Should the PC be Considered a Technological Revolution? City Level Evidence from 1980-2000

Paul Beaudry
University of British Columbia and NBER

Mark Doms
Federal Reserve Bank of San Francisco

Ethan Lewis *
Dartmouth College

October 2008

*The authors would like to thank Daron Acemoglu, David Autor, David Green, Larry Katz, Thomas Lemieux, and seminar participants at UBC, LSE, UCL, the NBER Summer Institute, and the New York Federal Reserve for very helpful discussions and to Meryl Motika for research assistance.

— Abstract —

The notion that the introduction and diffusion of IT capital may have constitute a technological revolution is widespread. This paper uses data across US cities to see whether the links between PC adoption, educational attainment and returns to skill conform to the prediction of a model of technological revolutions. Our simple neo-classical model implies, among other things, that a shift in technological paradigm will increase the returns to skill and that the adoption of new technology will be greatest where skill is most abundant and where the price of skill is initially low. Moreover the model implies that the returns to skill should temporarily become insensitive to increases in supply since, instead of putting downward pressure on returns, an increase in skill after a major technological innovation acts only to accelerate the transition to the new technology. Our main finding is that the cross-city patterns over the period 1980-2000 conform closely to these and other predictions, suggesting that the era of diffusion of PCs may well deserve its recognition as a technological revolution.

Key Words: Biased Technological Change, Relative Wages, Education, Technology Diffusion

JEL Class.: E13, O33, J30

Paul Beaudry	Mark Doms	Ethan Lewis
Department of Economics	Federal Reserve	Economics Department
University of British Columbia	Bank of San Francisco	Dartmouth College
997-1873 East Mall	101 Market Street, MS 1130	
Vancouver, B.C.	San Francisco, CA 94105	Hanover, NH 03755
Canada, V6T 1Z1		
and NBER		
paulbe@interchange.ubc.ca	mark.doms@sf.frb.org	ethan.g.lewis@dartmouth.edu

1 Introduction

There is a strong consensus regarding the long run relationship between education, the returns to skill and technological change. The consensus view is that technological change tends to be skilled biased and push up the return to education, while increases in the supply of educated workers puts downward pressure on the return.[1] The combined effect of these two forces offers a simple explanation for why the returns to education showed no marked trend over the full span of the 20th century as technological change and the supply of skill acted as offsetting forces.

Although the above view is a reasonable description of the long run, it does not follow that the relationship between education, the returns to skill, and technological change is as simple or as stable in the short or medium run. Even within the confines of a neo-classical theory, it is possible for an increase in education to have a very different effect on its return during a period of revolutionary technological change than over periods of incremental technological change. For example, in periods of shifts in technological paradigm, we will show that comparative advantage principles imply that the returns to skill should increase most where skills are most abundant, even though in the long run increases in skill should eventually cause a fall in its return. We will refer to such a possibility as an "episodic" view of economic progress, as it implies that the behavior of the economy varies over time depending on the phase of the technological cycle.

In this paper we use data from US cities to examine the links between technological adoption, educational attainment, and the returns to skill during a period of substantial technological change. We focus on the period 1980-2000, commonly referred to as the era of IT revolution, and therefore an ideal candidate period to evaluate the episodic view of the link between technological change and wages. In particular, our goal is to examine whether certain key economic relationships over this period depart from perceived long run patterns and instead conform to the predictions of a model which embodies the notion of paradigm shift. The model we consider is one where there are periodic arrivals of a major technological innovation, and the speed of adjustment to the new paradigm is endogenous and responds to price of skill. Since the model has a neo-classical structure, it has the property that an increase in educational attainment has the long run effect of reducing the returns to skill. However,

[1] See for example Author, Levy, and Murnane (2003), and Autor, Katz, and Kearney (2006), Goldin and Katz (2008). A related but distinct literature examines the extent to which technology adoption is affected by the relative supply of skilled labor. For example, see Comin and Hobijn (2004), Caselli and Wilson (2004).

in the medium run when the economy is transiting to a new technological paradigm, the relationship between educational attainment and its return is different from the long run relationship. Following the arrival of a new technology that raises returns to skill, the model predicts that increasing average educational attainment in many cases will not help to push down the return to skills. Instead, an increase in average educational attainment only increases the speed of transition to the new technology. One of the model's more surprising predictions is that the arrival of a new skill-biased technology will not lead to the returns to skill to be larger where it is adopted more intensively.

The main result of the paper is that the cross-city patterns of PC adoption, educational attainment, and the returns to skill conform closely to those implied by a model of technological revolutions. The model on which we base our empirical exploration, described in the next section, extends the technological adoption models of Caselli (1999) and Acemoglu (2006)[2] to local labor markets and examines the model's implications for how the supply of skill affects technology adoption and changes in its return.[3] We assess the model with data on PC use, skills, and returns to skill for a sample of 230 U.S. metropolitan areas over the main period of diffusion of the PC, from 1980 to 2000. The structure of the model stems from the observation that a new means of production is initially attractive only to localities facing particular configurations of factor prices. In such a situation, it will be optimal for one locality to adopt the new technology - if it has a comparative advantage in doing so - while in a different locality it may be optimal to maintain an old technology. The model predicts that it is localities with the highest educational attainment which adopts the new technology most aggressively and thereby witness the greatest rise in returns to skill. However, the rise in returns to skill as a result of technology adoption will not be so great so as to create a positive association between the supply of skill and the return to skill. Instead, after the introduction of the new technology, there will be a range of skill supplies over which the returns to skill do not vary.

[2]Also see Beaudry and Green (1998, 2003, and 2005).

[3]While models of endogenous technology adoption used here have generally played a secondary role in the literature on skilled biased technological change and the diffusion of PCs, there is a tradition in the economic history literature that advocates a similar approach for understanding periods of major technological change. For instance, Goldin and Sokoloff (1984) use a similar model to explain regional outcomes across the US during the industrial revolution. They find that factor price differences in 1830 between the northern and southern U.S. states (due to crop differences) help explain the differential patterns of industrialization, and that payments to those factors rose fastest in areas where technology was adopted most aggressively. The writing of Habakkuk (1962) also reflects similar ideas. Given this previous work, it is of interest to examine whether simple neoclassical principles applied to technology adoption offer a unified way of understanding many major technological episodes, including the most recent, the IT revolution.

Empirically, we begin by documenting wide and persistent differences in skill and technology use across a large sample of U.S. cities from 1980 to 2000, and we show that these differences do not arise from industry composition. Using our data and instruments, we find that it is in cities where high school educated workers are more abundant (and cheap) relative to college educated workers that PCs were adopted most intensely. In our data, we take the intensity of PC use as our main indicator of adoption of the new technology.[4] It is also these cities that experience the greatest increase in the returns to education. That is, it is cities that possess a more abundant supply of college educated workers that adopted PCs most intensively and saw the returns to college increase fastest. Moreover, we document that the downward slope across cities between the supply of skill and the return to skill which existed in 1980 dissipated by 2000.[5] However, as the model suggests, the high PC adopting cities are not observed to have higher returns to education in 2000 than their slower adopting counterparts. This observation contrasts with common intuition regarding the likely correlation between PC use and returns to education, but it is consistent with the endogenous technology adoption framework used here to model technological revolutions.

In our analysis, we take care to contrast the implications of our model with several alternative explanations of the same facts. Although some of our observations can also be explained by alternative mechanisms, we emphasize the difficulties other explanations face in reproducing them all. For example, one potential explanation for our observation that relative wages equalized across cities between 1980 and 2000 is that trade and labor mobility across cities increased during this time period, leading to a greater increase in the returns to education in cities with more educated labor. While this explanation cannot be completely ruled out, we show that the actual movement of people and industries across cities do not provide much support for it. Similarly, we show that the most common production function specifications

[4] Doms and Lewis (2006) examine a variety of factors related to technology diffusion and find that the most important factor in understanding the large variance in PC use across cities is the supply of human capital. The present paper extends Doms and Lewis (2006) by specifying and testing the mechanism that drives these differences in PC adoption.

[5] Our results may appear to conflict with more aggregate approaches, such as Katz and Murphy (1992), Krusell, Ohanian, Rios-Rull and Violante (2000) and Autor, Katz, and Kearney (2006), which suggest that there is a stable relationship between the supply and skill and its return over long periods of time, and hence the notion of technological revolution as used here may be misplaced. However, as shown in Card & Dinardo (2002) and Beaudry & Green (2005), when the aggregate US experience is split between a pre-1980 period and a post 1980 period, the relationship between the supply of skill and its return changes drastically. In particular, the relation changes in a direction consistent with the evidence presented here, as the downward effect of increased education on its return is generally insignificant in the post-1980 period. Moreover, our results are also consistent with those reported in Fortin (2006), where post 1980 state level variation in the aggregate supply of skill in not found to have any negative effects on the return to skill.

with PC-skill complementarity can explain some, but not all, of the cross-city patterns we document for the wages of high school and college educated workers.

The remaining sections of the paper are structured as follows. In section 2, we present a model of technology adoption and derive a set of implications regarding local level interactions between returns to education, changes in returns to education, and technology use. We begin by treating local supplies of skills as exogenous, then we extend the model to allow for migration across cities in response to differential skill returns and thereby endogenize difference in skill across cities. In Section 3 we discuss the data and section 4 contains our empirical results. Section 5 examines several alternative explanations for our results. The remaining section offers concluding comments.

2 A neoclassical model of technology adoption

Consider an environment at time t where firms have access to a set of technologies to produce a final good denoted by Y_t. The production of Y_t requires inputs Z_t, where these inputs can be organized in different ways to produce output, each of these alternative organizations corresponding to a different technology. If we parameterize the different technologies by $\theta \in \Theta$, then the production possibilities facing a firm can be represented by:

$$F(Z_t, \theta), \qquad \theta \in \Theta_t$$

where for each $\theta \in \Theta$, the production function is assumed to satisfy constant returns to scale and concavity. In this case, a price taking firm will aim to maximize profits by solving the following problem,

$$\max_{Z_t, \theta_t} F(Z_t, \theta_t) - w_t Z_t$$

where w_t is the vector of factor prices. For such an environment, standard techniques can be used to prove the existence of a competitive equilibrium, where a competitive equilibrium can be defined as a set of prices, allocations and technology choices, such that, given prices, allocations and technology choices are optimal, and markets clear.

Let us now consider the situation with a set of distinct markets, indexed by i. Each of these markets is assumed to have access to the same set of technologies. We will begin by assuming that these markets differ in terms of the supply of at least a subset of the factors Z. The question we ask is how do the different markets react to a change in the set of choices, that is, a change in Θ_t. Obviously, the answer to this question depends on the nature of the change in Θ. In particular, given the time period that interests us, we want to examine the effects of having Θ extend to include a more skilled biased technology relative to the pre-existing choices. To this end, we focus on the case where initially there is only one dominant technology used across all markets. This technology uses as inputs skilled labor S, unskilled labor U and traditional capital K. The market for skilled and unskilled labor is initially assumed to be a purely local market, with exogenously fixed local supplies. Later will we will extend the analysis to the case where workers are mobile across markets, but where markets differ in terms of amenities and congestion. Throughout we assume that the market for K is a common market, where firms from all localities can rent the capital at the rate r^k. Finally, for clarity of presentation and ease of comparison with the existing literature, the pre-existing technology is assumed to have the following functional form:[6]

$$F^T(K,S,U) = K^{1-\alpha}[aS^\sigma + (1-a)U^\sigma]^{\frac{\alpha}{\sigma}}, \quad 0 < \alpha < 1, \quad 0 < a < 1, \quad 0 < \sigma < 1$$

In this environment, the initial returns to skill will differ across markets. In particular, the ratio of the market specific skilled wage w_i^S to the unskilled wage w_i^U will be given by:

$$\frac{w_i^S}{w_i^U} = \frac{aS_i^{\sigma-1}}{(1-a)U_i^{\sigma-1}}$$

where S_i and U_i represent the quantities of skilled labor available in market i. Given the form of the production function, it is possible for the wage of skilled workers to be less than that of unskilled workers. We therefore impose Assumption 1 to insure that we are focusing on economies where skilled workers are paid more than less skilled workers.[7]

Assumption 1: $\frac{S_i}{U_i} < (\frac{a}{(1-a)})^{\frac{1}{1-\sigma}}$

Now suppose that at a point in time, say at $t = 0$, a new technology becomes available. This technology has two characteristics that differentiate it from the traditional technology. First, it uses a different form of capital, which we denote as PC capital, and PC capital is assumed to be available on a common market at rental rate r^{PC}. Second, the new technology

[6] The analysis can be easily extended to cases where $\sigma < 0$.

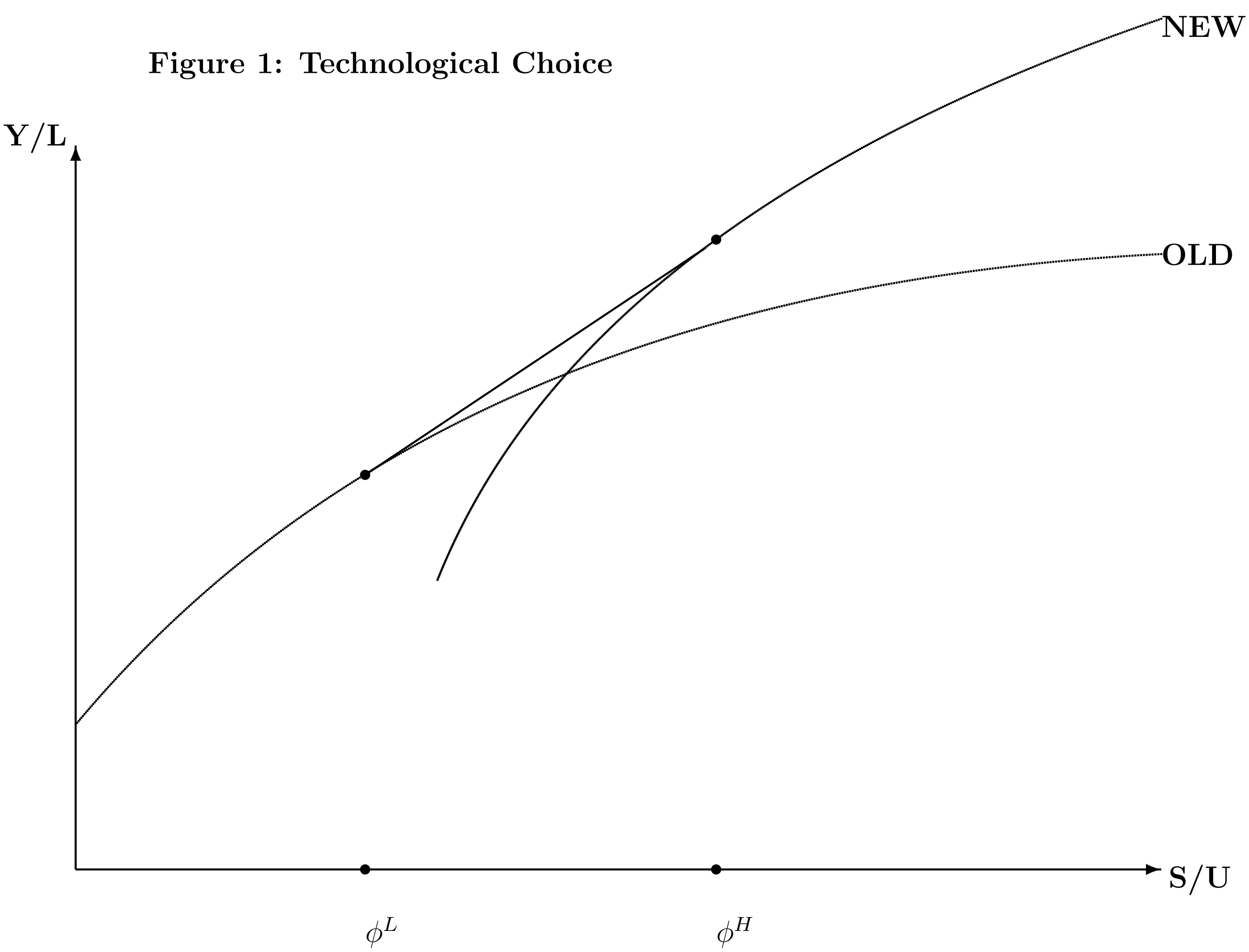
[7] If we assume that skilled workers can fill jobs designated for unskilled workers, this would guarantee that unskilled workers are never paid more than unskilled. For some of our results, this assumption could replace Assumption 1. However, certain of our results require the stronger restriction that the returns to skill are positive, which is implied by Assumption 1.

is assumed to be skilled biased relative to the old technology in the sense that at common factor prices, the new technology uses skilled labor more intensively (i.e., has a higher ratio of $\frac{S}{U}$) than the traditional technology. These features are captured in the following functional form for the new technology:

$$F^N(PC, S, U) = PC^{1-\alpha}[bS^\sigma + (1-b)U^\sigma]^{\frac{\alpha}{\sigma}}, \qquad a < b < 1$$

It is important to notice that the new technology does not necessarily dominate the old technology in the sense of producing more output at any input combination. In effect, for certain rental rates of capital r^K and r^{PC}, the old technology is more productive than the new technology when used with a small fraction of skilled workers, while the new technology is more productive (higher output per worker) when used with a high fraction of skilled workers. This property is depicted in Figure 1.

