. Dependent variable is new work variable ×100. F-value is for null hypothesis that all city effects (G=363) are jointly zero.

Table 5: New work and initial city characteristics (second-step estimates)

Estimated city component of new work, from first step	*means* 1970	1980	1990	(1) *no 1st stage controls*		(2) *w/ worker controls*		(3) *w/ ind. controls*		(4) *city- and time-vary.*		(5) *1990, 2000 only (T=2)*	
College graduate share	0.09	0.14	0.17	19.9		11.3		11.4		35.9		14.2	
(< h.s. share omitted)	(0.04)	(0.05)	(0.06)	(1.7)	**	(1.6)	**	(2.3)	**	(4.1)	**	(2.3)	**
High school graduate share	0.41	0.50	0.57	-3.2		-3.6		-3.6		-5.6		-2.6	
	(0.08)	(0.07)	(0.06)	(0.8)	**	(0.8)	**	(1.1)	**	(2.1)	**	(1.0)	**
Industry variety	-0.08	-0.05	-0.04	10.8		10.4		7.4		19.6		9.5	
	(0.01)	(0.02)	(0.01)	(2.5)	**	(2.2)	**	(3.1)	*	(7.2)	**	(3.8)	**
Log population density	4.90	5.04	5.11	-0.38		-0.14		0.33		0.85		0.17	
	(1.39)	(1.31)	(1.30)	(0.13)	**	(0.13)		(0.21)		(0.39)	*	(0.21)	
N (city–year observations)	363	363	363	1,089		1,089		1,089		1,089		726	
F (initial city characteristics)				73.6	**	39.8	**	15.7	**	38.2	**	16.7	**
R^2				0.986		0.988		0.995		0.877		0.987	

Notes: **–Significant at the 99% level of confidence; *–95% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression, where the weights are the inverse estimated asymptotic variance of the dependent variable. Dependent variable is the estimated city component of new work from first step (see Table 4). Each regression includes city (G=363) and year (T=3) effects. All initial city-year characteristics are measured in the census year previous to the first-step estimate. Regression in column (1) uses city-level new-work means as the dependent variable. Regressions in columns (2) and (3) use control variables in first step, as in Table 4. Regression in column (4) allows coefficients on worker-level control variables to vary over cities and years. Regression in column (5) excludes 1980 data, using same regressors as (2).

Table 6: New work location and unobserved worker-city sorting

	OLS (in first step)								2SLS (in first step)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimated geographic component of new work, from first step	*non-movers only*	*movers only*	*h.s. dropouts*	*h.s. graduates*	*college graduates*	*prime–age men*	*state of residence*	*state of birth*	*place of birth*
College graduate share	11.0	13.3	0.4	16.1	15.4	15.1	11.7	5.1	11.8
(< h.s. share omitted)	(1.7) **	(2.3) **	(1.7)	(2.1) **	(2.7) **	(2.8) **	(4.5) *	(2.1) *	(3.4) **
High school graduate share	-3.6	-3.4	-2.4	-5.6	-2.8	-4.0	-3.5	-3.6	-3.9
	(0.8) **	(1.3) **	(0.9) **	(1.1) **	(1.6) †	(1.5) **	(2.3)	(1.1) **	(1.8) *
Industry variety	11.6	7.0	11.8	13.0	2.2	18.2	11.6	4.4	3.5
	(2.5) **	(3.4) *	(3.2) **	(3.4) **	(3.9)	(4.5) **	(5.6) *	(4.0)	(6.8)
Log population density	-0.15	-0.11	-0.45	0.00	-0.30	-0.47	0.09	-0.28	-0.56
	(0.14)	(0.19)	(0.14) **	(0.17)	(0.19)	(0.22) *	(0.37)	(0.16) †	(0.25) *
N (location–year observations)	1,089	1,089	1,089	1,089	1,089	1,089	144	144	144
R^2	0.987	0.968	0.972	0.982	0.914	0.952	0.994	0.998	0.994
Partial–R^2 (first stage) [min, max]	–	–	–	–	–	–	–	–	[0.05, 0.59]
p (overidentification) [min, max]	–	–	–	–	–	–	–	–	[0.00, 0.00]

Notes: **–Significant at the 99% level of confidence; *–95% level; †–90% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression (see Tables 4 and 5). Each regression includes city (G=363) or state (G=48) and year (T=3) effects. All initial city-year characteristics are measured in the census year previous to the first-step estimate. Regressions (1)-(8) estimate geographic components of new work using OLS in first step. Regression (9) estimates geographic components using 2SLS in first step, where instruments are interactions between 65 place-of-birth dummies and 2 age-group dummies.

