DIRECTED TECHNICAL CHANGE:
AN EMPIRICAL INVESTIGATION

Alfonso Gambardella
Università Commerciale "L. Bocconi", Milan, Italy
agambardella@unibocconi.it

Marco Giarratana
Universitad Carlos III, Madrid, Spain
marco.giarratana@uc3m.es

First version: January 2004
This version: August 2005

We benefited from discussion or comments on earlier versions from Daron Acemoglu, Suma Athreye, Tim Bresnahan, Bruno Cassiman, Paul David, Giovanni Dosi, Andrea Fosfuri, Walter Garcia-Fontes, Bronwyn Hall, Sandeep Kapur, David Mowery, Woody Powell, Nathan Rosenberg, and Gavin Wright, along with seminar participants at the Sant'Anna School of Advanced Studies, Pisa, Carlos III, Madrid, the Science & Technology Workshop, SIEPR, Stanford University, the Innovation Workshop at the Haas Business School, Berkeley, and the EconChange Workshop in Barcelona. We thank Daron Acemoglu for providing us with the data on Child Labor, and Richard Florida and his research group for their data on the "tolerance index". Grid Thoma provided excellent research assistantship. Financial support from the DG *Science and Technology*, ECONCHANGE Project (Contract N. SERD-2002-00117) and the Italian Ministry of University and Research (MIUR), is gratefully acknowledged. The usual disclaimers apply.

Abstract

This paper employs a novel dataset of 146 US cities hosting industrial and technological activities to test some hypothesis about the determinants of skill premia. We argue that an important dimension of "directed technical change" is the creation of new organizational settings based on local knowledge spillovers and technological entrepreneurship as opposed to large integrated firms. We use measures of knowledge spillovers and the presence of large firms in our cities and find that, after controlling for endogeneity, they predict, respectively, higher and lower skill premia. We also find that in the long-run shifts in the relative supply of skills produce higher inequality between skills.

Keywords: Technology, Organization forms, Entrepreneurship, Wage Inequality

JEL: J31, O3

1. INTRODUCTION

Several studies have documented the growing inequality of earnings between skilled and unskilled workers in the US (e.g. Gottschalk, 1997; Gottschalk and Smeeding, 1997; Aghion, Caroli, Garcia-Peñalosa, 1999), and a leading explanantion has focused on skill-biased technical change (e.g. Berman, Bound and Machin, 1998; Machin and Van Reenen, 1998; Galor and Moav, 2000; Machin, 2000). In this context, Acemoglu (2002a and 2002b) has shown that while the classical effect of an increase in the relative supply of an input is to reduce its relative price, it also induces "directed technical change", viz. investments to make the input relatively more productive prompted by the opportunity to spread the higher productivity on a larger scale. Thus, an increase in relative supply can ultimately increase rather than decrease the relative marginal productivities of the inputs, and therefore their relative price.

The goal of this paper is to estimate these effects. Specifically, we estimate an equation in which the relative compensation of skilled and unskilled workers is a function of their relative supply and of variables that induce technical change biased towards one or the other input. The paper provides two contributions. First, it studies the impact of the relative input supply on inequality, and particularly whether directed technical change offsets the reduction in the relative productivity of the skilled workers produced by the increase in their relative supply. Second, it highlights an important source of skill-biased productivity improvements. Acemoglu himself (e.g. 1998) noted that the dominant production technologies or organizational forms of a given era have been biased towards skilled or unskilled workers (see also Goldin and Katz, 1998). For example, the rise of mass production demeaned the role of the skilled artisans and increased the relative productivity of the unskilled layman, while the opposite seems to be true of today's computer-based technologies (e.g. Bresnahan, 1999; Brynjolfsson, Bresnahan, and Hitt, 2001). Here we focus on the trends towards more entrepreneurial industrial and technological activities in the US since the end of the 1970s. Many high-tech and human-capital intensive industries now feature decentralized patterns of production and invention hinging on technology-based companies, entrepreneurship, clusters of firms, along with a focus on design, innovation, and skill-intensive activites (e.g. Audretsch and Feldman, 1996; Bhidé, 2000). This phenomenon, which is epitomized by Silicon Valley (Saxenian, 1994), and has spread to many regions of the world (Bresnahan and Gambardella, 2004; Arora and Gambardella, 2005), compares to the more traditional organization of the large integrated firms (Chandler, 1990). We will not enter into the debate whether one model is superior to the other, even though there are signs that the more entrepreneurial models may be occupying areas that were the province of the larger firms (e.g. Langlois, 2003). What is important for us is that they underlie different organizational modes with different potential sources of productivity gains.

As a stylized representation, the "Chandlerian" firms rely on size and product diversification – economies of scale and scope, as Chandler put it (see also Henderson and Cockburn, 1996). Size basically means that the company replicates given activities on a larger scale, which requires more routinary and repetitive tasks based on a structured division of labor (whether at the production, managerial or other levels). Similarly, economies of scope stem from synergies across existing firm assets, which may in part be a creative task, but relies to a good extent on making connections among existing products or resources. Moreover, in the Chandlerian world economies of scope are often attained through physical rather than human capital. By contrast, the clusters of specialist technology firms are by their own nature creative and design- or human-capital intensive. They rely on systematic generation of new designs rather than large scale replicas of the same patterns. It is then natural that this model exhibits a higher relative demand of more creative or educated workers. Even within the same industry, biotech specialist firms focus on R&D, while the large pharmaceutical companies also carry out drug testing, production and marketing. Similarly, in semiconductors, fabless or chipless firms exhibit a higher relative demand for skilled technological workers compared to larger semiconductor producers which also carry out other (more routinary) production and commercial activities.