Figure 1: Technological Choice



It is easy to see from Figure 1 that, when faced with the choice between the new and the old technology, localities with high ratios of skilled to unskilled workers will want to adopt the new technology, while those with low levels of skill may want to maintain the old technology. We refer to the situation where there is this non-trivial choice between the two technologies as a **technological transition**. A key aspect of our analysis is to highlight the different relationship between factor prices and factor supplies during a technological transition in comparison to periods where economies do not face such choices. Note that the arrival of a new technology does not necessarily give rise to a technological transition. In our formulation, if the price of PC capital is very high, then the old technology dominates the new one regardless of the local supply of skill. As shown in the appendix, the condition $r^{PC} > r^K(\frac{b^{\frac{1}{1-\sigma}}+(1-b)^{\frac{1}{1-\sigma}}}{a^{\frac{1}{1-\sigma}}+(1-a)^{\frac{1}{1-\sigma}}})^{\frac{(1-\sigma)\alpha}{\sigma(1-\alpha)}}$ guarantees that no economy (which satisfies Assumption 1) will want to adopt the new technology. In contrast, if the price of PC capital is very low, then the new technology dominates the old technology. This arises when $r^{PC} < r^k(\frac{1-b}{1-a})^{(\frac{\alpha}{\sigma(1-\alpha)})}$, in which case the new technology would be adopted by all profit maximizing firms regardless of local market conditions. Hence, a technological transition corresponds to a temporary period that arises after the initial development of technology and lasts until other developments allow the new technology to become dominant. Since we want to focus on implications of a technological transition, we pursue our analysis under Assumption 2.

Assumption 2: The rental price of PC satisfies $r^k(\frac{1-b}{1-a})^{(\frac{\alpha}{\sigma(1-\alpha)})} < r^{PC} < r^K(\frac{b^{\frac{1}{1-\sigma}}+(1-b)^{\frac{1}{1-\sigma}}}{a^{\frac{1}{1-\sigma}}+(1-a)^{\frac{1}{1-\sigma}}})^{\frac{(1-\sigma)\alpha}{\sigma(1-\alpha)}}$

Under Assumption 2, some localities will find it optimal to adopt the new technology while others will not. However, it is not the case that localities will either fully adopt the old or new technology. Instead, the adoption decision is characterized by three regions delimited by skilled to unskilled labor ratios, with the middle region involving localities where the new and old technology co-exist. As shown in the appendix, there exist critical values of skill ratios ϕ^L and ϕ^H ($0 < \phi^L < \phi^H$) such that if a locality is characterized $\frac{S_i}{U_i} < \phi^L$, then it maintains the old technology. If $\frac{S_i}{U_i} > \phi^H > \phi^L$, then the locality switches completely to the new technology. Finally if $\phi^L < \frac{S_i}{U_i} < \phi^H$, then both technologies co-exist in a competitive equilibrium, with the fraction of production done using the new technology being an increasing function of $\frac{S_i}{U_i}$. We refer to a locality where the technologies co-exits as a locality that is **experiencing the technological transition**.

Proposition 1 addresses the link between *PC* use and local skill supply. Since *PC* capital is used intensively in the new technology, it follows that the quantity of PCs per worker used in a locality is a monotonically increasing function of the ratio of skilled to unskilled workers.[8] This forms the basis of Proposition 1.

Proposition 1: After the arrival of a PC-based, skilled-biased technology, the ratio of PCs per worker will be an increasing function of a locality's ratio of skilled to unskilled workers.
Proof: The proofs of all propositions and corollaries are given in the appendix.

Proposition 1 indicates that skill biased technologies are adopted most aggressively by localities in which skill is relatively abundant, and therefore the observable aspects of the technology – such as here PC capital — are most prevalent in localities with more skill. This implication is the focus of Doms & Lewis (2006). Here, we want to go further and derive a set of additional implications in order to more closely examine the relevance of a biased technology adoption model for understanding differences in outcomes across localities. To this end, we first extend Proposition 1 slightly and derive a corollary that captures the incentive mechanism that leads to the different adoption decisions. Note that from an individual firm's perspective, the differential adoption decisions across localities must reflect different incentives induced by factor prices. In localities with initially high ratios of skilled to unskilled labor, the relative price of skilled labor is initially low (prior to the availability of the new technology), favoring the adoption of a technology which uses skill intensively. This implication is expressed in Corollary 1.

Corollary 1: The ratio of PCs per worker is a decreasing function of a locality's initial ratio of skilled to unskilled wages.

Proposition 1 and Corollary 1 focus on the effects of local market conditions on adoption decisions. We now want to change perspective and examine the more intriguing aspect

[8] We have presented what we believe to be the simplest model that delivers the propositions which we investigate empirically. However, on some dimensions, it is certainly too simplistic. For example, the particular formulation implies that adding more unskilled workers to a labor market while keeping the number of skill workers constant would lead to a decrease in the number of PCs. This implication is not robust as it results from the extreme assumption that there is no possibility of using PCs in the traditional technology. A slightly generalized formulation, which reverses this implication, is one where the traditional technology uses both structures and equipment. Assume that at time $t = 0$, the PC becomes available. The PC can then be used either as a substitute for equipment in the traditional technology, or it can be used in a new form of organization which is both skill-biased and uses PC more intensely than the traditional technology (in the sense that, as given factor prices, it uses a greater number of PC per worker). This later form of work organization is what we envision as the new technology. It can be easily verified that this alternative formulation is consistent with all of the propositions presented in the paper, but it does not imply that the number of PCs used would decrease in response to an increase in unskilled labor.

of the technological transition which is how the arrival of the new technology affects the relationships between factor prices and supply. In particular, we first want to emphasize how changes in the return to skill, as expressed by the change in the $\log(\frac{w_i^S}{w_i^U})$, varies across localities faced with the same technology options. This is captured in Proposition 2 and Corollary 2.

Proposition 2: The arrival of the skilled biased technology causes the returns to skill to increase most in localities where skill is abundant.

The content of Proposition 2 can be obtained by deriving the relationship between the return to skill and the supply of skill before and after the arrival of the new technology, and taking the difference between the two. This relationship is expressed analytically below and graphically in Figure 2. As can be seen, for localities with very low initial supply of skilled workers, relative wages do not change since the new technology is not adopted. Localities where $\phi^H < \frac{S_i}{U_i}$ experience the largest increase in the returns to skill since they switch entirely to the new technology which acts as an increase in the demand for skill. Finally, for localities in the partial adoption region $\phi^L < \frac{S_i}{U_i} < \phi^H$, the increase in the returns to skill is strictly increasing in the supply of skill since the endogenously induced demand for skill is increasing with skill.

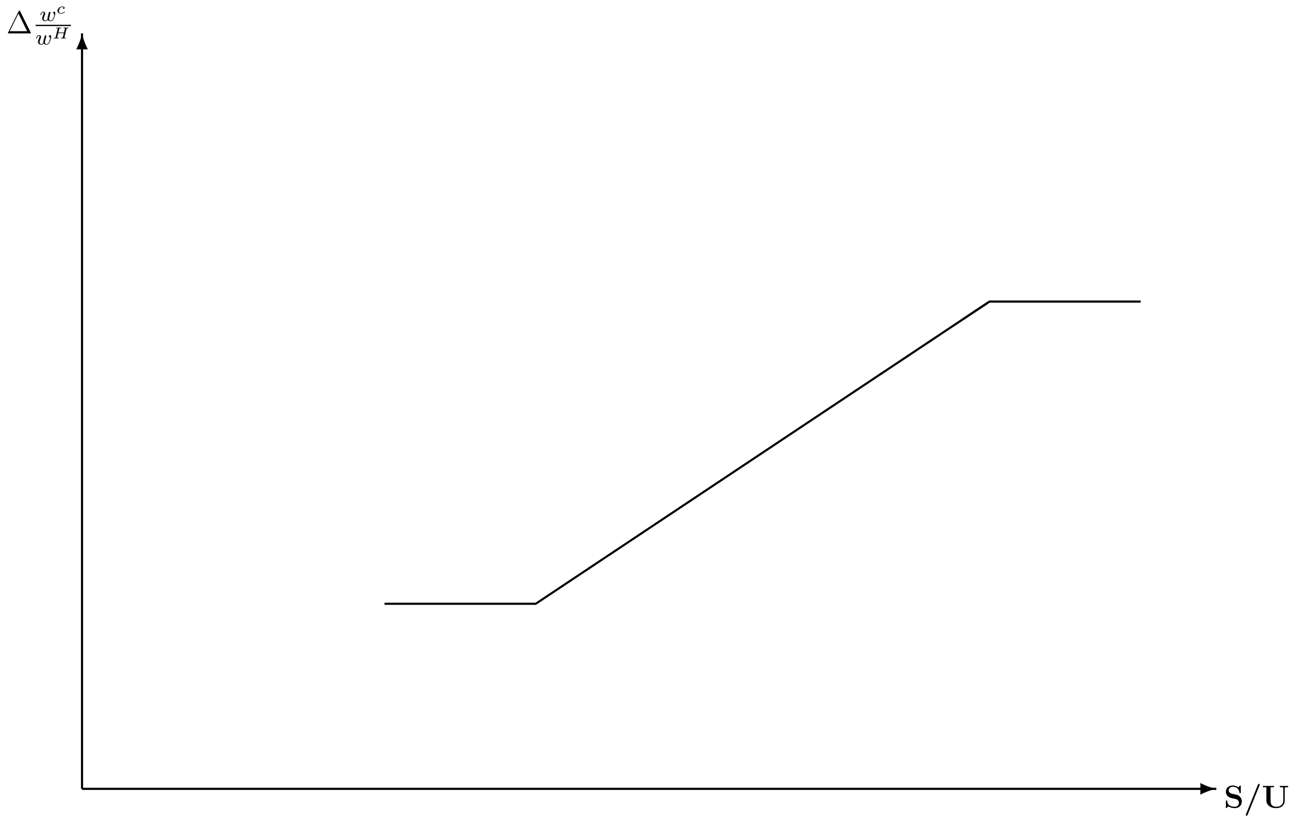
$$\begin{aligned}
\Delta \ln \frac{w_i^s}{w_i^U} &= 0 & if && \frac{S_i}{U_i} \leq \phi^L \\
\Delta \ln \frac{w_i^s}{w_i^U} &= (1-\sigma)[\log \frac{S_i}{U_i} - \log \phi^L] & if && \phi^L < \frac{S_i}{U_i} \leq \phi^H \\
\Delta \ln \frac{w_i^s}{w_i^U} &= (1-\sigma)[\log \phi^H - \log \phi^L] & if && \phi^H < \frac{S_i}{U_i}
\end{aligned}$$

Proposition 2 expresses how the arrival of the new technology induces a positive association between the supply of skill and changes in returns to skills. However, the proposition bypasses the channel through which this arises. Corollary 2 addresses this issue by combining Propositions 1 and 2 to highlight how it is the adoption of the PC-intensive technology that leads to increases in returns to skill.

Corollary 2: Returns to skill increase the most in localities which choose to adopt PCs most intensively.[9]

[9] Any approach taken to evaluate Corollary 2 must acknowledge the endogeneity between PCs and returns

Figure 2: Effect of Initial Supply on Change in Relative Wage



At first pass, Proposition 2 may appear counterintuitive since it predicts an increase in return to skill where supply is most abundant. However, this does not imply that the arrival of the new technology can cause the level of the return to skill to be positively related to supply. In fact, as stated in Proposition 3, even after the introduction of the skill-biased technology, the returns to skill must remain a weakly decreasing function of the supply of skill. Note that it is possible for the arrival of the skill biased technology to cause the disappearance of a negative relation between return and supply if localities are concentrated in the technology-mixing zone ($\phi^L < \frac{S_i}{U_i} < \phi^H$). However, in the absence of any externalities in adoption, the model implies that the relationship between returns to skill and supply of skill cannot be positive even after the introduction of the skilled-biased technology.

Proposition 3: The arrival of the skill biased technology cannot induce a positive association between the return to skill and the supply of skill.

The content of Propositions 2 and 3 can be easily inferred from Figure 1. Because the returns to skill in this figure are captured by the slope of the production function, we can note that the slope of the outer envelop is weakly decreasing in the fraction of skilled workers. This

to skill. Corollary 2 implies that its is the adoption of PCs induced by differences in initial supply of skill (or initial returns to skill) that causes increases in returns to skill.

is the content of Proposition 3. In contrast, if we consider the change in the return to skill induced by the new technology for an initial supply in the region ($\phi^L < \frac{S_i}{U_i} < \phi^H$), we see that the increase in the slope is larger for initial higher levels of supply. The reason is that the return to skill was initially more depressed in the higher supply localities and therefore the new technology allows for greater induced demand for skill in such areas. The content of Proposition 3 is depicted in Figure 3. In this figure we see that the availability of the new technology alters the relationship between returns to skill and supply. However, the slope of the new relationship is nowhere positive. Note that in the region $\phi^L < \frac{S_i}{U_i} < \phi^H$, the slope of the relationship is zero since the technological choice allows the reallocation of additional skill between the two technologies without affecting the returns.[10]

Now that we have examined the effects of skill supply on both technology adoption and relative wage change, we can therefore combine the two to obtain Corollary 3.

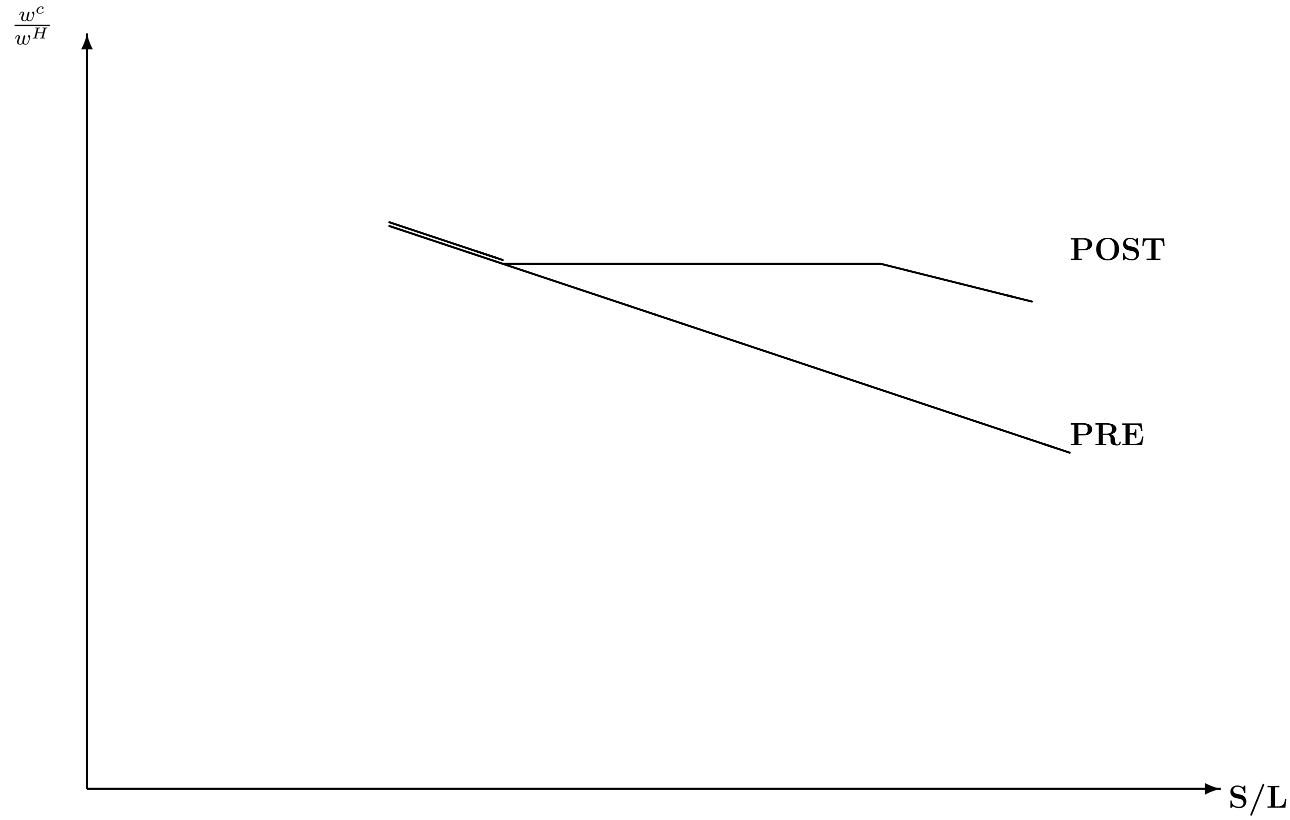
Corollary 3: The return to skill will not be larger in localities with more intensive use of PCs.