Table 7: New work location and unobserved city-year effects

Estimated city component of new work, from first step	*WLS (in second step)* (1)		(2)		(3)		*2SLS (in second step)* (4)		(5)	
College graduate share	12.6		12.0		10.9		14.4		15.5	
(< h.s. share omitted)	(1.8)	**	(1.8)	**	(2.1)	**	(4.4)	**	(2.8)	**
High school graduate share	-3.4		-2.8		-3.0		-4.7		-4.0	
	(1.2)	**	(1.0)	**	(1.1)	**	(1.6)	**	(1.5)	**
Industry variety	6.6		5.4		4.0		16.4		10.2	
	(2.5)	**	(2.4)	*	(2.6)		(16.6)		(4.3)	*
Log population density	0.08		0.03		0.00		-0.21		0.30	
	(0.21)		(0.14)		(0.14)		(0.49)		(0.28)	
Year–State effects (3x48)	X		–		–		–		–	
Year–Census region effects (3x9)	–		X		X		–		–	
Predicted employment growth	–		1.61		1.70		–		–	
			(0.54)	**	(0.57)	**				
Predicted new work employment	–		0.07		0.06		–		–	
			(0.04)		(0.05)					
Predicted patenting	–		0.44		0.38		–		–	
			(0.47)		(0.51)					
Other city–year characteristics	–		–		X		–		–	
N (city–year observations)	1,089		1,089		1,089		726		726	
F (additional regressors)	18.1	**	3.56	*	1.74	*	–		–	
R^2	0.992		0.990		0.990		0.986		0.987	
Lagged age structure	–		–		–		X		X	
Lagged industry structure	–		–		–		X		–	
Lagged city characteristics	–		–		–		–		X	
College grad. share, partial–R^2	–		–		–		0.17		0.47	
Industry variety, partial–R^2	–		–		–		0.04		0.56	
p (overidentification)	–		–		–		0.91		0.87	

Notes: **–Significant at the 99% level of confidence; *–95% level; †–90% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression (see Tables 4 and 5). Each regression includes city (G=363) and year (T=3) effects. All initial city-year characteristics are measured in the census year previous to the first-step estimate. F-value for test that: (1) year-state effects are jointly zero, (2) effects of predicted city-year characteristics are jointly zero, and (3) effects of other city-year characteristics are jointly zero. Other city-year characteristics are employment shares in six major industry and nine major occupation groups. Instruments in regressions (4) and (5) are 20-year lags of: (lagged age structure) share of city population in 3 age categories, (lagged industry structure) the 3 predicted city-year characteristics used in regressions (2) and (3), and (lagged city characteristics) the four main explanatory city-year variables.