Ultimately, because the technology-based entrepreneurial model relies to a greater extent on human capital, it will have greater incentives than the large Chandlerian firms to adopt technologies or organizational set-ups that improve the relative productivity of the skilled workers. Thus, for a given relative supply of skilled and unskilled labor, we expect that markets populated by many technology-based entrepreneurial firms will exhibit greater inequality between skills than markets in which there is a greater presence of large integrated firms. Our explanation also fits the evidence that the college premium first declined during the 1970s following the increase in the supply of skills. At that time the Chandlerian model was fully in place, and it was not common to leave one's employer to form a new firm. When the new model developed in the 1980s and the 1990s, the productivity of the skilled workers increased, and so did their premium (Levy and Murnane, 1992; Topel, 1997; Acemoglu, 1998). This also confirms that organizational changes are slower than technological changes (e.g. David, 1990). The organizational innovations that we are discussing took time to materialize, favoring the observed wage trends.

To summarize, our point is that Silicon Valley was itself a skilled-biased organizational innovation which ensued from the rise in the supply of educated workers in the US after 1950-1960. There are many stories of individuals whose inventions were not of interest to the (large) companies in which they were employed and left to create their own firms. This is because the Chandlerian model did not need so many skilled workers or so many new ideas. Its comparative advantage was in producing and commercializing fewer inventions on a large scale or in creating synergies with other existing productions within the firm, which required ordinary tasks along with workers who carried them out in a methodic rather than creative way. Thus, in the Chandlerian world a large supply of skills was redundant, with implied lower relative marginal productivity and lower returns. The Silicon-Valley model enabled the skilled individuals to escape this trap, as brilliantly discussed by Moore and Davis (2004) who highlight the features of a system in which engineers or scientists could try their ideas by giving birth to their own ventures. Moreover, the greater relative abundance of skilled workers "directed" their efforts towards an organizational model and related technologies that were human-capital intensive and that relied on local knowledge externalities arising from new firm formation, spin-offs, and human capital mobility across companies and institutions. This was again in contrast with the Chandlerian system which apart from being intensive in physical capital or routinary tasks, relied predominately on information generated within the firm (Henderson and Cockburn, 1996; Christensen, 1997).

To test these ideas, we constructed a dataset of 146 US cities. City-level data – as opposed to firm-level – are necessary because our study is about market wages. Moreover, cities seem to be an appropriate level of aggregation because systems of technology-based firms are often geographically localized, like Silicon Valley or the other regional clusters in or outside the US. Similarly, one or few large companies often affect the economy of specific locations (e.g. Agrawal and Cockburn, 2003). Moreover, even if there is a secular trend towards more entrepreneurial organizations of production, such a transtion is not instantaneous. Thus, different cities may entail a different degree of entrepreneurial vs Chandlerian activity, which provides a source of variation in the implementation of skill-biased productivity improvements.

The next Section presents the structure of the estimated equations. Section 3 discusses our empirical specifications, along with our proxies and instruments. Section 4 presents sample, variables, and the results. Section 5 concludes. There is a Data Appendix and a second Appendix reporting some robustness checks.

2. Relative Input Price Equation

The equations that we estimate stem from standard models of the formation of skill premia (Acemoglu, 2002a and 2002b). These models assume that the utility function of the consumers take the standard Dixit-Stiglitz form $U = \left(\sum_{i=1}^{n} \alpha_i q_i^{\rho} \right)^{\frac{1}{\rho}}$ where the q's are the consumption goods, the weights α sum up to 1, and $0 < \rho \leq 1$. Each q is produced by a different sector, which is assumed to specialize in the use of one input. The production function of the ith sector is $q_i = const \cdot y_i^{\beta} \left(\int_0^{A_i} x_{ij}^{1-\beta} dj \right)$, where y_i denotes the input, the x's are specialized machines that contribute to the productivity of the ith sector (and therefore of the ith input), and as customary in these models technical change is represented by the number of machines A_i employed by the industry. The firms in the sectors are price-taker. They maximize profits by choosing y_i and the x_i's given the price of the good p_i, the price of the input w_i, and the price of the machines r_i. The producers of the machines are monopolists who maximize profits by taking into account the optimal demand of the sectors. With a unit cost c_i for producing the x_i's, the price r_i is simply a mark-up on c_i. The profits of the machine producers serving the ith sector are $\pi_i = const \cdot p_i^{\frac{1}{\beta}} \cdot y_i$. They increase with the price of the good p_i and the supply of the input y_i.

The model is solved by first taking the ratio between the consumers demand for any two goods. The first order conditions of the consumers' problem yields $\frac{p_i}{p_j} = \frac{\alpha_i}{\alpha_j} \left(\frac{q_i}{q_j} \right)^{\rho-1}$.