Corollary 3 indicates that although PC adoption and increases in the returns to skill should go hand-in-hand (as stated in Corollary 2), such positive co-movement cannot induce an outcome where the returns to skill are higher in localities with high PC-intensity than localities with less PC-intensity. The reason for this result comes directly from the endogenous PC adoption in the framework. To be more precise, PCs are adopted more aggressively in one locality versus another only because the cost of skill is lower. Therefore, PC capital cannot be more intensely used in a locality with a higher cost of skill. By contrast, if the adoption and subsequent use of PCs were viewed as an exogenous phenomena (as is the case in many papers), then it would be natural to expect to find a positive association with PC use and returns to skill (assuming that PCs are a skill-biased technology). Hence, this prediction nicely illustrates how a model of endogenous technology adoption differs from more conventional models with exogenous technological change.[11]

[10] This mechanism is identical to the underlying factor price equalization zones in international trade theory.

[11] The implication of endogenous adoption stated in the previous propositions and corollaries would be modified if the adoption process involved externalities. For example, suppose there existed a network type externality associated with the adoption of the new technology, so that as more local production was done with the new technology, the productive performance of the new technology increased. In such a case, it would be possible to have returns to skill being positively correlated to PC-intensity. Also, the returns to skill could be, at least over a range, an increasing function of the supply of skill.

Figure 3: Effect of Supply on Relative Prices



From Figure 3, one can also observe that an increase in skill has a different effect on its return prior to the arrival of the new technology than during the technological transition. As noted in Proposition 4, an increase in supply has less of an effect on the return to skill during the transition, since for localities experiencing the transition, an increase in supply causes them to allocate a greater fraction of the workforce to the new technology, thereby allow the returns to skill to stay constant as supply increases.

Proposition 4: A local increase in the supply of skill has a less negative effect on the return to skill during a technological transition. In particular, for localities experiencing the transition, an increase in skill only accentuates the adoption of the new technology and does not decrease the return to skill.

2.1 Allowing for labor mobility

We have so far emphasized the role of differences in skill mix at the city level on technology adoption and on the subsequent induced change in wages. However, it is also important to examine whether the above results would still hold if skill mix could endogenously react to changes in relative wages. To this end, we extend our model to allow workers to move between localities in response to differences in expected utility. However, even if workers are perfectly mobile across localities this will not necessarily lead to equal wages since localities may have other attributes that affect their attractiveness.[12] Moreover, these local characteristics will generally be valued by skill and unskilled workers differently which will cause skill mix to vary geographically. To capture the notion that some localities have a comparative advantage in attracting a more skilled labor force, let us consider the situation where there are two types of housing in each city, one that is attractive to skilled workers and one that is attractive to unskilled workers; and let us assume that localities differ in their capacity to supply one type of housing versus the other.[13] To be more precise, let preferences of workers depend on their consumption of a traded good, c, and their consumption of housing services h^S or h^u. For a skilled worker, preferences are captured by the utility function $\log(c) + \upsilon \log(h^s)$, and for an unskilled worker preferences are given by $\log(c) + \upsilon \log(h^u)$. The consumption good is assumed to be tradeable across cities and to be produced using the technologies discussed in the previous section; that is, technological change will take the form of a change in the technology set available to produce the tradeable consumption good. The two types of housing are assumed to be locally produced with the production possibility set of a locality given by

$$[d_i(H_i^s)^\psi + (H_i^u)^\psi]^{\frac{1}{\psi}} = E_i \qquad \psi > 1$$

where H_i^s is the total supply of housing directed to skilled workers in market i, H_i^u is the total supply of housing directed to unskilled workers in market i, and E_i is the land endowment of market i. For simplicity, we assume that land is the only input used to produce the non-tradeable good called housing. Note that the parameter d_i in the production possibility set governs a city's comparative advantage for providing housing that is attractive to skilled versus unskilled workers. A city with a low value of d has a comparative advantage in suppling housing that is desirable to skilled workers. Given a competitive market for housing, the production possibility set implies that the relative prices for housing services will satisfy:

$$\frac{p_i^s}{p_i^u} = d_i(\frac{H_i^s}{H_i^u})^{\psi-1} = d_i(\frac{S_i h_i^s}{U_i h_i^u})^{\psi-1}$$

[12] As discussed in Black, Kolesnikova and Taylor (2004), amenities can help explain differences in skill mix, differences in returns to skill, and differences in housing prices across cities.

[13] The assumption that housing preferences are directly associated to skill is a shorthand way of of capturing the differential demands of the two groups for local amenities. An alternative, but more complicated, approach is to endogenously derive the different demands through income effects.

where $S_i h_i^s$ represents the total demand for housing by skilled workers in market i, and $U_i h_i^u$ represents the total demand for housing by the unskilled.

Workers take wages and house prices for each locality as given when deciding where to locate and how much housing to consume. Utility maximization on the part of workers implies that the indirect utility of a skilled worker in locality i is given by $(1+\upsilon)\log(w_i^s) - \upsilon \log p_i^s$, and the indirect utility for an unskilled worker is given by $(1+\upsilon)\log(w_i^u) - \upsilon \log p_i^u$; where p_i^s and p_i^u are the prices of housing in locality i. The associated quantities of housing demanded in locality i is $h_i^s = \frac{\upsilon w_i^s}{(1+\upsilon)p_i^s}$ for a skilled individual and $h_i^u = \frac{\upsilon w_i^u}{(1+\upsilon)p_i^u}$ for an unskilled individual. Since workers will migrate to ensure equal utility across localities, it implies that relative skill premiums will be proportional to the difference in relative housing prices between any two localities i and j, as expressed below:

$$\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}) = \upsilon[\log(\frac{p_i^s}{p_i^u}) - \log(\frac{p_j^s}{p_j^u})]$$

The above two relationships can be combined to provide a simple expression for the relative supply of skill as a function of the relative wages between two localities i and j. This is given by the equation below where we can see that a locality will attract more skilled workers if it either has a comparative advantage in supplying desirable housing (low d_i) or if the skill premium is high.

$$\log(\frac{S_i}{U_i}) - \log(\frac{S_j}{U_j}) = \frac{\psi}{\psi - 1}[\log(\frac{d_j}{d_i}) + (\frac{1+\upsilon}{\upsilon})(\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}))]$$

If we now consider the situation before the technological transition, we know that demand for skills by firms producing the tradeable consumption good in locality i is given by $\frac{aS_i^{\sigma-1}}{(1-a)U_i^{\sigma-1}}$. Combining the relative demand relation with the supply relationship allows to express relative wages and relative supplies as a function of the comparative advantage parameters d. It is easy to verify that is this case, the mobility of workers will lead the locality with a low d_i to have both a greater relative supply of skilled versus unskilled workers, and a lower return to skill. This represents the situation prior to the arrival of a new technology. The question we now want to address is how allowing for the mobility of workers across cities affects our results regarding technological transition. This is stated in Proposition 5.

Proposition 5: When skill supply across localities adjusts to equate utility, a technological transition will still lead to (1) PC per worker being higher where relative skill supply is initially highest, (2) the returns to skill to increase most where skill supply is initially highest, and (3) it will not lead to a situation where returns to skill are positively associated with either higher levels of PC per capita or the supply of skill.

Proposition 5 indicates that the results from the previous section, where workers were assumed immobile, can be extended to allow for worker mobility.[14] The intuition for why these results extend easily to a case with mobility can be inferred by adding to Figure 3 a positive sloped supply curve for skill. In this modified figure, localities with low d's would have relative supply curves for skill that are translated to the right relative to those for localities with a high value of d. The technological transition causes the relative demand for skill to pivot upward, moving along the supply for skill schedules where, among any two localities experiencing the transition, there would a greater increase in return to skill where initial supply is highest. This is not to say that allowing worker mobility has no effect on the equilibrium outcome. Clearly it would have quantitative effects. However Proposition 5 indicates that it would not change the qualitative implications of Propositions 1 through 3. Interestingly, when worker mobility is modeled along the lines presented here, a technological transition would be a period where relative skill supplies diverge across localities; with high skill cities becoming even more highly skilled. This arises since the arrival of the new technology counters the decreasing returns to skill and thereby favors in migration of skilled workers.[15] Such an outcome could at at first pass be mistaken for the emergence of some form of positive externality associated with skill. However, the current model offers an explanation to such a divergence by the arrival of a new technology choice which does not involve any externalities.

[14] The only Proposition which can not be directly extended to the case of worker mobility is Proposition 4 since it is no longer meaningful to talk about an exogenous increase in the supply of skill.

[15] The evidence on whether or not there has been a divergence across cities in the relative supply of skill over 1980-2000 is somewhat mixed. If one follows Berry and Glaeser (2005) and defines skilled workers as workers with at least a 4 year college degree, then there appears to have been divergence. In contrast, if one follows Katz and Murphy and to also include individuals with a non-BA post-secondary degree into the skilled group, then the data does not reveal any divergence. Further, these results are also sensitive to whether relative skill supplies are measured in logs or in levels. Since the main focus of this paper is not whether a technological transition causes divergence in the geographical distribution of skill, we do not pursue this point in the empirical section.

3 Data

Section 2 highlighted several implications of viewing technological adoption as driven by principles of comparative advantage. Our goal now is to examine whether city-level outcomes observed over the 1980-2000 period exhibit the patterns implied by such a model. We choose to focus on this period for several reasons. First, this is a period often considered one of technological revolution due to astounding technical progress and diffusion of information technology. Hence it is a perfect candidate period to see whether our neoclassical model of technological adoption is relevant. Second, it is a period in which returns to education increased substantially, and skill-biased technological change is often considered to be one of the reasons behind this increase. Therefore, it is particularly relevant to examine whether this period is best characterized as reflecting the effects of exogenous technological change (in line with much of the literature which treats the extent and bias of technological change as an exogenous driving force) or whether instead it reflects a process of endogenous choice of production techniques.[16]

The city-level data we use can roughly be divided into two categories; technology and demographic. The technology data is derived from establishment-level information on technology use and is described in more detail in Doms and Lewis (2006). About 160,000 establishment-level observations per year are used to compute the PC intensity (PCs per employee) of each city in our sample (230 cities).[17] Our PC intensity measure is an industry-adjusted measure; in computing the PC intensity, we control for the 3-digit SIC industry interacted with 8 establishment size classes, for a total of over 1,800 interactions.[18]

We focus on PCs instead of other IT technologies for several reasons. First, businesses spent

[16] Note that it could be possible that the extent of bias of technical change is endogenous at the national level, but not at the city level since markets across cities are well integrated. In such a case, our approach of focusing on city-level outcomes would not identify elements of endogenous technological change. In other words, our empirical work evaluates the joint hypothesis that technological adoption is a phenomena that reacts to market conditions and that the labor markets in cities across the US are not perfectly integrated.

[17] To increase the precision of our city-level measures, the 1990 PC intensity measure uses data from 1990 and 1992 and the 2000 PC intensity measure relies on 2000 and 2002 data. Doms and Lewis (2006) define "city" primarily as consolidated metropolitan statistical areas (CMSAs). The logic was to derive city definitions that corresponded to the idea of a local labor markets. In some cases, CMSAs were modified to more closely capture the concept of a labor market. Our results throughout the paper are insensitive to how we define labor markets in especially large, contiguous areas, such as in and around New York City.

[18] The SIC-size interaction allows for the possibility that, for instance, large banks perform different operations than small banks. As described in Doms and Lewis (2006), our city-level measures of PC intensity are strongly correlated with other measures that control for 4-digit SIC and for measures that also control for the firm to which an establishment belongs.

about 90 percent more money on PCs during the 1990s than on other types of computers. Also, spending on PCs is likely correlated with other information technology spending, such as spending on software, computer networking equipment, printers, et cetera. Finally, we were able to obtain consistent measures of PCs over this period.[19]

Figure 4 shows a scatter plot of the city-level PC measures for 1990 and 2000 (the 1990 results are shown along the horizontal axis and the 2000 results are shown on the vertical axis). The axes in Figure 4 are scaled to the San Francisco Bay Area, a city that consistently ranked very highly in nearly all measures of technology that we examined. For instance, in 1990, the mean establishment in San Francisco had .12 more PCs per employee than the mean establishment in Scranton, PA (the city that frequently ranked among the lowest of our sample of cities) after controlling for industry and size differences across the two cities. In 2000, the difference in PC intensity between the Bay Area and Scranton increased to .16. One item to note about figure 4 is that the differences in PC intensity are persistent over time: for example, the correlation in PC intensity between 1990 and 2002 is 0.57.

Most of the city demographic information we use comes from the decennial censuses, specifically the public-use micro-data files for 1940, 1980, 1990, or 2000. Using these data, our measure of skilled labor is defined as workers who have a least a four year college degree plus one-half of those with at least some college education. Measures similar to this one have often been used in research examining the effects of skill-biased technological change, such as Katz and Murphy(1992), Autor et al.(2003), and Card and DiNardo(2002).[20]

Figure 5 shows a scatter plot of the log of skilled to unskilled workers for 1980 and 2000. As the figure shows, the measure of skill varies greatly across cities, but the ranking of cities according to skill remains nearly constant over time. As with the PC intensity data, there is great persistence in the skill mix over time as the skill mix in 1980 explains over 85 percent of the variance in 2000, although the mean skill share increases sharply. Doms and Lewis (2006) finds that one reason why the skill mix for a city changes by more or less than average is immigration. Cities such as Fresno, Stockton, and El Paso received relatively large numbers of unskilled immigrants over the past several decades, resulting in lower than average skill appreciation.

[19] Other measures of information technology were examined in Doms and Lewis (2006), including more refined measures of PC. The results in Doms and Lewis (2006) were very robust to choice of technology measure.

[20] Our sample of workers includes both men and women aged 16-65 with at least one year of potential work experience, employed, and not living in group quarters.

In most cases, our measure of relative wages is computed using the average log hourly earnings of people who report completing exactly 12 years of education and people who report completing exactly four years of post high school education.[21] We do not use the raw (log) wage gap as our measure of returns - though our results are robust to using that - but rather the wage gap adjusted for a fourth-degree polynomial in potential work experience, a female dummy, an immigrant dummy, and a dummy for people born after 1950. Appendix 2 has more details on the wage construction.

We also construct several city-level measures that we label as "city controls." These are the log of the size of the area's labor force and the percent of the area's workforce which is African American, female, Mexican-born, and U.S. citizens. Additionally, we construct "industry controls" which reflect the employment distribution across 12 major industry groups within each city.[22]

4 Empirical results

This section uses the city-level data described in the previous section to examine whether the patterns of PC adoption, returns to skill, and education levels conform to the predictions of the model outlined in Section 2. The model highlighted how forces of comparative advantage would cause cities with initial differences in educational attainment to react differently to a change in the technological paradigm. In the model, initial differences in skill mix were assumed to be exogenous, coming from, say, historical events unrelated to the new technological opportunities or from differences due to local amenities. The model emphasizes how exogenous differences in initial skill mix **cause** different outcomes, both in terms of adoption of the new technology and in changes in the returns to skill. However, in our empirical analysis, initial levels of educational attainment cannot a priori be taken as exogenous to the subsequent process of technological change. Instead, it is necessary for us to either make the case that initial levels of education can be treated as exogenous, or alternatively provide an valid instrumental variable strategy aimed at estimating the causal effect of a city's initial

[21] As a robustness check, we compute the returns to college using the sample of all workers with at least 11 years of education in Table 3b.

[22] These give share of employment in industry categories which corresponding roughly to one-digit SIC: agriculture & mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, "FIRE," business and repair services, other low-skill services, entertainment, and professional services. Note that in the PC regressions, these industry mix controls are *on top of* the detailed industry adjustment already performed on the dependent variable (3-digit SIC x establishment size). The industry mix controls therefore capture any additional indirect or "spillover" effects of industry mix.

level of education on subsequent adoption of technology and change in the returns to skill .