Table 8: New work location and other sources of city-year variation

	(1a)	(1b)	(1c)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Estimated city component	*Time-varying coefficients*			*By city population size (in 1970)*				*By industry*				*By Census division*		
of new work, from first step	*1980*	*1990*	*2000*	*<125k*	*<250k*	*<750k*	*>750k*	*Mfg.*	*Trd. svc.*	*Trn. & trd.*	*Oth. svc.*	*Ne. & Mw.*	*West*	*South*
College graduate share	2.2	3.3	6.5	11.3	10.2	15.8	16.2	13.1	42.4	2.2	1.7	5.9	13.5	14.8
(< h.s. share omitted)	(2.5)	(2.0) †	(1.6) **	(4.0) **	(3.8) **	(3.2) **	(4.3) **	(4.0) **	(4.7) **	(2.0)	(1.6)	(2.7) *	(3.8) **	(2.5) **
High school graduate share	-2.3	0.4	1.8	-1.9	-4.5	-2.8	-0.8	-4.5	-9.1	-1.1	-1.8	-5.4	-3.0	-1.8
	(0.8) **	(0.8)	(0.9) †	(2.0)	(2.0) *	(1.8)	(2.5)	(2.3) *	(3.0) **	(1.1)	(0.9) *	(2.0) **	(3.1)	(1.3)
Industry variety	7.4	6.8	6.6	6.7	8.7	1.8	6.6	12.8	9.3	8.4	-2.5	9.6	2.8	10.8
	(4.6)	(2.3) **	(4.2) †	(5.9)	(3.4) *	(4.1)	(7.0)	(5.6) *	(6.3)	(3.6) *	(2.4)	(3.7) **	(8.5)	(2.7) **
Log population density	-0.08	0.00	0.06	-0.14	0.49	0.37	-0.13	0.74	0.77	0.46	-0.15	-0.15	-0.21	0.40
	(0.10)	(0.10)	(0.11)	(0.19)	(0.38)	(0.35)	(0.24)	(0.36) *	(0.46) †	(0.19) *	(0.12)	(0.34)	(0.23)	(0.21) †
N (city–year observations)	–	–	1,089	312	321	267	189	1,089	1,089	1,089	1,089	460	204	425

Notes: **–Significant at the 99% level of confidence; *–95% level; †–90% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression (see Tables 4 and 5). Columns (1a), (1b), and (1c) report time-varying coefficients from a single regression. Each regression includes city (G=363), year (T=3), and census region × year (9×3) effects. All initial city-year characteristics are measured in the census year previous to the first-step estimate. Regression (7) includes workers employed in communication, finance, insurance, and business services. Regression (8) includes workers employed in transportation and wholesale trade. Regression (9) includes workers employed in all other service sectors. Regression (10) includes cities in the Northeast and Midwest U.S. Census divisions.

Table 9: New work location and alternate measures of new work, city education, and industry variety

		(1)	(2)			(3)	(4)			(5)	(6)
Estimated city comp.		*base*	*alt.*			*base*	*alt.*			*base*	*alt.*
of new work	*means*	*n.w.*	*n.w.*		*means*	*n.w.*	*n.w.*		*means*	*n.w.*	*n.w.*
College graduate	0.14	11.3	6.4	Log total coll.	10.6	2.1	1.7	Log total coll.	10.6	1.9	1.6
share	(0.06)	(1.6) **	(1.2) **	graduates	(1.2)	(0.3) **	(0.2) **	graduates	(1.2)	(0.3) **	(0.2) **
High school	0.49	-3.6	-5.0	Log total h.s.	11.9	-2.3	-1.9	Log total h.s.	11.9	-2.2	-1.9
graduate share	(0.10)	(0.8) **	(0.6) **	graduates	(1.1)	(0.2) **	(0.2) **	graduates	(1.1)	(0.2) **	(0.2) **
–1 x (Herf. index of	-0.06	10.4	8.4	Log industry	-0.02	5.4	3.4	Log industry	-1.64	0.16	0.12
industry empl.)	(0.02)	(2.2) **	(2.1) **	count	(0.03)	(1.0) **	(1.0) **	count > *1k*	(0.88)	(0.05) **	(0.05) *
Log pop. density	5.0	-0.14	0.31	Log pop.	5.0	0.09	0.55	Log pop.	5.0	0.07	0.57
	(1.3)	(0.13)	(0.12) **	density	(1.3)	(0.21)	(0.22) *	density	(1.3)	(0.15)	(0.22) **
F (city chars.)		39.8 **	48.1 **			34.0 **	33.9 **			29.2 **	29.7 **

Notes: **–Significant at the 99% level of confidence; *–95% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression (see Tables 4 and 5). Each regression includes city (G=363) and year (T=3), effects. All initial city-year characteristics are measured in the census year previous to the first-step estimate. Alternate new work measure in regressions (2), (4) and (6) uses *census-rules* definition in 2000 and *1991-update* definition in 1990.