After replacing the optimal demands for the x's in the production function of the q's, we have a second relationships between the ratios of any two prices and quantities,

$\frac{q_i}{q_j} = \left(\frac{p_i}{p_j}\right)^{\frac{1-\beta}{\beta}} \cdot \frac{A_i}{A_j}\frac{y_i}{y_j}$. Together these two conditions give the ratio of the prices of the two goods as a function of the ratios between the weights, the *A*'s and the inputs, viz.

$\frac{p_i}{p_j} = \left(\frac{\alpha_i}{\alpha_j}\right)^{\frac{\beta}{\beta\rho+(1-\rho)}} \left(\frac{A_i}{A_j}\frac{y_i}{y_j}\right)^{-\frac{\beta(1-\rho)}{\beta\rho+(1-\rho)}}$. From the first order conditions of the firms in the sectors, the ratio between the prices of any two inputs is $\frac{w_i}{w_j} = \left(\frac{p_i}{p_j}\right)^{\frac{1}{\beta}} \cdot \frac{A_i}{A_j}$. By replacing the expression for the ratio of the prices of the goods one obtains

$$\frac{w_K}{w_L} = \left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\gamma}{\sigma}} \left(\frac{A_K}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{K}{L}\right)^{-\frac{1}{\sigma}} \qquad (1)$$

where *K* and *L* represent the supply of skilled and unskilled workers and their indices, and we reparameterized the model so that $\sigma \equiv \beta\left(\frac{\rho}{1-\rho}\right)+1$ is the elasticity of substitution between the two inputs and $\gamma \equiv \frac{1}{1-\rho}$ is the elasticity of substitution between the two goods associated with the two inputs.

Equation (1) embodies several effects. First, the direct effect of an increase in the relative supply $\left(\frac{K}{L}\right)$ is to reduce inequality (the standard input-substitution effect). Second, inequality increases with the relative consumer demand for the good that uses *K* more intensively, as implied by the ratio between the two weights. Third, the effect of technical progress $\left(\frac{A_K}{A_L}\right)$ depends on the elasticity of substitution between the two inputs. If they are gross complements ($\sigma < 1$), technical progress that makes *K* relatively more productive induces a higher demand for unskilled labor, whose ultimate effect is to reduce the ratio of the two prices. By contrast, if they are gross substitutes, it further raises the demand for K which increases inequality.

In the long-run, $\left(\frac{A_K}{A_L}\right)$ is also affected by the relative supply. The producers of the machines will allocate their resources between the production of machines for the L- and K-intensive sectors till the marginal profitability of the two sectors are equal. In the steady state this is equivalent to equalizing their profits. Acemoglu (2002b) shows that this yields

$$\frac{A_K}{A_L}=\left(\frac{v_K}{v_L}\right)^{\sigma}\left(\frac{\alpha_K}{\alpha_L}\right)^{\gamma}\left(\frac{K}{L}\right)^{\sigma-1} \tag{2}$$

where the v's are factors that affect technical change biased towards one or the other input. Note that with gross substitution an increase in the relative supply induces technical change directed towards the more abundant input. This is because the gross substitution induces a higher demand for K which implies that the higher productivity of the machines used in the K-sector will spread on a larger scale.

By replacing (2) in (1), we obtain

$$\frac{w_K}{w_L}=\left(\frac{\alpha_K}{\alpha_L}\right)^{\gamma}\left(\frac{v_K}{v_L}\right)^{\sigma-1}\left(\frac{K}{L}\right)^{\sigma-2} \tag{3}$$

which is the long-run expression for the input price ratio. Equation (3) also shows that if the gross substituton is strong enough ($\sigma > 2$) the long-run effect of an increase in relative supply is to increase inequality.[1]

[1] Acemoglu (2002b) develops a more general version of the model which shows that with state-dependece in the innovation possibility frontier an upward sloping effect of increases in the relative supply occurs for $\sigma > 2 - \delta$ where $\delta \in [0, 1]$ is the parameter accounting for state-dependence. Thus, the long-run increase in inequality can also occur for $\sigma < 2$.

3. EMPIRICAL SPECIFICATIONS, PROXIES, AND INSTRUMENTS

3.1 Specification I

We estimate two empirical specifications. In Specification I, which we discuss in this Section, we assume that we do not observe A_K and A_L, but we observe the variables that affect $\frac{A_K}{A_L}$, viz. $\frac{\alpha_K}{\alpha_L}$, $\frac{v_K}{v_L}$, and the relative supply $\frac{K}{L}$. Specification II, which we discuss in Section 3.3, assumes that we observe A_K and A_L. Specification I amounts to estimating (3), while Specification II amounts to jointly estimating (1) and (2). In both cases we have to take into account the endogeneity of $\frac{K}{L}$. Compared for instance to countries, both skilled and unskilled workers can move easily across towns in response to incentives of various sorts, including changes in the price of their services. In Specification II we also have to take into account the endogeneity of $\frac{A_K}{A_L}$. We will discuss our instruments in Section 3.4.