To be explicit about how we will process to address potential endogeneity issues, let $\Delta X_{i,2000-1980}$ represents the change in outcome X (either PC adoption or change in the college-high school ln wage gap) in city i between 1980 and 2000 and let $\ln(\frac{S}{U})_{i,1980}$ represent the initial educational mix of the area. We will evaluate the model's prediction about the impact of initial skill mix on X by estimating linear models of the form:

$$\Delta X_{i,2000-1980} = \gamma_0 + \gamma_1 \ln(\frac{S}{U})_{i,1980} + \epsilon^x_{i,2000-1980} \tag{1}$$

where $\epsilon^x_{i,2000-1980}$ represents other factors not captured by the model that affect the change in X. There are at least two reasons why in the data $\ln(\frac{S}{U})_{i,1980}$ could be correlated with $\epsilon^x_{i,2000-1980}$ and thereby cause OLS estimates of γ_1 to be biased. First it may the case that workers can partially predict $\epsilon^x_{i,2000-1980}$ and move in anticipation. Second, there may be some other factor that is both correlated with educational attainment in 1980 and with determinants of ΔX_i between 1980 and 2000. For concreteness, we will refer to a cities with a high level of this factor as "innovative" cities. In the empirical model, we will represent this factor with the (unobserved) random variable $\eta_{i,1980}$ and loading factor γ_x for outcome X. Incorporating this into (1):

$$\Delta X_{i,2000-1980} = \gamma_0 + \gamma_1 \ln(\frac{S}{U})_{i,1980} + \gamma_x \eta_{i,1980} + \mu^x_{i,2000-1980} \tag{1'}$$

The error term from (1) has been expanded as $\epsilon^x_{i,2000-1980} = \gamma_x \eta_{i,1980} + \mu^x_{i,2000-1980}$, where $\mu^x_{i,2000-1980}$ is defined to be orthogonal to $\ln(\frac{S}{U})_{i,1980}$. However, the endogenous error component, $\eta_{i,1980}$ is not: an area's innovative attributes may have favored demand for skill in locality i prior to the 1980-2000 period and attracted more educated workers to i. Notice $\eta_{i,1980}$ is a common factor which potentially affects different outcomes of interest with different loadings.

One empirical strategy we could adopt to explore endoegneity bias would be to use educational attainment as of 1940, $\ln(\frac{S}{U})_{i,1940}$, an instrument for $\ln(\frac{S}{U})_{i,1980}$.[23] This would be a valid instrumental variable strategy if $\ln(\frac{S}{U})_{i,1940}$ is known to be uncorrelated with $\eta_{i,1980}$. However, a priori it is again unclear that $\ln(\frac{S}{U})_{i,1940}$ is uncorrelated with the error term

[23] 1940 has the advantage of not only long predating the computer revolution, but also the massive expansion in college attainment in the U.S. (associated, e.g., with the World War II "G.I." Bill.)

in (1'). Therefore, rather than just report such IV result, we will also present a transformed control function specification which provides considerable insight into the magnitude of the endogeneity problem. In particular, we will use the fact that 1980 education mix can be decomposed as the 1940 skill mix plus the change in skill between 1940 and 1980 ($\Delta \ln(\frac{S}{U})_{i,1980-1940}$). Substituting this in to (1') gives rise to (1").

$$\Delta X_{i,2000-1980} = \gamma_0' + \gamma_1' \ln(\frac{S}{U})_{i,1940} + \gamma_2' \Delta \ln(\frac{S}{U})_{i,1980-1940} + \gamma_x \eta_{i,1980} + \mu^x_{i,2000-1980} \qquad (1'')$$

Our interest in specification (1") derives from the idea that the potential endogeneity of $\ln(\frac{S}{U})_{i,1980}$ with respect to $\eta_{i,1980}$ is unlikely to bias the coefficients on each component of 1980 skill mix in exactly the same way. For example, the reverse causality story should most likely create a strong correlation between $\eta_{i,1980}$ and the part representing **changes** in skill mix leading up to the PC revolution, that is, $\Delta \ln(\frac{S}{U})_{i,1980-1940}$. In contrast, in such a story the endoegenity bias related to $\ln(\frac{S}{U})_{i,1980}$ should be weaker. Hence, if we compare estimates of γ_1' and γ_2' from (1") this will provide insight into the nature of the endoegneity problem. In particular, under the null hypothesis that there is no endoegenity problem, then the coefficients on γ_1' and γ_2' should be identical. Therefore testing whether the OLS estimates of γ_1' and γ_2' in (1") are equal provides evidence on the validity of the OLS estimates. Note that this way of proceeding it is simple extension of a "control function" specification approach, i.e. adding a control for the "bad" (endogenous) variation. However, our approach generalizes this by not requiring us to take a stand on whether $\Delta \ln(\frac{S}{U})_{i,1980-1940}$ or $\ln(\frac{S}{U})_{i,1940}$ is the source of bad variation in 1980 skill mix.

In what follows, we will show that data do support the idea that there are "innovative cities" (which we are representing as $\eta_{i,1980}$), that is, there does appear to be a common factor which affects both a city's demand for skill in 1980 and its PC adoption between 1980 and 2000. However, we will also show that the data is quite supportive of the null hypothesis that this unobserved factor, $\eta_{i,1980}$, does not affect educational attainment as of 1980. In fact, we will show that the data fail to reject the equality of the OLS coefficients γ_1' and γ_2' in (1"), which is consistent with $\eta_{i,1980}$ not being correlated with skill mix as of 1980. We interpret this finding as indicating that there are innovative cities with respect to PC adoption, but the extent of innovativeness is not related the level of skill mix: there are innovative cities with both both initially high and low levels of educational attainment.[24] Hence we will thus argue

[24]While this inference may at first pass seem inconsistent with migration incentives, it might be explained

that OLS estimation of (1) can be reasonably used to evaluate the predictions of the model. Given the contentiousness of this statement, we will probe the robustness of the results in other ways as well.

Our exploration of the relationships implied by the theory will broadly follow the order of the Propositions and Corollaries in the model section. We begin by examining the determinants of technology adoption as measured by the diffusion of PCs (Proposition 1). The first results echo those presented in Doms and Lewis (2006) by looking at the link between PC adoption and the local supply of skill. We then go further by examining whether returns to skill in 1980, that is at the beginning of the diffusion process for PCs, are negatively associated with the intensity of PC use in 2000 (Corollary 1). We next turn to examining implications of the model for changes in the returns to skill, as measured by the ratio of college wages to high school wages. In particular, we explore whether returns to skill increased most where skill was initially most abundant (Proposition 2). The model also implies that the initially negative relationship between returns and skill supply will become less negative and possibly entirely disappear. However, as indicated by Proposition 3, this relationship is predicted to remain non-positive. These implications are examined by documenting the returns-skill relationships for each of the years 1980, 1990 and 2000 separately. We then go back and examine whether the returns to skill increased most where PC were adopted most aggressively (Corollary 2), while simultaneous verifying whether the relationship between city-level PC intensity and returns to skill does not exhibit a positive relationship (Corollary 3). Recall that these Propositions were shown to be robust to allowing local supplies of skill to react to changes in the returns to skill (Proposition 5).

4.1 PC adoption and local market conditions

Table 1 reports regression results motivated by Proposition 1 and Table 2 reports results motivated by Corollary 1. Table 1 examines whether PCs were adopted more intensively in cities where the initial skill composition was high; the dependent variable is the within industry ratio of PC-per employees as of 2000 adjusted for industrial composition, as described in the data section and appendix. This variable is treated as ΔX_i in Equation (1) since PC intensity in 1980 was zero. Column (1) reports the findings from the regression of this measure of technological adoption on the ratio of college to high school workers in 1980.

by that fact that innovative cities pay high wages for all workers and therefore such cities tend to attract all type of workers.

As can be seen, there is a strong positive correlation between PC intensity in 2000 and the educational attainment of the labor force in 1980. Obviously this does not prove a causal a link from skill composition to PC intensity, since the initial skill mix may be correlated with omitted factors. As a first attempt to explore the endogeneity of $\ln(\frac{S}{U})_{i,1980}$ with respect to $\epsilon^{PC}_{i,1980-2000}$, in Column (3) we report the instrumental variable estimation of the relationship where educational attainment as of 1940 ($\ln(\frac{S}{U})_{i,1940}$) is used as an instrument. Since using the 1940 data forces us to change or sample of cities, in Column (2) we report the OLS estimate for the more restricted sample. Interestingly, the coefficient on $\ln(\frac{S}{U})_{i,1980}$ is almost identical whether it is estimated by OLS or by IV. This is a first piece of evidence suggesting that the correlation between $\ln(\frac{S}{U})_{i,1980}$ and the error term in (1) ($\epsilon^{PC}_{i,1980-2000}$) may not be very strong. In particular, consider the potential endogeneity due to workers anticipating $\epsilon^{PC}_{i,1980-2000}$, and choosing to reallocate accordingly. If this were the case, we would expect the the IV estimation would result in a much lower estimate of γ_1. Instead, OLS and IV estimates are almost identical.

We make this point even more saliently in Column (4) where we exploit the fact that $\ln(\frac{S}{U})_{i,1980}$ can be expressed at the sum of $\ln(\frac{S}{U})_{i,1940}$ and $\Delta\ln(\frac{S}{U})_{i,1980-1940}$. This allows us to include both of these terms separately in the regression and estimate it by OLS. It is interesting to note that the correlation between $\ln(\frac{S}{U})_{i,1940}$ and $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ is significantly negative in our data, which reflects less educated cities catching up with more educated cities during the 1940-80 period. The important element to focus upon in Column (4) is that the estimates on both components of $\ln(\frac{S}{U})_{i,1980}$ are almost identical (p-value associated with them being different is .13). This reinforces the observation in Column (3) that sorting of skilled workers in anticipation of high values of $\epsilon^{PC}_{i,1980-2000}$ – "reverse causality" – is unlikely since we would expect this to be much more severe for $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ than for $\ln(\frac{S}{U})_{i,1940}$. Since the coefficients on both terms are almost identical, this source of bias seems to be small. Workers likely had some sense that the arrival of PCs would raise demand for skills, but they seem not to have anticipated its differential impact on demand for skills across different labor markets.

Now consider the second potential source of bias related to the idea that third factors correlated with skill mix are the true drivers of PC adoption. Under this hypothesis, it is again difficult to formulate a story whereby both $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ and $\ln(\frac{S}{U})_{i,1940}$ would be subject to the same bias, especially given their negative correlation. The most plausible case would be one where the bias should be greatest for $\Delta\ln(\frac{S}{U})_{i,1980-1940}$. Since the coefficient on both terms in Column (4) are not significantly different, this suggests that $\ln(\frac{S}{U})_{i,1980}$ is not correlated with $\epsilon^{PC}_{i,1980-2000}$

In columns (5) and (6) of Table 1 we add a set of city level control variables to the specifications of Column (1) and (3).[25] If educational attainment as of 1980 is uncorrelated with other determinants of PC-adoption ($\epsilon^{PC}_{i,1980-2000}$), as the results from Columns (2) and (3) suggest, then adding additional variables to the regression should not generally affect our estimates of γ_1, and this is true regardless of whether these variables are exogenous or not with respect to $\epsilon^{PC}_{i,1980-2000}$. As we can see, adding the set of city controls does not substantially change the estimates of γ_1. In the same spirit, in the last two columns of the table we also add a set of industry controls to the two specifications. We again see that the estimate of γ_1 is not significantly affected, and we see that the coefficients on both $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ and $\ln(\frac{S}{U})_{i,1940}$ are approximately equal, although in Column (6) there is indication at conventional levels against the strict equality of coefficients.

In Table 2 we turn to examining the link between the adoption of PCs over the period 1980-2000 and the initial college-high school wage gap, which we shall call the "price of" or "returns to" skill. Corollary 2 implies that PCs should be adopted more intensively in cities where the price of skill is initially low, since such city have a comparative advantage in adopting the new skill biased technology. In Column 1 we regress PC intensity in 2000 on the the ratio of college to high school wages in 1980. Here we see a significant negative relationship between the initial price of skill and the adoption of PCs. There is again no reason to consider this result as providing evidence of causality. In columns (2) and (3), we show that this negative association is robust to the inclusion of our set of city controls and the industry controls introduced in Table 1. In columns (4)-(6) we estimate by IV the relationship between the price of skill in 1980 and the subsequent adoption of PCs. In column (4) we use college share in 1940, $\ln(\frac{S}{U})_{i,1940}$, as the instrument for 1980 price of skill. In column (5), we use $\ln(\frac{S}{U})_{i,1980}$ as the instrument. In column (6) we use both $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ and $\ln(\frac{S}{U})_{i,1940}$ simultaneously and we report the associated over identification test. As can be seen, in all three cases we recover a significant negative relationship between the initial price of skill and PC adoption, and the coefficient is similar whether we use $\ln(\frac{S}{U})_{i,1980}$, $\ln(\frac{S}{U})_{i,1940}$ or when we use both $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ and $\ln(\frac{S}{U})_{i,1940}$ as instruments. Although the coefficient and the returns to skill in 1980 are similar in columns (4)-(6), they are much greater in absolute value that those reported in columns (1)-(3) of Table 2. We believe that this pattern likely

[25] Estimating the relationship by IV, using 1940 educational attainment as the instrument, also gives very similar to that of Column (3) when controls are added.

reflects the presence of innovative cities. To see this, it is helpful to express the price of skill in 1980 as a negative function of the supply of skill in 1980 and as a function of other factors that affect the demand for skill in 1980, which we continue to denote by $\eta_{i,1980}$:

$$\ln(\frac{w^s}{w^U})_{i,1980} = \alpha_0 - \alpha_1 \ln(\frac{S}{U})_{i,1980} + \eta_{i,1980} \qquad \alpha_1 > 0 \quad (2)$$

The problem we are facing is can be expressed in the PC version of equation (1'):

$$PC_{i,2000} = \gamma_0 + \gamma_1 \ln(\frac{S}{U})_{i,1980} + \gamma_2 \eta_{i,1980} + \mu^{PC}_{i,1980-2000} \qquad \gamma_2 > 0 \quad (3)$$

In other words, there may be a common "innovative city" factor which simultaneously raises the demand for skill in 1980 and PC adoption. Such a factor will bias what would otherwise be a negative relationship between PC adoption and the initial price of skill towards zero. This can also be seen by substituting (2) into (3):

$$PC_{i,2000} = (\gamma_0 + \frac{\gamma_1 \alpha_0}{\alpha_1}) - \frac{\gamma_1}{\alpha_1} \ln(\frac{w^s}{w^U})_{i,1980} + (\frac{1}{\alpha_1} + \gamma_2)\eta_{i,1980} + \mu^{PC}_{i,1980-2000} \qquad (4)$$

The causal effect of the price of skill on PC adoption is $-\frac{\gamma_1}{\alpha_1}$, but because of the positive correlation between $\ln(\frac{w^s}{w^U})_{i,1980}$ and $\eta_{i,1980}$ the OLS estimate of this relationship will be muted. However, we found evidence in Table 1 that college share as of 1980 is uncorrelated with $\eta_{i,1980}$, making it a potentially valid instrument for $\ln(\frac{w^s}{w^U})_{i,1980}$. The fact that our three IV strategies give statistically similar estimates of the effect, again suggests that either educational attainment in 1980 is uncorrelated with $\eta_{i,1980}$, or both components of $\ln(\frac{S}{U})_{i,1980}$ reflect the same biases. Given that the later seems implausible, we conclude that the innovative factor varies significantly across cities, but educational attainment as of 1980 does not appear to systematically reflect this spirit. Examples of particular areas support this interpretation. Charlotte, NC had, by our measure, slightly below average skills in 1980, yet very high returns to skill in 1980 and PC adoption 1980-2000. At the other end, Provo-Orem, UT was one of the most educated markets in our sample in 1980, had very low returns to college in 1980, and below average PC adoption 1980-2000. Salinas, CA is similar to Provo.

4.2 Changes in returns to skill and initial skill level

A central prediction of the model presented in Section 2 is that, during a technological transition, the relationship between returns to skill and the supply of skill takes the following form. Increases in the returns to skill should be greatest where supply is initially most abundant (Figure 2 and Proposition 2). Furthermore, such a relationship should arise due to to a flattening of the relationship between returns to skill and the supply of skill (Figure 3). Accordingly, in Table 3a we report results from regressions aimed at documenting the relationship between changes in the return to skill over the period 1980-2000 and the level of skill in 1980. In column (1) we report the OLS relationship without any additional control variables. In columns (2) and (3) we add our set of city level control variables and a set of industry level control variables. As can be seen, there is a strong positive association between the skill mix in 1980 and the change in returns to skill over the 1980-2000 period. The data supporting this relationship is presented in Figure 6. While we have argued above that the 1980 skill level may reasonably be considered exogenous to the subsequent forces determining PC adoption, we now want to explore whether the same is true for the change in returns to skill over the period 1980-2000. Accordingly, in columns (5)-(7) we estimated the relationship between changes in the returns to skill and initial educational attainment by decomposing the initial educational attainment into its two components. As before, we expect that endogeneity of skill mix should manifest itself in the form of different coefficients on the two components of $\ln(\frac{S}{U})_{i,1980}$. In contrast, if we obtain similar estimates on both components it provides evidence against endogeneity. As can be seen in the Table 3a, the estimated effects of both components are similar, and this is robust to the inclusion of both city and industry level variables. Hence, if one were to conjecture that the OLS relationship is driven by individuals agglomerating in 1980 in cities with high expected increases in the returns to skill, this would be difficult to believe since such a prediction of increased returns would need to have been equally well formulated in 1940.