Table 10: New work location and measurement of occupation

Estimated city component of new work, from first step	*Occ.–year title share*		*Occ.–year employment*		*Occ.–year number of titles*			*Occ. censoring*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	3^{rd} quintile	*5^{th} quintile*	*1^{st} quintile*	*5^{th} quintile*	*1^{st} quintile*	*3^{rd} quintile*	*5^{th} quintile*	*not allocated*
College graduate share	0.51	46.3	12.3	-5.1	10.3	25.5	-2.8	12.2
(< h.s. share omitted)	(0.33)	(5.5) **	(2.5) **	(0.9) **	(3.2) **	(3.4) **	(1.5) *	(1.7) **
High school graduate share	0.16	-5.4	-0.56	-0.39	-1.6	-6.3	0.90	-3.5
	(0.16)	(3.4)	(1.41)	(0.56)	(1.9)	(1.9) **	(0.82)	(0.8) **
Industry variety	0.91	31.9	11.0	0.24	-6.4	10.4	8.1	10.5
	(0.78)	(5.4) **	(5.2) *	(1.86)	(7.2)	(6.3)	(3.5) *	(2.4) **
Log population density	-0.01	-1.32	-0.36	-0.23	0.56	-0.26	-0.42	-0.14
	(0.03)	(0.50) **	(0.21) *	(0.08) **	(0.33) *	(0.26)	(0.11) **	(0.13)
M (workers, in 000s) (first step)	1,310.9	1,675.4	1,784.5	1,721.7	909.1	1,448.3	2,843.5	7,867.9
# occupation–years (first step)	176	280	1036	23	310	251	278	1,239
Mean new title share, 2000	1.1	21.8	4.6	1.4	7.4	6.7	1.4	5.5
	(0.5)	(20.2)	(10.0)	(1.4)	(14.9)	(16.2)	(2.4)	(13.0)
Mean new title share, 1990	13.1	17.6	7.3	7.5	6.7	10.4	7.0	7.8
	(10.9)	(15.0)	(12.8)	(6.5)	(16.1)	(16.2)	(7.1)	(12.6)
Mean new title share, 1980	3.9	28.4	8.2	8.7	4.4	11.5	8.6	8.3
	(1.2)	(14.0)	(13.2)	(4.6)	(14.7)	(14.5)	(6.8)	(12.8)

Notes: **–Significant at the 99% level of confidence; *–95% level; †–90% level. Robust standard errors in parentheses, adjusted for clustering on city (G=363). Each numbered column is a separate WLS regression (see Tables 4 and 5). Each regression includes city (G=363) and year (T=3) effects. N=1,089 in each regression. All initial city-year characteristics are measured in the census year previous to the first-step estimate. In regressions (1) and (2), sample is divided by the new title share ν in occupation-years (for the first and second quintiles, this share is zero). Mean values of the new title share ν are calculated at the U.S. Census detailed occupation level. In regressions (3) and (4), sample is divided by total employment in occupation-years. In regressions (5), (6), and (7), sample is divided by the number of titles in occupation-years.

A Data appendix

A.1 Occupation titles

The *DOT* comparisons involve the third (1965), fourth (1977), and revised fourth (1991) editions, as well as two editions of the *Conversion Table of Code and Title Changes* (U.S. D.O.L., 1979 and 1991). The 1991 *DOT* and the 1991 *Conversion Table* are easily obtained in electronic format (ICPSR study number 6100). The 1977 *DOT* and 1979 *Conversion Table* were processed for this paper, from hard-copy versions, using optical character recognition. I did not directly use the 1965 *DOT*, as all necessary information is available in the 1977 *DOT* and the 1979 *Conversion Table*. In both 1977 and 1991, I consider only base titles, excluding master titles (which encompass many individual titles) or alternate titles (which are alternate names for the base titles).

An advantage of the *DOT* series is that the nine-digit codes assigned to titles remain consistent over time, except for changes that are universally covered by the *Conversion Tables*.