To obtain our empirical specification of (3), let $w \equiv \log\left(\frac{w_K}{w_L}\right)$, $\alpha \equiv \sum_i \phi_i \log\left(\frac{\alpha_K}{\alpha_L}\right)_i$, $v \equiv \sum_i \eta_i \log\left(\frac{v_K}{v_L}\right)_i$, and $k \equiv \frac{K}{L}$, where $\left(\frac{\alpha_K}{\alpha_L}\right)_i$ are factors that affect the relative weights of the demand for K- and L-intensive productions in the cities, $\left(\frac{v_K}{v_L}\right)_i$ are factors that affect productivity biases towards one or the other factor, and ϕ and η are parameters. Since we have no way to distinguish between factors that affect demand *vs* productivity biases, we assume that the α and v factors are the same, and we denote the log of these ratios as ξ_I. We then rewrite $\alpha \equiv \sum_i \phi_i \xi_{1i}$, and $v \equiv \sum_i \eta_i \xi_{1i}$. The equation that we estimate is

$$w = const + \sum_i \chi_{1i}\xi_{1i} + (\sigma - 2)k \qquad (4)$$

where $\chi_{1i} \equiv \gamma\phi_i + (\sigma - 1)\eta_i$ is a set of parameters to be estimated. We cannot identify γ, but we can identify the elasticity of substitution σ. We can therefore assess whether the ultimate effect of an increase in the relative supply of skills is to increase inequality ($\sigma > 2$).[2] In addition, we can check the internal consistency of the model. We will distinguish between variables that account for productivity improvements biased towards skilled or unskilled labor. If $\sigma > 1$, we expect that the former have a positive impact on inequality, while the latter have a negative impact. The reverse will be true if $\sigma < 1$.

We also estimate (4) jointly with an equation for the relative supply of skills, viz.

$$k = const + \sum_i \chi_{2i}\xi_{2i} + \chi_w w \tag{5}$$

where ξ_2 are factors that affect the location of skilled *vs* unskilled workers, the χ_2 are the corresponding parameters to be estimated, and χ_w is the elasticity of the relative supply of skills.

3.2 Empirical Proxies for Chandlerian *vs* Silicon-Valley Cities

Before deriving Specification II we discuss our empirical proxies for the intensity of the two models, Silicon-Valley and "Chandlerian-ness", in our sample cities. As noted in the Introduction a typical characteristic of the technology-based entrepreneurial clusters is local knowledge externalities, which stem from the systematic rise of new firms, spin-offs, and human capital mobility across companies and institutions (e.g. Saxenian, 1994; Audretsch and Feldman, 1996). By contrast, the large firms tend to exploit information generated within their organization boundaries and they use outside knowledge and information less intensely (Henderson and Cockburn, 1996; Christensen, 1997).

[2] As noted in the previous footnote, 2 is an upper bound for σ to induce long-run increases in inequality.

We then employ the share of citations in the patents of a given city to patents produced in the same location as a proxy for the extent to which the area is characterized by a technological entrepreneurship model. The details of this variable are in the Appendix. Patent citations are often used to measure knowledge flows (e.g. Jaffe, Trajtenberg and Henderson, 1993; Hall, Jaffe and Trajtenberg, 2001). We therefore take local citations as evidence of knowledge spillovers and interactions among organizations in the same locations. We will use this proxy along with others capturing the importance of technological activities in the area, like trademarks, number of patents, share of high-tech output of the region. These variables are common to Chandlerian and Silicon-Valley cities, as they account for the importance of technological activities. Our local citations variable is the only one that, given the technological activities of the area, enables us to distinguish whether they are performed through one or the other model.

We also employ variables that account for the importance of Chandlerian firms. Our key proxy will be the share of firms with more than 1500 employees located in the city. Alternatively, we employed the average firm size in the city, but this variable is less informative because the average firm size is affected by the many very small operations that one can find in all cities, and thus varies less markedly with the presence of larger firms. In addition, we employed the share of trade union members on the total employment in the State of the city and the percentage of votes to George Bush in the city during the 2000 Presidential election. The share of trade union members is another proxy for the presence of large firms, as unions are more important in companies of larger size. It also ensures that our share-of-large-firm variable does not affect the wage ratio because it proxies for the unionization of the area. Unfortunately, the share of trade union members is only available at the State level. This is why we also employed the Bush measure, which is a city-level measure, and accounts for the fact that "big business" tends to be associated with Republican cities.

To summarize, in Specification I local citations proxy for the factors that affect the unobserved A_K, i.e. α_K and v_K, while the share of large firms, unionization and Bush's votes proxy for the factors that affect the unobserved A_L, i.e. α_L and v_L. In Specification II, instead, we use local citations as a measure of A_K and the three other proxies as measures of A_L.