As an additional check on whether the results are driven by anticipated returns, columns (4) and (8) of Table 3a correct for the endogenous selection into particular metro areas using Dahl's (2002) flexible control function approach. The idea behind Dahl's approach is that place of birth is a strong predictor of a person's eventual labor market, but within demographic categories, place of birth is orthogonal to potential returns to skill (the maintained assumption). In practice, the correction involves controlling for a polynomial in the probability that an individual is observed in a particular metropolitan area, estimated from the observed distribution of individuals across metro areas by demographic category (age x race x gender) and state or country of birth.[26] Consistent with our argument that endogeneity is not a major problem here, columns (4) and (8) of Table 3a reveal that the selection correction has little effect on the results.

Another issue is whether the relationship in Table 3a simply reflects some unobserved difference between more- and less-educated markets which results in differential trends in returns to college which predate the PC revolution. We might be worried, say, that earlier generations of technologies favoring skilled workers were adopted more quickly in these permanently more "innovative" cities. We investigate this empirically in Table 3b. Obtaining reliable estimates of the returns to college by area prior to 1980 turns out to be quite difficult because wage samples for years before 1980 tend to be small. To deal with this, we have combined the 1940 and 1950 censuses, which report geography in identical detail. Even with the stacked data, the number of wage observations, especially for college graduates, is quite small in many metropolitan areas.[27] To increase the sample further, rather than including only those with exactly 12 and exactly 16 years of education (as we do in all other tables), we estimated the college-high school wage gap in a regression which included all individuals with at least 11 years of education. To deal with the skill heterogeneity in this larger group, we conditioned our city-specific estimates of the returns to college on a linear term in years of education.

Table 3b reports the results. First, to make sure our earlier results remain with the change in estimation methods, columns (1)-(3) of Table 3b report regressions of the change in log wages between 1980-2000, estimated with the new method, on 1980 college share. A significant, positive relationship remains with and without controls. In contrast, there is no evidence of a positive relationship between 1980 skill share and the change returns 1940/50-1980. Thus it appears the relationship we are uncovering is exclusive to the recent period of rapid technological change. This supports the earlier evidence that skill mix in 1980 may have been exogenous to other forces affecting returns.[28]

[26]The details of how this is implemented mirror Beaudry, Green and Sand (2007) and are described in Appendix 2.

[27]In the 1940 and 1950 public use Census samples, only a small subset of so-called "sample-line" respondents were asked to report their wages.

[28]We also tried looking at the 1970 Census, where the sample is a bit larger but the number of identifiable metropolitan areas is smaller. Using these data, there is also little trend in returns associated with college share, or a weak positive trend. To the extent the trend is positive in the 1970s, it may reflect the beginnings of the computer revolution.

While the patterns highlighted in Tables 3a and 3b are consistent with our model, they could be consistent for the wrong reason. In particular, greater increases in returns to skill arising where skill is most abundant could arise from three different underlying movements in the returns-supply relationship. It could arise from a positive relationship between the return to skill and supply of skill becoming more positive over the period 1980-2000. Alternatively, it could come from the relationship changing from being negative to being positive. Finally, it could arise from the relationship becoming less negative, without becoming positive. As we have emphasized, only the third case is consistent with our model of technological revolutions (Proposition 4).

To examine these implications, columns Tables 4a and 4b report estimates of of the relationship between returns to skill and the skill mix in the years 1980, 1990, and 2000. Table 4a reports OLS estimates. With no controls, the estimated relationship between returns to skill and the supply of skills falls from -0.07 in 1980 to 0.00 in 2000.Adding city controls, in columns (4)-(6), and city and industry controls, in columns (7)-(9), this "flattening" pattern, consistent with proposition 4, is strengthened.[29]

Table 4b presents some additional specifications, including instrumental variables estimates. Columns (1)-(4) of the table report OLS estimates for the subset of 157 metropolitan areas in which the instrument - the log of college/noncollege in 1940 - is observed. In this subsample, approximately the same flattening is observed between 1980 and 2000 (columns (1) and (2)) as in Table 4a. Columns (3) and (4) correct for selection bias using the same procedure as before. Columns (5)-(8) report instrumental variables estimates for the same specifications. In all cases, we see that there was a significantly negative relationship between relative skill supply and its return in 1980, and over the following twenty year period, this relationship became less negative. In no case is the point estimate significantly different from zero in the year 2000. Viewing the data through the lens of the model suggest that as of 2000, the US economy appeared to still be in the technological transition phase where increase in the supply of skill have little or not effect of the returns to skill, but instead favor a faster adoption of the skilled biased technology as was seen in Table 1.

[29]The point estimates for 1980 correspond to an elasticity of substitution between college and high school labor in the range of 6-16 (inverse of point estimate). This is smaller than estimates of the elasticity obtained from aggregate data (e.g., Katz and Murphy, 1994) but similar to estimates based on cross-area variation, such as from research on the local labor market impact of immigration.

4.3 PC adoption and Returns to skill

The previous two subsections documented that PCs were adopted more intensively and the returns to skill increased most where skill was initially most abundant. In this subsection we begin by completing the triangle and verifying that returns to skill actually increased most where PCs were adopted most aggressively (Corollary 2). We then examine one of the model's most surprising predictions, that is, even though the returns to skill increase most where PCs are adopted most aggressively, this should not lead to a situation where cities that use PCs most intensively have the highest returns to skill (Corollary 3). In fact, if our inferences from Tables 4a and 4b are correct that the economy is still in a transition phase as of 2000, we should find no systematic link between PC use and city level returns to skill in 2000.

In Table 5 we report the results of regressions of the change in the returns to skill over the period 1980 to 2000 on PC intensity as of 2000. In the first four columns we report OLS results, with and without city and industry level controls. In all three columns we see that returns increased most in cities which adopted PC most aggressively. In columns (5)-(7) we estimate the relationship by IV. In columns (5) and (7) we use $\ln(\frac{S}{U})_{i,1940}$ as the instrument for PC adoption, and in Column (6) we use $\ln(\frac{S}{U})_{i,1980}$. Column (7) parallels column (4) in the inclusion of controls. When estimating by IV, we find that the effect of PC adoption on returns to skill varies little with the different choices of instruments. We do nevertheless find that the IV estimates are significantly larger than the OLS estimates. This is again what one should expect if (i) there are factors other than educational attainment that favored PC adoption and are correlated with returns to skill and (ii) educational attainment as of 1980 is uncorrelated with these other factors.

In Table 6 we examine the relationship between the returns to skill in 2000 and PC use. The first four columns of the table report OLS estimates. The first column has no additional controls, the second column looks at the set of cities identified in the 1940 census, and columns (3) and (4) add city and industry controls. Interestingly, we do not see any sign of a positive association between the intensity of PC use and the returns to skill at the city level, which is consistent with the predictions of our model of technological revolutions. In Columns (5),(7) and (8) use present IV estimates of this relationship where in Column (5) we use educational attainment in 1940 as the instrument for PC use in 2000, and column (6) uses educational attainment in 1980. Column (7) also adds city and industry controls, and column (8) adds the correction for the endogenous selection into particular markets. Given our interpretation of the results in previous tables, we should expect that the IV estimate of the association between PC use and returns to skill to be more negative that the OLS estimate. The reason being that the presence of innovative cities should induce a positive co-variance between PC use and returns to skill, while instrumenting with past skill mix should eliminate this positive bias as we have argued that initial skill mix is likely uncorrelated with the unobserved "innovative" factor. Consistent our interpretation of omitted factors, that is what we find. The IV estimates in Table 6 are more negative than the OLS estimates, suggesting no positive link between returns to skill and PC use as of 2000 as implied by the model. If anything, the IV estimates suggest a possible weak negative relationship.

4.4 Quantifying Effects

Our model of endogenous technology adoption has the property that localities with different levels of educational attainment should adjust differently to a technology revolution, with more educated cities adopting the new technology more aggressively and thereby experiencing a greater increase in the returns to skill. The evidence present in Tables 1-6 provides support for this view. However, from these tables it is difficult to assess the importance of this process in relation to the overall increase in the returns to education observed over the period 1980-2000.

It is interesting to ask, for example, what our estimates imply regarding the magnitude of increased returns to education in a high educated city versus a low educated city. To take the two extremes, we can compare the predicted outcomes for Hickory, NC, our sample's least educated area in 1980, and Tallahassee FL, our sample's most educated area. In 1980, the fraction of college workers was 16% in Hickory, while it was 44% in Tallahassee.[30]. Using the estimate of Column (1) of Table 3, we predict find that returns to schooling in Hickory increased by .14 log points and in Tallahassee by .23 log points. The difference between these two different cities is a rather sizeable .09 log points. To get a sense of this magnitude it is insightful to compare it to the average increase in the returns to college across the entire US over the 1980-2000 period. In our data the average increase in returns to college is .17 log points. Hence, the cross-city differences implied by our estimates are about half as big as

[30]That Tallahassee was the most educated city in 1980 may be unexpected to some readers. However, recall that our measure of college share includes half of individuals with some college education but no four-year degree.

the average increase in returns to college over the period.[31]

5 Alternative explanations

The evidence presented in Tables 1 to 6 conforms well to the implications of the model of endogenous technology adoption we presented in Section 2. However, such evidence does not imply that this model is correct since the data may be consistent with alternative interpretations. In this section we explore two plausible alternative explanations to the observations we highlighted. First, we discuss whether the observed change in wage patterns could be driven by increased trade integration. Second, we discuss whether these patterns could reflect ongoing capital skill complementarity – as modeled by Krusell, Ohanian, Rioss-Rull and Violante (2000) – instead of resulting from a more drastic change in production possibilities as emphasized by our model of technological revolutions.

5.1 Trade Integration

Some readers may have noticed that our model's predictions are identical to those of a simple model of a small economy going from autarky to free trade. There are no tariff barriers between U.S. markets, but trade between them is not costless. So a sudden fall in the cost of trading in the economic integration of U.S. labor markets over this period - perhaps enabled by information technology - could generate the same empirical patterns as our model. So how will be able to distinguish this trade integration story from the predictions of our model?

The answer is that if the cost of trading between cities really diminished significantly over the period, we ought to see a reallocation of industries across markets consistent with it. Industries should move to take advantage of the initial differences in factor mix (and prices) across cities which were sustained by the initially higher cost of trade between cities. A useful framework to evaluate this empirically is a version of Card and Lewis's (2006) identical decomposition of cross-sectional differences in skill mix into "between" and "within" industry components:

[31] Here is another comparison: a market which increased its 1980 college share (really, $\ln(S/U)$) by one standard deviation - about 30 log points - is predicted to have experienced two percentage points faster growth in the returns to college 1980-2000. A 30 log point increase is not out of the question: the average area experienced a 65 log point change in college share between 1980 and 2000.

$$\frac{CE_i}{N_i} - \frac{CE}{N} = \sum_j \frac{CE_j}{N_j}\left(\frac{N_{j,i}}{N_i} - \frac{N_j}{N}\right) + \sum_j \frac{N_{j,i}}{N_i}\left(\frac{CE_{j,i}}{N_{j,i}} - \frac{CE_j}{N_j}\right)$$

$\frac{CE_i}{N_i} - \frac{CE}{N}$ is a city's "excess college share," that is, the difference in college share between city i and the country as a whole. Industries are indexed by j, so $\frac{CE_j}{N_j}$ represents the college-intensity of sector j in the nation as whole, and $\frac{CE_{j,i}}{N_{j,i}}$ is the college-intensity of sector j in city i. A city's excess college share can be decomposed into, first, a ("between") term which reflects the city's excess of college-intensive industries and, second, a ("within") term which reflects the city's excess of college-intensity within sectors.

The decomposition above is cross-sectional, but we are interested in whether changes in industry mix are consistent with increased integration. Differencing the industry mix portion of the "within" term provides a measure the extent to which changes in a city's industry mix absorb its initial excess skill supply:

$$\sum_j \frac{CE_j}{N_j}\left[\Delta\left(\frac{N_{j,i}}{N_i}\right) - \Delta\left(\frac{N_j}{N}\right)\right]$$

For the purpose of empirical implementation, $\frac{CE_j}{N_j}$ represents the college equivalent share for industry j in 2000 and the term in brackets is equal to the change in employment share in industry j and city c compared to industry j in the nation as a whole between 1980 and 2000. This statistic tells us how much changes in industry mix in i are expected to increase demand for college equivalent labor in i relative to the nation as a whole. If geographic integration is occurring between 1980 and 2000, then this statistic will be high for cities with a high initial college share, as college-intensive industries disproportionately relocate to them to take advantage of the low price of skill labor in those markets, and it will be low for cities with a low initial college share. In the extreme where trade integration leads to shifts in industry mix which fully absorb initial differences in skill mix (and nothing else changes), it will be exactly equal to the city's excess college share $\frac{CE_i}{N_i} - \frac{CE}{N}$.

In order to evaluate the extent to which this is the case, Figure 7 plots this between industry statistic against 1980 excess college share. Differences in the between industry statistic tend to be small compared to the differences in college share: note the difference in scale in the y- and x-axis. Put another way, industry mix is evolving in roughly the same way in all markets, at least as it pertains to the inquiry about increased trade integration. There is also little sign that the growth in a city's college-intensive industries has a positive relationship with its initial college share. One of the least educated cities in our sample, Scranton, PA, has one of the larger increases in college-intensive industries between 1980 and 2000. More generally, the coefficient from the regression line, which can be interpreted as the proportion of initial differences in college share absorbed by changes in industry mix, is 0.01 with a standard error of 0.02.[32] One can also plot this statistic against each area's initial returns to skill. The trade integration story predicts a negative relationship between this and initial returns to skill: skill-intensive sectors will want to "move out" (in relative terms) of an area with initially high returns to skill to an area with a lower price of skill. The actual relationship with initial returns is slightly positive. Thus, at the level of industry detail available in the censuses (roughly three-digit SIC),[33] there is no evidence that cities are "integrating" over this period. While some research suggests that integration may show up in changes in product quality below the level of industry detail we can observe (Schott, 2004), it is not very encouraging for the trade integration view that there is virtually no action at the level of industry detail we do observe.

Another difficulty with the increased inter-city trade explanation is that, controlling for industry structure, there should be no systematic differences in PC-use across cities: cities with more educated workers should have more skill and PC intensive industries but should not use PCs more intensively within an industry.[34] However, as was documented in Table 1, there is a strong positive link between PC use within industries and the local supply of skill.

5.2 Capital-Skill complementarity

We now turn to examining whether the observation of greater increases in the returns to skill where skill is most abundant could be explained by a capital skill complementarity mechanism along the lines presented in Krusell, Ohanian, Rios-Rull and Violante (2000). In this framework, the driving force is a decrease in the price of equipment, but instead of affecting the choice of technology, it operates within a unique production structure $g(\cdot)$ of the form

[32] The regression used the same weights used in the regression presented in tables 1-6.

[33] Industry codes are more detailed than this in the 2000 Census. We use Lewis (2003)'s mapping between industry codes in the 1980 and 2000 Censuses.

[34] Recall that in the case of the PC data, industry is observed at an even higher level of detail than in the census - four-digit SIC.

$$g(PC,S,U) = ((c(PC^{\nu} + S^{\nu})^{\frac{\mu}{\nu}} + (1-c)U^{\mu})^{\frac{1}{\mu}}$$

where PC represent equipment. As in the main text, we let the price of this capital be given by r^{PC} and we assume that each locality can take this price as given when deciding how much equipment capital to allocate to production. In this framework, the return to skill in a local market is given by $\frac{w_i^S}{w_i^U} = \frac{g_S(PC_i,S_i,U_i)}{g_U(PC_i,S_i,U_i)}$, where the quantity of equipment (PC), is taken to be endogenously determined by the marginal product condition $g_{PC}(PC_i, S_i, U_i) = r^{PC}$. Hence, by solving out for the optimal level of PC, the returns to skill in a locality can be stated as a function of the price of equipment and the ratio $\frac{S_i}{U_i}$ in the locality. Let us denote this relationship as $\frac{w_i^S}{w_i^U}(r^{PC}, \frac{S_i}{U_i})$. The issue we want to address is whether, in this framework, a decrease in the price of equipment would have a greater or smaller effect on the returns to skill in a locality with a greater level of skill. In order words, we want to ask whether $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ is greater or small than zero under the capital skill complementarity hypothesis.[35] If it is smaller than zero, it could rationalize the observation of greater increases in the returns to skill where skill is most abundant. Recall that the capital-skill complementarity assumption, which requires that ν be less that μ and less than zero, guarantees that $\frac{\partial \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC}}$ is less that zero.