Section 3 of the 1991 *Conversion Table*, "NEW DEFINITIONS IN THE *DOT*," lists 830 new base titles. I merge these into the full list of 12,741 *DOT* base titles. For each title, the 1991 *DOT* also reports corresponding (three-digit) 1980 standard occupational classification (SOC) detailed occupations. I match 1980 SOC detailed occupations to 1980 census detailed occupations, using technical documentation from the U.S. Census (1980a). Census detailed occupations are the same in the 1980 and 1990 census. I then collapse the data by 1990 census detailed occupation.

Section 3 of the 1979 *Conversion Table*, "CODES AND TITLES OF NEW DEFINITIONS," lists 2,292 new titles, of which 1,152 are new base titles. I match the latter category to the full list of 12,695 *DOT* base titles. Unfortunately, I do not have data on corresponding (three-digit) detailed occupations for 1977 titles. Instead, I rely on information from the 1991 *DOT* and *Conversion Table*. Crucially, section 1 of the 1991 *Conversion Table*, "OCCUPATIONAL CODE AND/OR TITLE CHANGES," and section 2, "OCCUPATIONAL DEFINITIONS DELETED FROM OR COMBINED WITH ANOTHER IN THE REVISED FOURTH EDITION," allow me to map nearly every 1977 title into a 1991 title. I then use the 1980 SOC detailed occupations, reported in the 1991 *DOT*, combined with the SOC-census detailed occupation crosswalk, to collapse the data by 1980 census detailed occupation.

A disadvantage of the *DOT* series is that not all census detailed occupations appear in the *DOT*-SOC-census crosswalk. In the 1980 and 1990 IPUMS, 4.4% and 2.0%, respectively, of sampled workers (age 16-70 with non-missing detailed occupation, excluding Alaska and Hawaii) report detailed occupations that do not have new work information from the *DOT*. These workers are excluded from the analysis.

I construct an alternate version of the 1991 new title list based on discussion in the 1991 *DOT* that indicates that titles added during the last of a series of inter-revision reviews were most likely to represent new work that appeared in the 1980s. The *DOT* flags 89 titles that were added during this last update in 1991; I call this alternate set of new titles the *1991-update* definition of new work.

I identify new work in 2000 by comparing titles in the 1990 and 2000 *Classified Index of Industries and Occupations*. I use two algorithms to identify new titles. In the first method, I initially perform a string match, allowing for typographic differences (as advised by Scopp, 2003), followed by a manual review of the remaining occupation titles. The final list contains 840 new titles. Titles in this list include "web designer," "data recovery planner," "pharmacoepidemiologist" (studies drug outcomes in large populations), "dosimetrist" (determines proper doses in radiation therapy), "AIDS counselor," and "polymerization kettle operator" ("controls reactor vessels to polymerize raw resin materials to form phenolic, acrylic, or polyester resins," according to the *DOT*).

An alternate, completely orthogonal algorithm relies on detailed internal census documents, obtained from Tom Scopp and Marisa Tegler at the U.S. Census Bureau. These data report, for each 2000 title, an indicator for whether it was new to the 2000 classification system, the corresponding 1990 detailed occupation code, and further information about why changes, if any, occurred. I classify a title to be new under the *census-rules* definition if it is in the intersection of the following sets: (a) the title is new in 2000, (b) the title does not have a corresponding 1990 detailed occupation code, (c) the title is not an alternate title, and (d) comments do not indicate that the new title was added due to a split, adjustment, or coding error of a previous title. The *census-rules* definition contains 814 new titles.

In the text I cite the appearance, in 2000, of detailed occupation 111, *network systems and data communication analysts*, as evidence that new occupations indeed followed actual innovations. Another example is detailed occupation 104, *computer support specialists*, which contains workers who provide technical assistance to users of desktop computers and database software. Desktop computers, such as the IBM PC and Apple //, and commercial database software, such as Oracle and DB2, did not widely appear until the mid 1980s. Clearly, new types of work appeared around this time to support these new innovations. Given the decennial nature of the census, it seems reasonable that they were first cataloged for census 2000.