3.3 Specification II

We assume that $A_K = const \cdot Z_K^{\mu_K}$, where Z_K is our citation proxy for skill-bias technical change and μ_K is a parameter. Similarly, $A_L = const \cdot Z_L^{\mu_L}$, with μ_L being another parameter, and

$$z_L = \lambda_{SHARE1500} \cdot LOG(SHARE1500) + \lambda_{SMEMB} \cdot LOG(SMEMB) + (1-\lambda_{SHARE1500}-\lambda_{SMEMB}) \cdot LOG(BUSH) \quad (6)$$

where $z_L \equiv log(Z_L)$ is a weighted average of the logs of the three variables that compose the index (share of large firms, unionization, Bush's vote), and the λs are their weights which will be estimated. By using the notation of Section 3.1, equation (1) becomes

$$w = const + \sum_i \frac{\gamma_i}{\sigma} \xi_{1i} + \frac{\sigma - 1}{\sigma} \left(\mu_K z_K - \mu_L z_L \right) - \frac{1}{\sigma} k \quad (7)$$

where $\gamma_i \equiv \gamma\phi_i$ and $z_K \equiv logZ_K$. We can estimate $\gamma\phi_i$ (but not γ alone), σ, μ_K and μ_L. That is, we can assess the extent of technical change biased towards skilled or unskilled labor. Like in the previous specification, k is endogenous, and we used instruments to identify σ.[3]

We make the important assumption that while A_K is endogenous, the index for A_L is not. The reason is that the formation of technology clusters is a more recent phenomenon, and the size of the cluster can typically expand or retract according to relatively short-run factors

[3] In the empirical analysis we also tried to estimate our equations by leaving the three variables unconstrained with no change in the results.

like cyclical expansions or retractions in the formation of new firms, the availability of venture capital, technological opportunities and the rate of innovation. In other words, they can be a response to shocks that may contemporaneously affect salaries and the relative supply of labor. By contrast, the establishment of large companies in a region is a long-term phenomenon. As mentioned earlier, large companies in specific industries have made the tradition of certain locations, and it is unlikely to observe abrupt changes in the presence and position of these firms in the area. Correspondingly, one is unlikely to observe rapid changes in the extent to which the area pursues skill *vs* unskill-biased productivity improvements.

To deal with the endogeneity of z_K we use (2) after replacing $\left(\frac{A_K}{A_L}\right)$ with the z. By solving for z_K one can re-write (2) as

$$z_K = const + \sum_i \frac{\gamma_i + \sigma\eta_i}{\mu_K}\xi_{1i} + \frac{\sigma - 1}{\mu_K}k + \frac{\mu_L}{\mu_K}z_L \qquad (8)$$

Specification II then jointly estimates (7) and (8). We also estimate (7) and (8) together with the relative supply (5).

3.4 Instruments

We employ two key instruments for the relative supply of skills. The first one is a composite variable that summarizes a set of State laws for allowing children to obtain a work permit. The second variable is the percentage of population in the city with English ancestry.

The Child Labor (CL) variable was constructed by Acemoglu and Angrist (2000) who employed it as an instrument for the average level of schooling in a State. It is the minimum years that a child has to attend school before obtaining a work permit. They collected these data since 1914, and discuss extensively how CL provides a good proxy for the extent to which a given State supported the growth of education. We use this measure for 1959 (but any other year in the 1950s worked equally well) which is well before the period to which our regressions refer. We can safely argue that the distance between the time in which the Laws were issued and the phenomena that we are trying to explain with our regressions is large enough to make the creation of the Laws exogenous for our purposes. The CL variable has become more similar across States in recent years, as one would expect given that all the US States, along with many other countries, have now recognized the importance of widespread youth education. We posit that differences in the early perception of the importance of youngsters education may reflect exogenous differences in the relative supply of skills across US States. Acemoglu and Angrist constructed another variable, the minimum years of compulsory school attendance. We tried it as an instrument as well but found that it did not perform as well as the Child Labor Laws. We then focussed on the latter.

Since the CL variable varies across States but not cities, we sought variables that moved across the latter as well to increase the power of our instruments. The rationale for using the US population with English ancestry is that it has a higher share of educated individuals than the population as a whole. It is worth noting that we are using ancestry characteristics, not contries of birth that is a measure typically used by studies on immigration and cultural diversity (e.g. Ottaviano and Peri, 2004). US Census data released in 1998 show that the share of US population with English ancenstry holding a bachelor degree or higher is 28.4% against 20.3% for all the ancestry groups.[4] Similarly, by combining these data with data for the labor force by ancestry group,[5] we computed the ratio between the population with bachelor degree or higher and the labor force net of the population with bachelor degree or higher, which mimics our relative supply $\frac{K}{L}$. This ratio is 33.1% for the population with English ancenstry *vs* 20.2% for the entire population. Further evidence is provided by the

[4] www.census.gov/population/socdemo/ancestry/table_01.txt

[5] www.census.gov/population/socdemo/ancestry/table_03.txt

sociological literature. In Alba and Moore (1982), "protestants with origins from the British Isles (colloquially, 'WASPs')" represent with 22.9% the most important ethnic group of their sample of 2,480 American men appointed to elite jobs. This percentage raises to 31% if we restrict their sample to the individuals with a four-year college degree. In their conclusions, the two authors find that WASPs are over-represented in elite occupations and particularly in elite occupations in business. This is confirmed by more recent studies (e.g. Kalmijn and Kraajkaamp, 1996), and the pattern has even reinforced as discussed by recent reports in the press (*New York Times*, 2005a and 2005b). In sum, people with English ancestries seem to be more inclined towards education for reasons partly unrelated to short-run demand-for-skill factors (social reasons, values, etc.) showing a more rigid and stable demand for education.[6]

To be sure, the share of English ancenstry in the current population of the cities could be endogenous. As educated workers move to cities with a higher demand for skills, and these workers are to a greater extent of English origin, the higher share of English ancestry in the city may be a consequence rather than a cause. But the presence of social groups in certain locations, particularly for groups that have come to the US for many generations, is unlikely to be explained by short-run migration patterns. It is hard to think that a certain US city has a British root because of recent trends in the demand for skills in the area. It is more natural that English origins have been a long-standing characteristic of the city, and this influences the supply of skills because English families have a greater bent for education.