Deriving an explicit expression for $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ turns out to be quite difficult. However, in the particular case where $\mu = 1$ and therefore $g(\cdot)$ takes the form $g(PC,S,U) = (PC^{\nu} + S^{\nu})^{\frac{c}{\nu}} U^{1-c}$, it is possible to show analytically that $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ is greater than zero.[36]. Hence, in this case, we know that capital skill complementarity cannot generate greater increases in the returns to skill where skill is most abundant. In the more general case where $\mu \neq 1$, we examined the sign of $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ numerically. Looking over a very wide range of parameters considered relevant in the literature, we found no cases where capital skill complementarity could cause greater increases in skill where capital is most abundant. Accordingly, we quite confident to infer that capital skill complementarity does not offer a plausible explanations to the observations we presented.[37]

[35] Since returns to skills are determined by properties of the first order derivative of the production structure, $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ implicitly depends of third order derivative of the production structure.

[36] The proof of this statement, which is tedious but not difficult, is available from the authors

[37] In this discussion we focused on the sign of $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC},\frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$, while in the empirical section we focused

5.3 Other Explanations?

The above discussions indicates that our observations, especially our observation of greater increases in returns to skill where skill is most abundant, is not easily explained by either increased trade integration or by a simple model of capital-skill complementarity. However, there certainly remains – as is almost always the case–some alternative mechanisms that could explain the same observations. In this sense, we cannot claim that the cross city patterns in PC adoption and returns to skill imply that 1980-2000 period was a period of technological revolution reflection the endogenous adjustment to a new technological paradigm. Nevertheless, we believe that our model provides a set of relevant and quite demanding criteria by which one may want to assess the likelihood of a technological revolution. The interesting finding is that this set of criteria appears to be met over the 1980-2000, precisely at a point in time where many believe a paradigm shift has occurred.

6 Conclusion

The notion that certain periods correspond to technological revolutions is widespread. However it is often unclear what characteristics should be met for a period to be considered a technological revolution. A common idea is that such a period should reflect a paradigm shift in the method of production. Accordingly in this paper we used a simple neo-classical model of adjustment to a new technological paradigm to highlight a set of implications that could define a technological revolution. We then examined whether these implications were observed in the US over the period 1980-2000 when the PC was introduced and diffused. Our finding is that cross-city patterns regarding the returns to skill, the levels of education and the adoption of PCs conformed quite closely to the implications of the model. For this reason we believe that the period 1980-2000 may well deserve the designation of a technological revolution. One of the main implications of our model is that the link between supply of factors and their returns should be quite different during a technological revolution than in the long run. As we know from many studies, over the long run the supply of skill places downward pressure on the returns to skill. However, as implied by the model, during a period of adjustment to a new technological paradigm that is skill-biased this should not

on effects on changes in $\ln \frac{w_i^S}{w_i^U}$. To be consistent, we have redone verified that the observation of greater increases in skill where skill is most abundant is robust to whether or not skill is measured as $\frac{w_i^S}{w_i^U}$ or $\ln \frac{w_i^S}{w_i^U}$.

generally be the case. Instead, during such a period, we should observe that returns to skill increase most where skill is most abundant, and we should see increase in skill lead to a faster diffusion of the new technology without placing downward pressure of returns. Moreover, the model predicts that endogenous diffusion should not lead to a situation where returns to skill are highest in localities that have adopted the new technology most aggressively. Using a combination of data sources, we found evidence supporting all of these predictions.

Is it important to know whether a period is a technological revolution? We believe that the answer is yes since such knowledge is useful to better guide and evaluate policies. For example, the arrival of a major skilled biased technology as modeled here will cause an increase in the returns to skill and therefore increased wage inequality. From a historical data, a reasonable policy to combat such increased inequality is to favor greater accumulation of skill. However during a technological revolution, as we have shown, an increase in the supply of skill is unlikely to reduce wage inequality as the increase in skill instead acts simply to accelerate the adoption of the skill biased technology. In the absence of any recognition that the period may be a technological revolution, and that therefore historical relations may be temporarily inactive, one may evaluate policies inappropriately and make wrong choices. In particular during a period of a technological revolution, our analysis suggest that policy makers need to be very patient if they want to see increases in supply have negative impacts on prices. Moreover, during such periods policy makers may need to choose alternative policies than that suggested by long run patterns if one desires immediate result.

References

Acemoglu, Daron (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*, November 1998, vol 113, 1055-1089.

Acemoglu, Daron (2002). "Directed Technical Change." *Review of Economic Studies* 69: pp. 781-810.

Acemoglu, Daron (2006). "Equilibrium Bias of Technology." Massachusetts Institute of Technology mimeo, June.

Altonji, Joseph and David Card (1991). "Effects of Immigration on Labor-Market Outcomes of Less-Skilled Natives" in John Abowd and Richard Freeman, eds., *Immigration, Trade, and Labor*. Chicago: University of Chicago Press, p. 201-234.

Atkinson, Anthony B. and Joseph E. Stiglitz (1969). "A New View of Technological Change." *Economic Journal* 79(3): September, pp.573-78.

Autor, David, Lawrence F. Katz and Melissa S. Kearney (2006). "The Polarization of the U.S. Labor Market," mimeo.

Autor, David H., Frank Levy and Richard J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118 (4): November 2003, p. 1279-1334.

Basu, Susanto and David Weil (1998). "Appropriate Technology and Growth," *Quarterly Journal of Economics*, November, 113(4),pp. 1025-54.

Beaudry, Paul and David Green (1998). "What is Driving US and Canadian Wages: Exogenous Technical Change or Endogenous Choice of Technique?" (December), NBER working paper 6853.

Beaudry, Paul and David Green (2003). "The Changing Structure of Wages in the US and Germany: What explains the differences?" *American Economic Review*, June , pp 573-603.

Beaudry, Paul and David Green (2005). "Changes in U.S. Wages, 1976-2000: Ongoing Skill Bias or Major Technological Change?" *Journal of Labor Economics*, vol 23 , pp. 609-648.

Beaudry, Paul, David Green, and Benjamin Sand (2007). "Spill-Overs from Good Jobs," National Bureau of Economic Research Working Paper 13006, April.

Benhabib, Jess and Mark Spiegel (2005), "Human Capital and Technology Diffusion," forthcoming, Handbook of Economic Growth, Philippe Aghion and Steven Durlauf, eds., North-Holland, Amsterdam.

Berry, Chirstopher R. and Edward L. Glaeser (2005). "The Divergence of Human Capital Levels Across Cities." National Bureau of Economic Research Working Paper 11617, September.

Black, Dan, N. Kolesnikova and L. Taylor (2004). "Understanding the Returns to Education when Wages and Prices Vary by Location," working paper.

Card, David (1995), "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in Louis N. Chiristofides, E. Kenneth Grant , and Robert Swidinsky, Eds., Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp. Toronto: University of Toronto Press, p. 201-222.

Card, David (1999). "The Causal Effect of Education on Earnings," Handbook of Labor Economics. Volume 3A, 1999, p 1801-1863. Amsterdam, New York and Oxford: Elsevier Science, North-Holland.

Card, David and John E. DiNardo (2002). "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20 (4): October 2002, pp. 733-83.

Card, David and Ethan Lewis (2005). "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts." NBER Working Paper no. 11552, August 2005.

Caselli, Francesco (1999). "Technological Revolutions," *American Economic Review*, 89:1, pp. 78-102.

Caselli, Francesco and W. John Coleman (2006). "The World Technology Frontier." *American Economic Review* 96(3): June, pp. 499-522..

Caselli, Francesco and Daniel Wilson (2004). "Importing Technolgy." *Journal of Monetary Economics* 51(1): pp. 1-32.

Comin, Diego and Bart Hobijn (2004). "Cross-Country Technology Adoption: Making the Theories Face the Facts." *Journal of Monetary Economics* 51: pp. 39-83.

Dahl, Gordon (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6): pp. 2367-2420.

Doms, Mark and Ethan Lewis (2006). "The Diffusion of Personal Computers Across the U.S., 1990-2002." Working Paper Series 2006-18, Federal Reserve Bank of San Francisco.

Gerschenkron, A. (1962). *Economic Backwardness in Historical Perspective.* Cambridge, Belknap Press of Harvard University Press.

Goldin, Claudia and Larry Katz (2008). *The Race Between Education and Technology.* Cambridge and London: Harvard University Press, Belknap Press.

Goldin, Claudia and Kenneth Sokoloff (1984). "The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850." *Quarterly Journal of Economics*, 99 (August): pp. 461-488.

Habakkuk, H. J. *American and British Technology in the Nineteenth Century.* Cambridge: Cambridge University Press, 1962.

Kane, Thomas J. and Cecilia Elena Rouse (1995). "Labor Market Returns to Two- and Four-Year College," *American Economic Review* 85(3): June 1985, p. 600-614.

Kane, Thomas J. and Cecilia Elena Rouse (1999). "The Community College: Educating Students at the Margin Between College and Work." *Journal of Economic Perspectives* 13(1): Winter, pp. 63-84.

Katz, Lawrence F. and Kevin M. Murphy. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics* 112: pp. 291-303.

Krueger, Alan (1993). "How Computers Have Changed the Wage Structure: Evidence from Microdata 1984-1989." *Quarterly Journal of Economics* 108(1): pp. 33-60.

Krusell,Per, Lee E. Ohanian, Jose - Victor Rios-Rull, and Giovanni L. Violante *Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis*, Econometrica, 2000,1029-1055.

Lewis, Ethan (2003)."Local, Open Economies Within the US: How Do Industries Respond to Immigration?" Federal Reserve Bank of Philadelphia Working Paper no. 04-01, December 2003.

Lewis, Ethan (2005). “Immigration, Skill Mix, and the Choice of Technique.” Federal Reserve Bank of Philadelphia Working Paper, May 2005.

Manuelli, Rodolfo and Ananth Seshadri (2004), “Frictionless Technology Diffusion: The Case of Tractors.” working paper.

Moretti, Enrico (2004), “Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data,” *Journal of Econometrics* 121(1-2): July-August 2004, p. 175-212.

Nervis, Allan (1962). *The State Universities and Democracy.* Urbana, IL: University of Illinois Press, 1962.

Schott, Peter K. (2004). “Across-Product versus Within-Product Specialization in International Trade.” *Quarterly Journal of Economics* 119(2): May 2004, p. 647-678.

Welch, Finis (2000), “Growth in Women’s Relative Wages and in Inequality Among Men: One Phenomenon or Two?,” *Papers and Proceedings of the American Economic Association*, pp. 444-449.

Table 1: PC Adoption 2000 and 1980 Education

	OLS	OLS	IV:1940	OLS	OLS	OLS	OLS	OLS
	1	2	3	4	5	6	7	8
$\ln(\frac{S}{U})_{1980}$	0.15	0.14	0.16	-	0.11	-	0.11	-
	(0.01)	(0.01)	(0.02)		(0.02)		(0.02)	
$\Delta \ln(\frac{S}{U})_{1980-1940}$	-	-	-	0.12	-	0.09	-	0.09
				(0.02)		(0.02)		(0.02)
$\ln(\frac{S}{U})_{1940}$	-	-	-	0.14	-	0.12	-	0.11
				(0.01)		(0.02)		(0.02)
City Controls	-	-	-	-	Yes	Yes	Yes	Yes
Ind. Controls	-	-	-	-	-	-	Yes	Yes
Observations	230	157	157	157	157	157	157	157
R^2	0.44	0.45		0.45	0.60	0,61	0.64	0.64
F-Stat on Diff. in Coef.				2.35		5.95		1.99
p-value				0.13		0.02		0.16

The dependent variable is the number of PCs per worker at the city level corrected for industry composition. The variable $\ln(\frac{S}{U})_{1980}$ corresponds to the log of the ratio of college equivalent workers to high school equivalent workers. The instrument used in Column (3) is the ratio of college equivalent workers in 1940. The reported p-value is associated to the test of equality of coefficients on $\ln(\frac{S}{U}_{1940})$ and $\Delta \ln(\frac{S}{U})_{1980-1940}$. The city level controls correspond the the log of the labor force, the unemployment rate, the fraction of city population which is female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force.

Table 2: PC Adoption (2000) and Returns to Skill(1980)

	OLS	OLS	OLS	IV:1940	IV:1980	IV:1940 & Δ 1940-80
	1	2	3	4	5	6
$\ln(\frac{W^S}{w^U})_{1980}$	-0.27	-0.33	-0.19	-2.10	-2.69	-2.34
	(0.13)	(0.10)	(0.08)	(0.50)	(0.81)	(0.58)
City Controls	-	Yes	Yes	-	-	-
Ind. Controls	-	-	Yes	-	-	-
Observations	230	157	157	157	157	157
R^2	0.05	0.49	0.60	-	-	-
Over-ID test						0.30
p-value						0.58

The dependent variable is the number of PCs per worker at the city level corrected for industry composition. The variable $\ln(\frac{W^S}{W^U})_{1980}$ corresponds to the log of the ratio of the adjusted log wage gap between college and high-school educated workers. See Appendix 2 for details of the adjustment, which controls for a quartic in potential experience and dummies for female, foreign-born, and born after 1950. The instrument used in Column (4) and (5) are the relative abundance in skill in 1940 and 1980 respectively. Column (6) includes the change in the abundance of skill over the period 1940-80 as an extra instrument. The reported p-value is associated to the over-identifcation test. The city level controls correspond the the log of the labor force, the unemployment rate, the fraction of city population which is female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force.

Table 3a: Change in returns to skill, 1980-2000, and initial skill mix.

	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	1	2	3	4	5	6	7	8
$\ln(\frac{S}{U})_{1980}$	0.07	0.08	0.09	0.08	-	-	-	-
	(0.02)	(0.02)	(0.03)	(0.04)				
$\Delta\ln(\frac{S}{U})_{1980-1940}$	-	-	-	-	0.07	0.08	0.08	0.06
					(0.02)	(0.02)	(0.03)	(0.04)
$\ln(\frac{S}{U})_{1940}$	-	-	-	-	0.07	0.09	0.09	0.08
					(0.02)	(0.02)	(0.03)	(0.04)
City Controls	-	Yes	Yes	Yes	-	Yes	Yes	Yes
Ind. Controls	-	-	Yes	Yes	-	-	Yes	Yes
Endo. Location	-	-		Yes	-	-		Yes
Observations	157	157	157	157	157	157	157	157
R^2	0.18	0.45	0.50	0.37	0.18	0.45	0.50	.37
F-Stat on Diff. in Coef.					0.03	0.19	0.40	1.81
p-value					0.87	0.67	0.53	0.18

The dependent variable is the change in the adjusted (log) wage gap between college and high school educated workers. See Appendix 2 for details of the adjustment, which controls for a quartic in potential experience and dummies for female, foreign-born, and born after 1950. The variable $\ln(\frac{S}{U})_{1980}$ corresponds to the log of the ratio of college equivalent workers to high school equivalent workers. The reported p-value is associated to the test of equality of coefficients on $\ln(\frac{S}{U})_{1940}$ and $\Delta\ln(\frac{S}{U})_{1980-1940}$. The city level controls are the the log of the labor force, the unemployment rate, and the fractions of city population which are female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force. A yes in the row denoted Endo. location implies that we have applied the correction discussed in the text associated with the endogenous choice of location.

Table 3b: Changes in returns to skill, 1980-2000 and 1940/50-1980

Dep. Var:	OLS 80-2000	OLS 80-2000	OLS 80-2000	OLS 1940-80	OLS 1940-80	OLS 1940-80
	1	2	3	4	5	6
$\ln(\frac{S}{U})_{1980}$	0.06	0.07	0.09	-0.00	-0.06	-0.03
	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.06)
City Controls	-	Yes	Yes	-	Yes	Yes
Ind. Controls	-	-	Yes	-	-	Yes
Observations	157	157	157	157	157	157
R-squared	0.098	0.509	0.416	0.000	0.068	0.136

The dependent variable is the change in the adjusted (log) wage gap between college and high school equivalent workers over the period 1980-2000 (columns (1)-(3)) or 1940/50-1980 (columns (4)-(6)). Adjusted wages were computed using the sample of wage earners with at least 11 years of education by regressing individual log wages on a linear term in years of education, a quartic in potential experience, dummies for foreign-born and female, a vector of metro area dummies, and a vector of metro area dummies interacted with the number of college-equivalent units supplied by each individual. The estimated coefficients on the latter vector provided the adjusted returns. Because the sample which was asked to report their wage in the 1940 Census was small, it was combined with the 1950 census; in the stacked 1940/50 sample, log wages were also adjusted for a year effect.

The variable $\ln(\frac{S}{U})_{1980}$ corresponds to the log of the ratio of college equivalent workers to high school equivalent workers. The city level controls are the the log of the labor force, the unemployment rate, and the fractions of city population which are female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force.