An earlier version of this paper (Lin, 2007) used a series of census technical papers (1968, 1972, and 1989) to try to identify new work at the three-digit detailed occupation level. Though the *DOT* allows greater precision, it is reassuring that many of the detailed occupations identified by both methods are the same.

These data are available by request.

A.2 Worker- and city-level data

I use three census extracts from the IPUMS (Ruggles et al., 2009): the 1980 5% state sample, and the 1990 and 2000 1% unweighted samples. These choices are dictated by the availability of both geographic and occupational information. Only the 5% state sample in 1980 reports consistent public-use microdata area of residence (conspuma); the 2000 5% sample reports fewer occupational categories. I exclude Alaska and Hawaii, people under 16 and over 70, and workers without an identified occupation. (Occupation is reported for workers age 16 or greater who have worked within the previous five years, excluding workers new to the labor force who have yet to secure a first job.) The number of observations in each year is 5,909,772 (1980), 1,329,710 (1990), and 1,562,904 (2000).

To construct a city panel using information spanning 1970 to 2000, I create county-based aggregate geographic units that can be consistently identified, based on county group in the 1970 IPUMS (cntygp97) and consistent public-use microdata area (conspuma) in the 1980, 1990, and 2000 IPUMS. (This step requires county composition files, available from IPUMS.) In order to better capture local labor markets, I further aggregate units within the same metropolitan area, using 2003 core-based statistical areas (CBSA), county groupings defined by the U.S. Office of Management and Budget (U.S. Census, 2003). Each conspuma/cntygp97 unit is uniquely assigned to one of 363 CBSA-based aggregates, which I refer to as "cities" throughout the paper. In cases where a county group overlaps multiple metropolitan areas, I assign the county group to the metropolitan area containing most of its population in 1990.

City-level data come from a variety of sources. I use the 1970, 1980 and 1990 IPUMS to compute some statistics related to industry composition, in particular the predicted employment growth, predicted new work employment, and predicted patenting indexes used in Table 7. I compute city population, land area, aggregate educational attainment, and other variables by aggregating county-level data from three editions of the U.S. Census' *City and County Data Book* (1972, 1983, and 1994), assembled by Haines and the ICPSR (2004). In particular, all county averages are weighted by county population, so that (for example) computed city population density is as experienced by the average person, rather than the average areal unit. Finally, I compute industry variety using summary files from the U.S. Census of Population and Housing (1970, 1980b, and 1990a). Data on industry employment by county are taken from the fourth summary file—Tables 58 and 62 in 1970, Tables 57 and 58 in 1980, and Tables 61 and 63 in 1990. I then aggregate the county information from the summary tables in the same way as the county data books.

B Theory appendix

Here, I outline a static, general equilibrium, economic geography model that helps interpret the observed spatial distribution of new work as coming from economies of density. Locations with initial stocks of educated workers and industry variety attract more new work. In the formal model presented, this is because of a backward linkage, as educated workers demand more products that use new work intensively, but do not consider the effect on local prices of their location decision. Similarly, a greater variety of goods-producing plants demand more new work inputs. The static model is closely related to Helpman (1998); Redding and Sturm (2008) use a similar strategy to simulate the before-and-after economic geography of

German division.

As emphasized in the main text, a number of micro-foundations can be used to generate agglomeration economies. The main point here is that adaptation to new technologies is a potential form of agglomeration economies distinct from, say, benefits from sharing indivisible goods; other micro-foundations could be employed to make the same point.

B.1 Setup, preferences, and technology

Consider two locations, labeled 1 and 2. Each location is endowed with a non-traded good, supplied inelastically across locations, with quantities h_1 and h_2. A population of educated workers L, mobile across locations (l_1 and l_2), supplies labor inelastically to traded goods production.[21] They consume services h and differentiated varieties of the traded good x. There are N total varieties of the traded good, each produced by a separate firm. Further, each variety is produced using a distinct production activity. Therefore, there is a one-to-one relationship between the varieties of traded goods, the number of plants, and the number of types of work. Traded goods production is footloose, constrained by $n_1 + n_2 = N$.