We collected a few other instruments. This was necessary for Specification II where we add A_k as an endogenous variable and therefore need more instruments to support the more complex econometric structure. Moreover, in Section 4.4 we also endogenize per capita income, which further enhances our quest for instruments. To preserve comparison we employed the full set of instruments in all our specifications. Our first additional instrument is

[6] We also tried the share of the population with English, German or Irish ancenstry as an instrument, instead of just English. The results did not change.

Florida's (2002) "tolerance index", a measure of the political, social and cultural tolerance of the US cities, and a set of variables summarizing the town weather. As discussed in our Data Appendix the tolerance index combines the share of the gay population in the city with measures of melting pot, inter-racial integration, and "bohemian-ness". Florida discusses why tolerant cities attract educated people. Our prior is that the cultural traits of the cities sink their roots well into the past, and they come before the location of human capital, which may stem from relatively "recent" events like the rise of information technology.[7] In short, we posit that the tolerance of the city is another instrument for the supply of skills. We also collected two measures of the cities' weather: their average annual temperature and their average monthly days of clear sky. We do not have any particular rationale for these measures. Weather variables are clearly exogenous to economic conditions. The warmer Southern US regions are less high-tech oriented and less inclined to technological entrepreneurship. Thus, the average annual temperature separates them from the Northern regions, and the days of clear sky single out California.

4. Empirical Results

4.1 Sample and variables

Since we are interested in assessing the effects of different firm types, we constructed a sample of cities with a fair amount of variation along this dimension. We selected US cities that hosted Fortune 500 US companies or companies in the INC list of the 500 US fastest growing private, not-quoted firms in 1998, 1999 or 2000.[8] Both lists are used extensively by the literature as references for managerial and young firms. We selected the top 100 cities in

[7] As evidence of the stability of the cultural attitudes or ideals across cities, take for example the stability of the political vote. Cities that are clearly democrat or republican have been so for many years (e.g. in Mass. or Texas). Cities that are mixed, have been mixed for an equally long time.

[8] To be eligible for the 500 INC. list, a US company should: a) be privately held, not public or subsidiary, not a holding, regulated bank or utility; b) have at least a five years sales history with sales more than 200,000$. See http://www.inc.com/resources/inc500/index.html.

each list ranked respectively by number of Fortune 500 or INC firms. Since a city could be in both lists, we ended up with a sample of 146 cities.[9] The rationale for this criterion is that we wanted to have cities in which there was enough industrial activity to be meaningful for our analysis.[10]

Table 1 lists all the variables that are employed in our analysis. Appendix 1 describes our data and their sources. Table 2 provides descriptive statistics.

TABLES 1 AND 2 ABOUT HERE

Our measure of the relative supply of skills is the ratio in 2000 of the population in the city with a four year university degree and its total population net of the population with a four year degree. The creation of proxies for the wage of skilled and unskilled workers was not a trivial matter because wage data by skills or educational levels were not available for our cities. In the end, we thought that the best way to proxy for them was to use the wages of managers and production workers. We employed the 1998-2000 averages of the US BLS data on occupational wages by metropolitan areas (MA) under the categories "management occupations" and "production workers". We then assigned to each city in our sample the wage data of the MA in which the city is located according to the Metropolitan Areas definition of the US Bureau of Census.[11]

Management Occupations is a fairly wide class that includes many categories of managerial jobs, from CEOs to marketing managers, production managers and construction managers. Similarly, Production Workers accounts for a wide definition of workers who deal with production activities. The rationale for using managers and production workers is that

[9] We had to discard two cities because of missing values in some of the other variables.

[10] Our sample includes 66% of all the US metropolitan areas with population higher than 150,000. Among the excluded cities with more than 150,000 inhabitants, the three largest ones are Las Vegas, NE, Honolulu, HI and Long Beach, CA.

[11] See http://www.census.gov/population/estimates/metro-city/99mfips.txt. Since some cities in our sample belong to the same MA, they will have the same dependent variable. Yet, as we shall see below, we also estimated our equations for the 75 MAs to which our 146 cities belong, with no change in results.

the inequality observed by the literature takes the form of a premium for college education, and these two occupational categories are representative of such educational levels. At the same time, unlike other occupations that we could select from the BLS data (e.g. various types of engineers, computer scientists, or several occupations with lower educational levels) they are not industry-specific, nor they are specific to larger or smaller firms. Managers and production workers can be found in any location irrespective of its industrial or economic specialization, and in firms of any size. We could have constructed wage indices of skilled and unskilled workers by combining several BLS occupations. But this meant that we had to make some imputations about which categories fell into one type or the other.