Table 4a: Return to skill and skill supply: 1980, 1990, 2000

	1980 OLS	1990 OLS	2000 OLS	1980 OLS	1990 OLS	2000 OLS	1980 OLS	1990 OLS	2000 OLS
	1	2	3	4	5	6	7	8	9
$\ln(\frac{S}{U})_{1980}$	-0.07	-	-	-0.10	-	-	-0.16	-	-
	(0.01)			(0.01)			(0.02)		
$\ln(\frac{S}{U})_{1990}$	-	-0.06	-	-	-0.06	-	-	-0.10	-
		(0.01)			(0.01)			(0.02)	
$\ln(\frac{S}{U})_{2000}$	-	-	0.00	-	-	-0.02	-	-	-0.03
			(0.02)			(0.02)			(0.03)
City Controls?				Yes	Yes	Yes	Yes	Yes	Yes
Industry Controls?							Yes	Yes	Yes
Observations	230	230	230	230	230	230	230	230	230
R-squared	0.13	0.08	0.00	0.42	0.38	0.47	0.50	0.42	0.53

The dependent variable is the adjusted log wage gap between college and high school educated workers in 1980, 1990, or 2000. See Appendix 2 for details of the adjustment, which controls for a quartic in potential experience and dummies for female, foreign-born, and born after 1950. The variable $\ln(\frac{S}{U})_t$ corresponds to the log of the ratio of college equivalent workers to high school equivalent workers. The city level controls are the log of the labor force, the unemployment rate, and the fractions of city population which are female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force.

Table 4b: Return to skill and skill supply: Additional Specifications

	1980 OLS	2000 OLS	1980 OLS	2000 OLS	1980 IV: 1940	2000 IV:1940	1980 IV:1940	2000 IV:1940
	1	2	3	4	5	6	7	8
$\ln(\frac{S}{U})_{1980}$	-0.05	-	-0.05	-	-0.08	-	-0.07	-
	(0.02)		(0.02)		(0.02)		(0.02)	
$\ln(\frac{S}{U})_{2000}$	-	0.02	-	0.00	-	-0.01	-	-0.03
		(0.02)		(0.02)		(0.03)		(0.03)
Endo. Location	-	-	Yes	Yes	-	-	Yes	Yes
Observations	157	157	157	157	157	157	157	157
R^2	0.10	0.01	0.10	0.00				

The dependent variable is the adjusted log wage gap between college and high school educated workers in either 1980 or 2000. See Appendix 2 for details of the adjustment, which controls for a quartic in potential experience and dummies for female, foreign-born, and born after 1950. The variable $\ln(\frac{S}{U})_{1980}$ corresponds to the log of the ratio of college equivalent workers to high school equivalent workers. The instrument used in Columns (5)-(4) is the ratio of college equivalent workers in 1940. The regressions are weighted by the square root of the size of the labor force. A yes in the row denoted Endo. location implies that we have applied the correction discussed in the text associated with the endogenous choice of location.

Table 5: Changes in Returns to Skill 1980-2000 and PC adoption.

	OLS	OLS	OLS	OLS	IV:1940	IV:1980	IV:1940
	1	2	3	4	5	6	7
PCs/Worker 2000	0.18	0.23	0.17	0.07	0.45	0.50	0.67
	(0.07)	(0.09)	(0.06)	(0.06)	(0.14)	(0.13)	(0.31)
City Controls	-	-	Yes	Yes	-	-	Yes
Ind. Controls	-	-	-	Yes	-	-	Yes
Observations	230	157	230	230	157	157	157
R^2	0.05	0.09	0.27	0.35			

The dependent variable is the change in the adjusted log wage gap between college and high school educated workers. See Appendix 2 for details of the adjustment, which controls for a quartic in potential experience and dummies for female, foreign-born, and born after 1950. The variable PCs/worker is the number of PCs per worker at the city level corrected for industry composition. The instruments used in Column (5)-(7) is the ratio of college equivalent workers in 1940 or 1980. The city level controls correspond the the log of the labor force, the unemployment rate, the fraction of city population which is female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force.

Table 6: Returns to Skill 2000 and PCs in 2000

	OLS	OLS	OLS	OLS	IV:1940	IV:1980	IV:1940	IV:1940
	1	2	3	4	5	6	7	8
Pcs/worker 2000	-0.02	0.06	-0.08	-0.06	-0.03	0.12	-0.81	-0.70
	(0.07)	(0.09)	(0.05)	(0.06)	(0.15)	(0.12)	(0.43)	(0.45)
City Controls	-	-	Yes	Yes	-	-	Yes	Yes
Ind. Controls	-	-	-	Yes	-	-	Yes	Yes
Endo. Location	-	-	-	-	-	-	-	Yes
Observations	230	157	230	230	157	157	157	157
R^2	0.00	0.00	0.47	0.52				

The dependent variable is the (log) of the relative wage of college versus high school educated workers in 2000. The variable PCs/worker is the number of PCs per worker in 2000 at the city level corrected for industry composition. The instruments used in Column (5)-(7) is the ratio of college equivalent workers in 1940 or 1980. The city level controls correspond the the log of the labor force, the unemployment rate, the fraction of city population which is female, African-American, and U.S. citizens. The regressions are weighted by the square root of the size of the labor force. A yes in the row denoted Endo. location implies that we have applied the correction discussed in the text associated with the endogenous choice of location.

7 Appendix 1: Proofs

The competitive equilibrium of an economy faced with the choice between two techniques of production will solve the following program.

$$\max_{\gamma,\rho,K,PC} K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}}+PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}}-r^{K}K-r^{PC}PC$$

subject to $0 \leq \gamma \leq 1$, $0 \leq \rho \leq 1$, $0 \leq K$ and $0 \leq PC$

Recall that each local economy take the price of the two types of capital as given. (Note: In this appendix the subscript i on S_i and U_i are dropped for simplicity) The main first order conditions associated with this this problem are

$$(1-\alpha)K^{-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}} \leq r^{K}$$
$$(1-\alpha)PC^{-\alpha}[a((1-\gamma)S)^{\sigma}+(1-a)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}} \leq r^{PC}$$

$$\begin{aligned}
&If \quad 0<\gamma<1\\
&K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}a\gamma^{\sigma-1} = PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}b(1-\gamma)^{\sigma-1}\\
&and \quad if \quad 0<\rho<1\\
&K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}a\rho^{\sigma-1} = PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}b(1-\rho)^{\sigma-1}
\end{aligned}$$

In order to characterize the solution to the problem, let us define ϕ_L and ϕ_L implicitly by the following two equations

$$(\frac{1}{r^{K}})^{\frac{\alpha}{1-\alpha}}a[a+(1-a)(\frac{1}{\phi^{L}})^{\sigma}]^{\frac{1-\sigma}{\sigma}} = (\frac{1}{r^{PC}})^{\frac{\alpha}{1-\alpha}}b[b+(1-b)(\frac{1}{\phi^{H}})^{\sigma}]^{\frac{1-\sigma}{\sigma}}$$

and

$$(\frac{1}{r^{K}})^{\frac{\alpha}{1-\alpha}}(1-a)[a(\phi^{L})^{\sigma}+(1-a)]^{\frac{1-\sigma}{\sigma}} = (\frac{1}{r^{PC}})^{\frac{\alpha}{1-\alpha}}(1-b)[b(\phi^{H})^{\sigma}+(1-b)]^{\frac{1-\sigma}{\sigma}}$$

Under assumption 2, the following values for γ, ρ , K and PC satisfy the first order conditions, including the associated complementarity slackness conditions, and hence constitute a solution to the problem.

If $\frac{S}{U} < \phi_L$, then $\gamma = 1$, $\rho = 1$ and $PC = 0$ and $K = (1-\alpha)^{\frac{1}{\alpha}}(r^{K})^{\frac{-1}{\alpha}}[a(S)^{\sigma}+(1-a)(U)^{\sigma}]^{\frac{1}{\sigma}}$

If $\phi_L \leq \frac{S}{U} \leq \phi_H$, then $\gamma = \frac{\frac{\phi_H U}{S} - 1}{\frac{\phi_H}{\phi_L} - 1}$, $\rho = \frac{\frac{\phi_H}{\phi_L} - \frac{S}{\phi_L U}}{\frac{\phi_H}{\phi_L} - 1}$, $K = (1-\alpha)^{\frac{1}{\alpha}} r^{K\frac{-1}{\alpha}} [a(\gamma S)^\sigma + (1-a)(\rho U)^\sigma]^{\frac{1}{\sigma}}$ and $PC = (1-\alpha)^{\frac{1}{\alpha}} r^{PC\frac{-1}{\alpha}} [b((1-\gamma)S)^\sigma + (1-b)((1-\rho)U)^\sigma]^{\frac{1}{\sigma}}$

If $\frac{S}{U} > \phi_H$, then $\gamma = 0$, $\rho = 0$ and $K = 0$ and $PC = (1-\alpha)^{\frac{1}{\alpha}} (r^{PC})^{\frac{-1}{\alpha}} [b(S)^\sigma + (1-b)(U)^\sigma]^{\frac{1}{\sigma}}$

This characterization of the solution to the optimization problem will be used in the proofs of the Propositions and Corollaries.

Proof of Proposition 1: From the above solution to the maximization problem, we know that PCs per worker will be given by:

$If \quad \frac{S}{U} > \phi_H, \quad \frac{PC}{S+U} = (1-\alpha)^{\frac{1}{\alpha}} (r^{PC})^{\frac{-1}{\alpha}} [b(\frac{S}{S+U})^\sigma + (1-b)(\frac{U}{S+U})^\sigma]^{\frac{1}{\sigma}}$

which under assumption 1 is an increasing function of $\frac{S}{U}$

$If \quad \phi_L \leq \frac{S}{U} \leq \phi_H, \quad \frac{PC}{S+U} = (1-\alpha)^{\frac{1}{\alpha}} r^{PC\frac{-1}{\alpha}} [b((1-\gamma)\frac{S}{S+U})^\sigma + (1-b)((1-\rho)\frac{U}{S+U})^\sigma]^{\frac{1}{\sigma}}$

where $\gamma = \frac{\frac{\phi_H U}{S} - 1}{\frac{\phi_H}{\phi_L} - 1}$ and $\rho = \frac{\frac{\phi_H}{\phi_L} - \frac{S}{\phi_L U}}{\frac{\phi_H}{\phi_L} - 1}$

which again is an increasing function of $\frac{S}{U}$

and finally, if $\frac{S}{U} < \phi_L$, $\frac{PC}{S+U} = 0$.

Hence, PC per worker is a weakly increasing function of $\frac{S}{U}$.

Proof of Corollary 1: Since by Proposition 1, PCs per worker is an increasing function of a $\frac{S}{U}$ and that the initial ratio of skilled to unskilled wages is given by

$$\frac{w^S}{w^U} = \frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}$$

If follows that PCs per worker is an increasing function of the locality's initial ratio of skilled to unskilled wages

Proof of Proposition 2: Before the arrival of the new technology, the relationship between returns to skill and the supply of skill is given by

$$\ln(\frac{w^S}{w^U}) = \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}})$$

After the arrival of the new technology, the relationship is given by

$$\begin{array}{llll}
\ln(\frac{w^S}{w^U}) = & \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}) & if & \frac{S}{U} \leq \phi^L \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{a(\phi_L)^{\sigma-1}}{(1-a)}) = \ln(\frac{b(\phi_H)^{\sigma-1}}{(1-b)}) & if & \phi^L < \frac{S}{U} \leq \phi^H \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{bS^{\sigma-1}}{(1-b)U^{\sigma-1}}) & if & \phi^H < \frac{S}{U}
\end{array}$$

Hence the relationship between the initial supply of skill and the change in the returns to skill is given by

$$\begin{array}{llll}
\Delta \ln \frac{w^S}{w^U} = & 0 & if & \frac{S}{U} \leq \phi^L \\
\Delta \ln \frac{w^S}{w^U} = & (1-\sigma)[\log \frac{S}{U} - \log \phi^L] & if & \phi^L < \frac{S}{U} \leq \phi^H \\
\Delta \ln \frac{w^S}{w^U} = & (1-\sigma)[\log \phi^H - \log \phi^L] & if & \phi^H < \frac{S}{U}
\end{array}$$

This implies that the change in return to skill is positively related to the initial supply of skill.

Proof of Corollary 2: Corollary 2 follows directly by combining Propositions 1 and 2.

Proof of Proposition 3: After the arrival of the skill biased technology, the relationship between the supply of skill and the return to skill is given by

$$\begin{array}{llll}
\ln(\frac{w^S}{w^U}) = & \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}) & if & \frac{S}{U} \leq \phi^L \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{a(\phi_L)^{\sigma-1}}{(1-a)}) = \ln(\frac{b(\phi_H)^{\sigma-1}}{(1-b)}) & if & \phi^L < \frac{S}{U} \leq \phi^H \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{bS^{\sigma-1}}{(1-b)U^{\sigma-1}}) & if & \phi^H < \frac{S}{U}
\end{array}$$

Since this function exhibits a negative relationship between the supply of skill and the return to skill, the arrival of the skill bias technology cannot lead to a positive association between the return to skill and the supply of skill.

Proof of Corollary 3: Corollary 3 follows directly from Proposition 3 and 1.

Proof of Proposition 4: Before the arrival of the new technology, $\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}}$ is equal to $1-\sigma$. After the arrival of the new technology, is

$$\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} = 1-\sigma \quad if \quad \frac{S}{U} \leq \phi^L$$

$$\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} = 0 \quad if \quad \phi^L < \frac{S}{U} \leq \phi^H$$

$$\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} = 1-\sigma \quad if \quad \phi^H < \frac{S}{U}$$

Hence the effect of an increase in supply of skill on the returns to skill is (weakly) less after the arrival of the new technology. In particular, the effect is zero when the economy is in technology transition zone $\phi^L \leq \frac{S}{U} \leq \phi^H$.

Proof of Proposition 5:

Let us begin by proving point (2) of the proposition by contradiction. We now reintroduce subscripts on S and U since we need to compare different localities. As shown in the text, endogenous mobility implies

$$\log(\frac{S_i}{U_i}) - \log(\frac{S_j}{U_j}) = \frac{\psi}{\psi-1}[\log(\frac{d_j}{d_i}) + (\frac{1+\upsilon}{\upsilon})(\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}))]$$

which means that if locality i has a greater increase in returns to skill that locality j, and it has a lower initial supply of skill, it has a greater (log) increase in the supply of skill. Now suppose this is the case and locality i has a initial skill level that is lower than locality j, but has a higher (log) increase in supply of skill. This would imply that an increase in supply has a greater effect after the arrival of the new technology than before, which would contradict Proposition 4. Hence, point (2) of the proposition must hold.

From the proof of point (2) of the Proposition, we know that a city with an initially higher fraction of skill to unskilled workers will maintain its higher skill ratio after the arrival of the new technology and hence from point (1) of the Proposition 5 follows from the relationship between PC per workers and skill given in Proposition 1. Finally point (3) of Proposition 5 follows directly from Proposition 3 and Corollary 3.

8 Appendix 2: Wage Construction and Adjustments

All wage variables were constructed from public use micro data from the Census of Population (various years) using employed workers between the ages of 16 and 65, with at least one year of "potential experience" (age - years of education - 6), positive hours worked, not living in group quarters, and having either exactly 12 or 16 years of education. For each education level, regression-adjusted metropolitan area average log wages were constructed. We adjusted for a quartic in potential work experience, a female dummy, a foreign-born dummy, and a dummy for being born after 1950 (which is roughly when a trend break in returns to schooling occurs). An area's adjusted returns to college is the difference between its area's college and high school adjusted mean log wages.

In regressions in Tables 2-4 and in Table 6, wages were further adjusted for the selection of labor market using a procedure similar to the one outlined in Dahl(2002). In essence, the procedure involves estimating the probability of living in a particular location conditional on place of birth and other demographic characteristics, and then entering these estimated probabilities as a polynomial control function into the wage regressions (details below). Dahl allowed for separate control functions for two types of resident types: "stayers" - those living in their state of birth - and "movers." We add a third resident type, foreign-born, who were excluded from Dahl's analysis.