Representative educated-worker utility is $U = h^{1-\mu}[(\int_{j=0}^{N} x_j^{\alpha})^{1/\alpha}]^{\mu}$ where $\sigma \equiv 1/(1-\alpha)$ is the constant elasticity of substitution between traded goods varieties, assumed greater than 1. Let $\mu \equiv N/(N+\delta)$, $\delta > 0$, so that the expenditure share devoted to traded goods increases with the number of varieties.[22]

Production of each variety of traded good is subject to plant-level scale economies, modeled as a fixed cost f in terms of educated labor l. Let β be the unit cost in educated labor, then $l = f + \beta x$, where both f and β are assumed greater than zero. After production, there are iceberg transport costs. For each variety, $t > 1$ units must be shipped for 1 unit to arrive in the other location. Location 1 residents pay p_1 for every locally produced variety but tp_2 for varieties imported from location 2.

B.2 Comparing equilibria

Profit maximization implies that relative plant prices of the traded good (p_1/p_2) must be equal to relative wages of skilled labor ($w \equiv w_1/w_2$). By free entry of firms, equilibrium output for each variety is the same. It follows that educated-worker demand is equal across varieties; therefore, $n_1/N = l_1/L$.

In equilibrium, where are production activities concentrated? Define the share of production activities located in region 1 as $v \equiv n_1/N$. Two equilibrium conditions can be derived. The first relates v, the location of production activities, to w, relative prices and wages. This condition provides a unique solution to relative

[21]For brevity, less-educated labor are omitted. For now, assume they are immobile and consume and produce another type of good.

[22]This is a key assumption: as N expands, the expenditure share devoted to housing $(1-\mu)$ falls. Without it, growth in varieties scales production in each region, failing to deepen agglomeration. Is this assumption plausible? Bils and Klenow (2001) find that variety growth leads to lower expenditure shares on non-innovating sectors (i.e., h). Note, too, that alternative utility specifications can also generate flexible expenditure shares–a CES aggregator over h and x, for example.

prices and wages w for each distribution of production activities v.[23]

$$1 = \frac{vw^{1-\sigma}}{vw^{-\sigma} + (1-v)t^{1-\sigma}}[\mu + (1-\mu)(v + \frac{1-v}{w})] + \frac{(1-v)(tw)^{1-\sigma}}{v(tw)^{1-\sigma} + 1 - v}[\frac{\mu}{w} + (1-\mu)(v + \frac{1-v}{w})] \tag{A}$$

Since skilled labor is mobile across regions, a second equilibrium condition requires that household utility is equal across regions.

$$u = 1 = \left(\frac{h_1}{h_2}\frac{1-v}{v}\right)^{1-\mu}\left(\frac{\mu w + (1-\mu)(vw + 1 - v)}{\mu + (1-\mu)(vw + 1 - v)}\right)^{\mu}\left(\frac{vw^{1-\sigma} + (1-v)t^{1-\sigma}}{v(tw)^{1-\sigma} + 1 - v}\right)^{\mu/(\sigma-1)} \tag{B}$$

Equilibrium is fully characterized by these two conditions, which determine two endogenous variables, v and w, in terms of parameters μ, σ, t, and h_1/h_2. I solve for equilibrium values of v and w numerically. I first calculate relative utility $u \equiv u_1/u_2$ for the entire range of values of v, the share of production activities in region 1. In equilibrium, it must be that $u = 1$, or else that all activity concentrates in one region (and $u = 0$ or $u = \infty$).

Figure A graphs relative location 1 utility (Equation B) for the marginal mobile skilled worker, all else equal, for small and large numbers of available activities. First, the dashed line represents utility when there are a relatively small number of available products and types of work, and it is meant to characterize a before-innovation state. The dotted line represents utility under a large number of available products and types of work, meant to capture the introduction of new work. The marginal worker's utility is high when location 1 contains few educated workers or little production variety, since the non-traded good h is cheap; it is low when location 1 is crowded, due to congestion costs. Over some range in between, utility rises with the concentration of production activities, given parameter values that generate strong enough linkages between the locations of consumption and production. In general, these are features of any model featuring agglomeration economies: a race between agglomeration economies and congestion costs. In order to be consistent with the existence of cities, over some range, the density of economic activity is utility-improving for the marginal mobile household.