The share of citations to patents in the same city (LOCITE) is our proxy for the degree of technological entrepreneurship in the city. It is worth stressing that in constructing LOCITE we ruled out the self-citations, i.e. the citations to patents owned by the same assignee. The share of firms with more than 1500 employees (SHARE1500), along with the share of trade union members (SMEMB) and Bush votes (BUSH), are our proxies for the Chandlerian-ness of the city. We employ several controls. Apart from the city's income per capita (GDPPOP), which controls for its general wealth, we employed different measures of its technological activity. We used the number of trademarks per firm (TMFIRMS), patents per GDP (PATGDP) and the Milken Pole High-Tech Index (TECHIX) for the city. The latter is the product of the share of high-tech output of the region and the share of high-tech output of the region on the high-tech output of the US (see the Appendix). These controls ensures that LOCITE really captures the importance of the high-tech entrepreneurial model rather than just the local technological activities. To control for inter-industry differences across our cities we used 11 dummies for the sectors with the highest number of trademarks in the city.

4.2 Specification I

Table 3 presents our estimation results of equation (4) and of the relative supply of skills, equation (5). The first column shows the OLS result with no control for the endogeneity of *k* in (4) and *w* in (5). The subsequent columns report the results of separate estimation of (4) and joint-estimation of (4) and (5) by using our instruments for *k* and *w* in the two equations. We employed a maximum-likelihood procedure, but GMM produced similar results.

TABLE 3 ABOUT HERE

The key empirical result of Table 3 is that our measure of skill-biased productivity improvements, LOCITE, has a positive and largely significant impact on inequality in all our instrumental variable (IV) regressions.[12] Our basic measure of Chandlerian-ness, SHARE1500, also has the expected negative sign, and it is statistically significant at $p < 5\%$. Our two other measures of Chandlerian-ness are negative in the two IV equations but they are not well measured. Most importantly, we obtained all our effects after controlling for the technological intensity of the city through patents, trademarks and the technology index. This suggests that the skill-bias does not stem from technological intensity *per sè*, but from a different organization of industrial activities based on technological entrepreneurship, start-ups, and related organizational innovations, as opposed to the classical large firm model.

Our estimated elasticity of substitution is 1.6. This is consistent with earlier findings which obtained elasticities of substitution between skilled and unskilled workers between 1.0-2.0 (Acemoglu, 2002a). As noted in Section 2, factors that raise the productivity of skill-intensive goods have a positive impact on the skill premium if and only if $\sigma > 1$ (gross substitution). Thus, our interpretation of the empirical results is internally consistent. We interpret LOCITE to be a proxy for the new human-capital intensive organization of technology-based industries. If so, given $\sigma > 1$, the impact on inequality ought to be positive, which is what we observe. Likewise, we take SHARE1500 to be correlated with larger firms

[12] Since for some observations LOCITE = 0, we used log(1+LOCITE).

and unskill-bias productivity improvements. Given $\sigma > 1$, it ought to have a negative impact on the skill premium, which is again the observed impact. Compared to the OLS equation, the IV results entail a lower σ and a higher impacts of LOCITE and SHARE1500 (in absolute value). This is the expected bias produced by the endogeneity of *k*. With no control for it we attribute a higher impact of changes in *k* on inequality. Once we control for the endogeneity of *k*, part of the impact on inequality attributed to *k* is produced instead by the underlying model of industrial organization. Finally, our IV regressions estimate that $(\sigma - 2)$ is –0.4. The nul hypothesis $\sigma = 2$ is rejected at $p < 5\%$. As discussed in Section 2, if $\sigma > 2$ the long-run effect of an increase in the relative supply of skilled workers is to increase the skill premium because the skill-biased productivity change induced by the increased relative supply of skills offsets the pressures on the skill premium produced by its increased relative abundance. Thus, from this estimation we cannot conclude that long-run increases in the relative supply of skills raise the skill premium.

Table 3 also shows that our main supply-shiftors, the CL dummies and ENGANC, have the predicted impacts. Of the CL dummies only CL9 is statistically significant at $p < 10\%$ in the IV supply equation. But the point estimates behave as expected. The CL7 dummy is positive and the CL8-9 dummies are progressively larger, which suggests that higher levels of the Child Labor Laws in the 1950s entail a higher relative supply of human capital. The low statistical significance is likely to stem from the fact that these variables move across States rather than cities (which entails a low variance in our sample). With more observations they would probably better measured. The ENGANC instrument is powerful. Of the other supply-shiftors AVTEMP is the most effective.

The estimated elasticity of supply is 1.7. This reflects migration patterns of skilled workers. Apart from US skilled workers, a major phenomenon in recent years has been the movement of Asian computer specialists or engineers to US locations with a growing demand for their services (e.g. Saxienian, 2002). This also means that when assessing the total impact of changes in k on w we also have to take into account that the change in w affects k and then feeds back on w again. It is not difficult to see that if $(\sigma-2)$ is the long-run impact on w of a unit shock on k, this effect gets multiplied by $[1-(\sigma-2)\cdot\chi_w]^{-1}$ via the supply function. From our estimated parameters in the third column of Table 3 the multiplier is equal to $(1.694)^{-1}$. This produces a long-run effect on w of a unit shock on k equal to –0.240, which is significant at $p < 1\%$ using the Wald test. Thus, the feedbacks from the supply side entail a higher long-run effect (– 0.240 > – 0.4) than if we did not take the supply effects into account. This is because the decline in w discourages the local supply of skilled workers, which pushes the skill premium up. We shall see that as we gradually make other variables endogenous in Specification II the effect of a supply shock becomes larger and eventually we obtain that $\sigma > 2$, i.e. supply shocks increase the skill premium in the long-run.