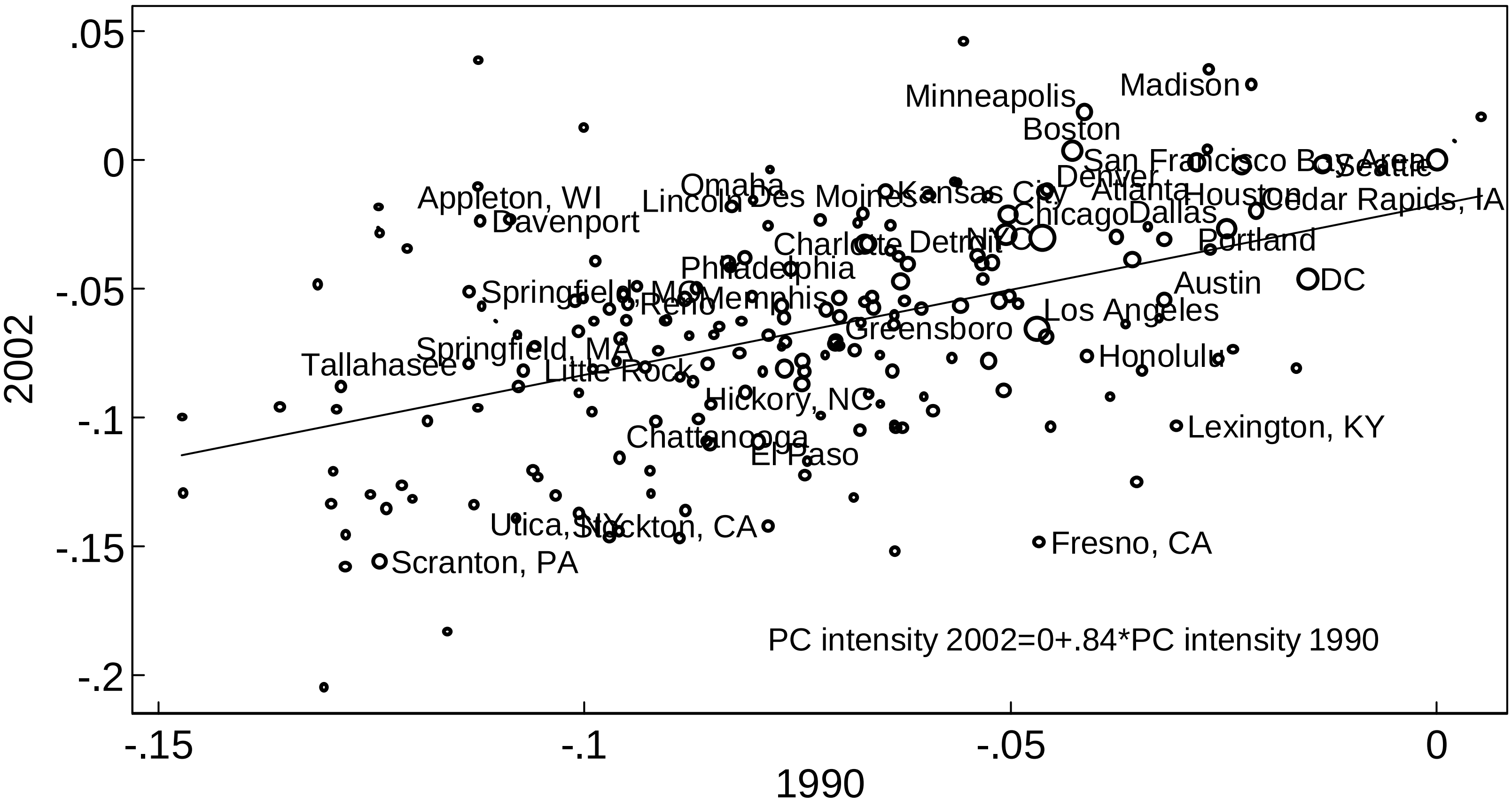
In our paper (and unlike in Dahl, 2002) metropolitan area, not state, is the unit of analysis. Metropolitan area of birth is not identified in the Census. To accommodate this, we follow Beaudry, Green, and Sand (2007) and define a "stayer" as any individual who lives in a metropolitan area which is at least partly in the individual's state of birth. For example, some counties in West Virginia are defined to be part of the Washington, DC metropolitan area. As a result, someone who was born in West Virginia and who lives anywhere in the DC area is considered a stayer.

Among the native-born (stayers and movers), an individual's probability of being located in the area was estimated as the share of people in the same demographic "cell" living in that metropolitan area. The demographic cells were defined by the interaction of two education categories (college, high school), two age categories (under/over 40), two race categories (white/nonwhite), gender, and state of birth. These shares were computed separately for movers and stayers. Among the foreign-born, the probability of living in a particular metropolitan area was allowed to vary across 18 region-of-origin groups: the 16 foreign regions used in Lewis (2003), plus the U.S. territory of Puerto Rico, plus all other foreign births (including other U.S. territories and U.S. natives born abroad of American parents). The immigrant portion of the selection correction thus corrects for endogeneity in a manner similar to the "supply/push"-type instruments used in research on the local labor market impact of immigration. Implicitly, the procedure treats Mexican immigration to border areas like Los Angeles as not very selected (the share of Mexicans living in LA is high) and treats Mexican immigration to Scranton, PA (where there are few Mexicans) as highly selected.

In the wage adjustment regressions, the control function for selection was again specified separately by education category. The function included dummies for resident type (mover, stayer, or foreign-born), dummies for all eight demographic category dummies (age x race x gender), and resident type-specific quadratics (following Dahl) in the estimated probability of being in the area.

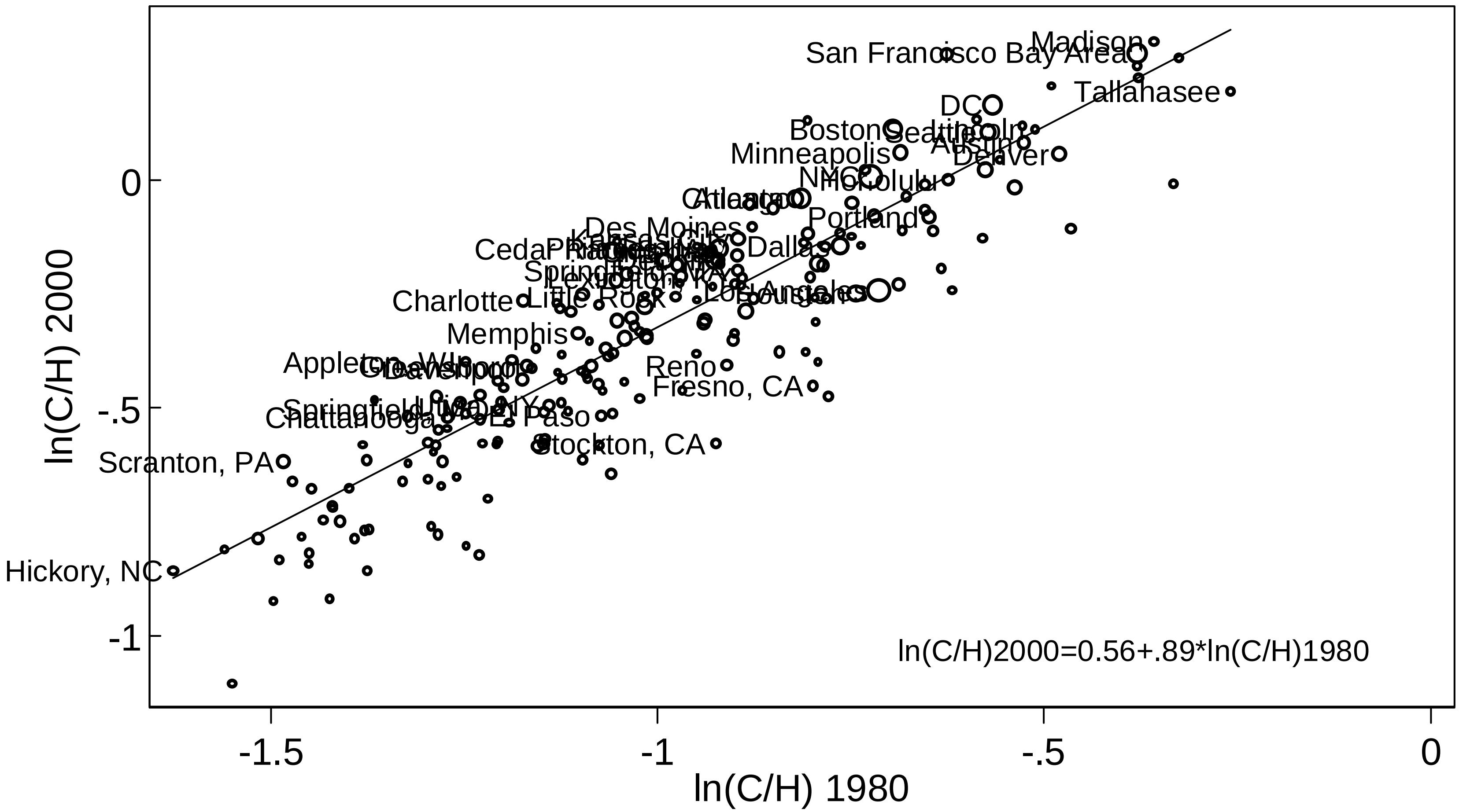
Figure 4: PCs per Employee by City in 1990 and 2000: Difference from the San Francisco Bay Area

(after controlling for industry and establishment size)



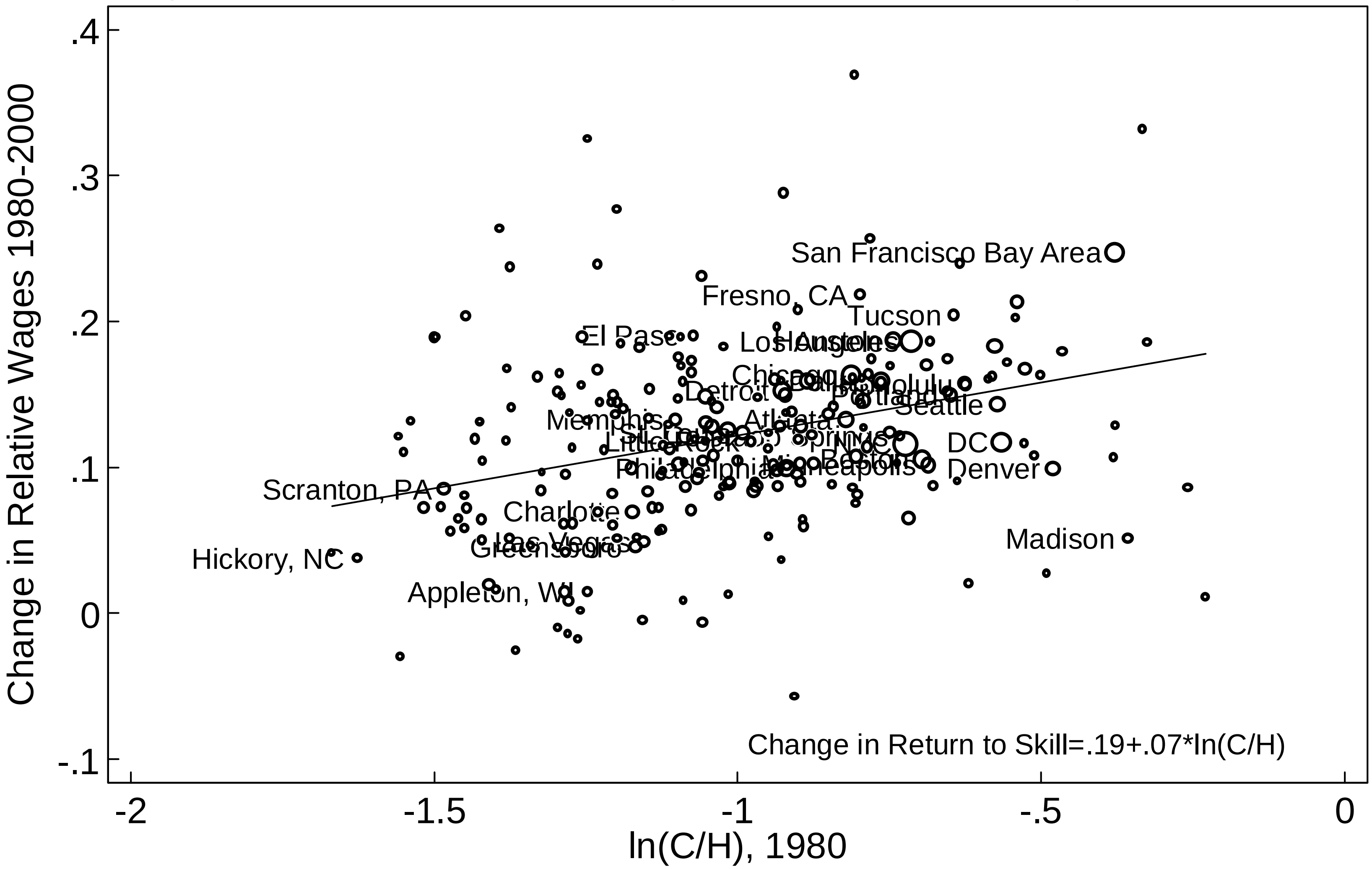
PC intensity computed from establishement-level regressions that control for 3-digit SIC industry interacted with 8 employment-size categories. The size of the markers in the figure are proportional to the square root of employment in the city. This same measure is used as weights in the regression. All data are from Harte-Hanks.

Figure 5: Log of College to High School Equivalents 2000 and 1980



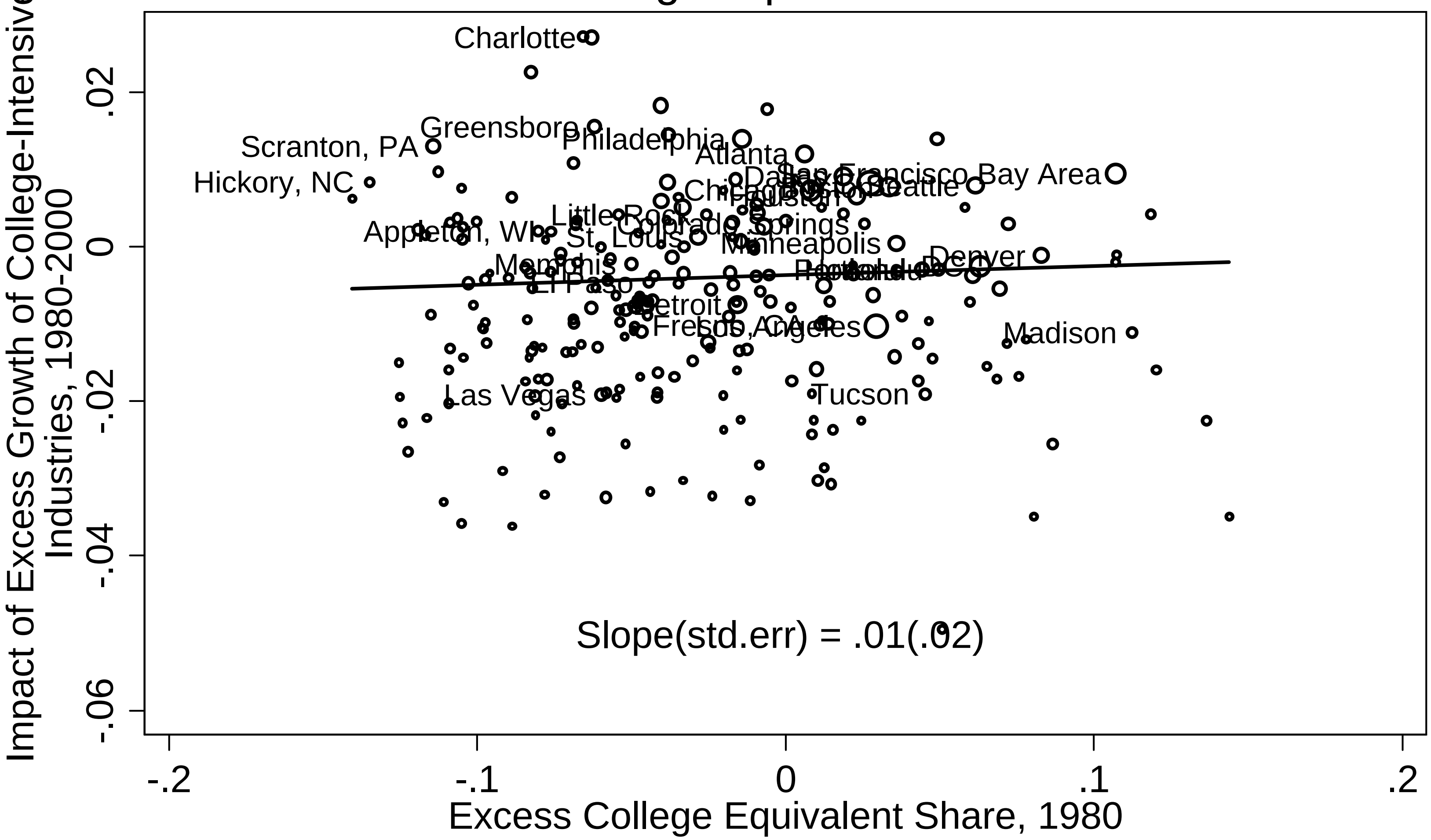
The size of the markers in the figure are proportional to the square root of employment in the city. This same measure is used as weights in the regression. College equivalents is defined as those with a college degree or more plus 1/2 of those with some college. All data come from the PUMS

Figure 6: Initial Skill and Changes in the Relative Wages, 1980-2000



C/H is the ratio of college equivalent workers to high school equivalent workers. The size of the markers in the figure are proportional to the square root of employment in the city. This same measure is used as weights in the regression.

Figure 7. Impact of Changes in Industry Mix on College Equivalent* Share



*College equivalent share is the share of workers with a college degree plus 1/2 of those with some college education but no degree. Excess share is the college equivalent share in the area minus the college equivalent share in all cities. The y-axis shows how much on college share; how this was calculated is described in the text. All calculations come from PUMS for 1980 and 2000.