An example of a stable spatial equilibrium in the before-innovation case is the intersection of the dashed line with $u = 1$ at point A. (See Helpman, 1998, for discussion of parameter value selection and multiple equilibria.) Here, location 1 contains more educated workers and industry variety; because of the demand linkage, more types of work are also used in location 1. Given an initial equilibria at point A, what is the likely spatial distribution of new activities after innovation (i.e., the introduction of new products and new types of work)? The likeliest answer is point B (assuming reasonable dynamics, such as slow adjustment to innovation). That is, new types of production activities appear disproportionately in the location that

[23] The relationship between v and w does not depend on either f nor β; they only scale the number of varieties and the level of production output.

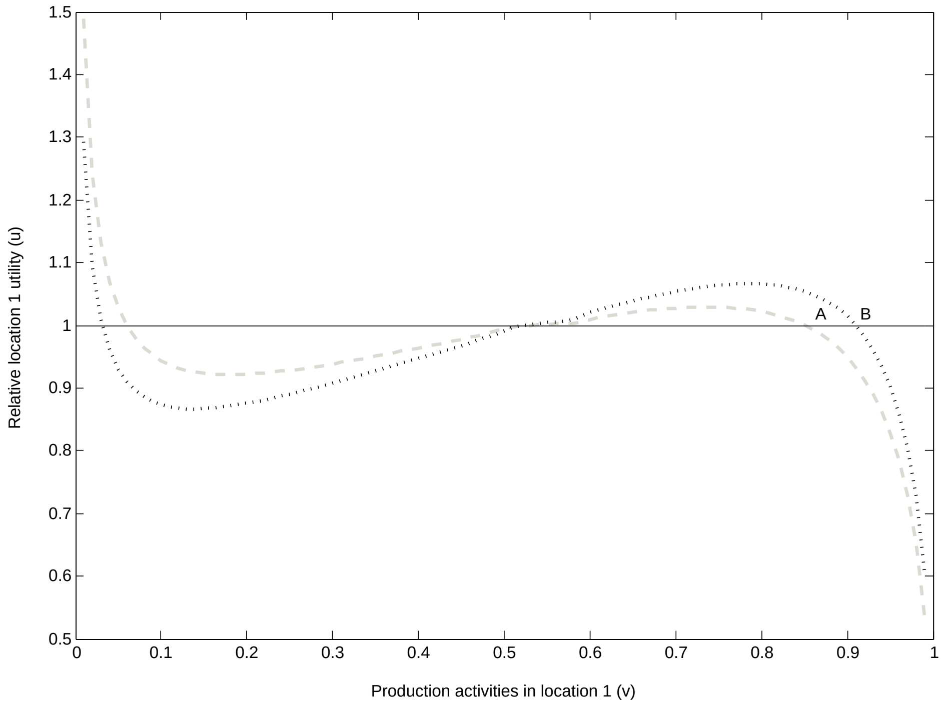


Figure A: Location 1 utility and concentration of production activities

initially produced a greater variety of products with a greater stock of educated workers. The demand linkage provides incentive for the marginal newly appearing activity to agglomerate in location 1. This is the main empirical prediction.

The model admittedly contains some unrealistic features. In particular: first, only educated workers demand, and are employed in the production of, traded goods. Including less-educated workers that demand fewer traded goods and are employed less intensively in the production of traded goods should not affect the main results, but may add some interesting interactions. For example, if (immobile) less-educated workers compete for the non-traded good, but do not consume or produce traded goods intensively, then a greater concentration of less-educated workers will deter new work from that location, because of congestion costs. Second, the effect of industry variety is technically indirect: locations with more industry variety employ more educated workers, which affects the location of new work. One way to introduce a direct, independent effect of industry variety would be to require a variety of new work inputs, supplied under imperfect competition, in the production of each variety of traded good (cf. the vertically linked industries model in Venables, 1996). This modification would add a second set of linkages; locations with greater variety in traded-goods production, and thus greater factor demand, would then be more advantageous places to locate new work inputs.