We checked the robustness of our results in three ways. We present these estimations in Appendix 2. First, as discussed in Section 3.2, we estimated (4) and (5) by employing AVSIZE in lieu of SHARE1500 as a proxy for the Chandlerian-ness of the cities. In the same spirit, there is evidence that patent self-citations proxy for larger firms (Hall, Jaffe and Trajtenberg, 2001). We also tried SELF in lieu of SHARE1500. In both cases, the empirical results did not change. Second, we aggregated all our observations for the 75 MA in our sample. As discussed earlier, the ratio of skilled and unskilled wages remained the same, and we averaged out all the other variables for the cities in the same MA. Again the results did not change. Finally, the innovation-driven activities of the entrepreneurial clusters entail more fickle and uncertain returns. By contrast, the larger firms are more reluctant to take up new inventive trajectories in a systematic way, and their comparative advantages often rest on their ability to develop given products, with less uncertain returns, on a larger scale. Thus, our proxies for Chandlerian and entrepreneurial cities ought to have the same impacts on the variability of the skilled wages ("within" inequality). We regressed the inter-quartile range of the skilled wages in our cities (RANGE) on several controls and on SHARE1500 and LOCITE. We found that the latter has a positive and significant impact, while the former has a negative impact or no impact according to the specification.

4.3 Specification II

In this Section we assume that LOCITE is endogenous, and present the results of the joint estimation of (7), (8), and (5). As discussed earlier, we assume that LOCITE is an observation, possibly a noisy one, of A_K. The alternative is to take v_K as endogenous, which amounts to basically the same assumption. What is crucial here is that we are taking the decision to organize technology-based industries like Silicon Valley to be a deliberate process, and we test the determinants and the effects of this choice. For identification, we need to assume that we also "observe" A_L, which we do by assuming that A_L is observed through the index z_L in expression (6). We tried different specifications of the index, e.g. z_L equal just to a parameter times LOG(SHARE1500), with LOG(SMEMB) and LOG(BUSH) entering independently the equations to be estimated. The results did not change. We settled with this specification because we thought that the index better captures A_L.

The results of this estimation are in Table 4. The specification of the estimated parameters follows exactly the way they enter equations (7) and (8). That is, the γ parameters are the impacts of the variables ξ_l in (7) (before normalizing them by σ), and they are the original impacts of the α variables in the inverse relative demand for skill (1). The parameters η are the impacts of the variable v in (2). Thus, they measure the impacts of the variables of interest on the incentives to undertake skill-biased productivity improvements, i.e. on $\left(\frac{A_K}{A_L}\right)$.

Of course, the structure of the model is such that the total impact of any variable ξ_l on z_K combines the demand impacts γ and the technology impacts η, as implied by the parameter

$\frac{\gamma_i + \sigma\eta_i}{\mu_K}$ in (8). In Table 4, the subscripts to the parameters γ and η denote the corresponding variable.

TABLE 4 ABOUT HERE

Table 4 confirms our earlier results. The estimated supply equation is similar to our previous estimations. The estimated μ_K is far higher than μ_L. This is suggestive of a greater impact of A_K than A_L, that is productivity improvements have been highly skill-biased. The weights λ_1 and λ_2 are statistically significant. They are between 0 and 1, and their sum is between 0 and 1. Our estimation then naturally picks reasonable weights.

The variables TECHIX, GDPPOP, and partly TMFIRMS, have positive and significant parameters γ, while the impact of PATGDP is statistically insignificant. The skill-biased impact of demand-side factors (GDPPOP) has some natural interpretation. Economic activities generate a good deal of local demand (backward and forward linkages). The greater technological intensity of a city is then likely to induce a more technologically-intensive local demand, which proves to be skill-biased. Similarly, richer local economies with higher per capita income have a relatively larger demand for more skill-intensive goods, like advanced services vis-à-vis food or manufactured goods. Interestingly enough, TECHIX and TMFIRMS, and above all GDPPOP, have a negative technological impact η, while the impact of PATGDP is still insignificant. One interpretation is that the higher scale of investments in skill-biased productivity improvements coming from the demand-side creates dimishing returns in the incentives to invest in skill-biased productivity improvements from the technology-side. For instance, higher purchase of skill-biased machines because of technological intensity or greater demand for skill-intensive goods creates greater incentives to raise unskill-biased technologies that have become less abundant.

The estimated σ is now 1.871. This is higher than the effect that we estimated earlier, which suggests that by making directed technical change endogenous its impact becomes more important and it almost offsets the classical impact of the relative supply (we now have $(\sigma - 2) = -0.129$). Moreover, we can now decompose the effect of a shift in the relative supply k on the demand for skills w into two effects. From (7), the direct effect of an increase in the relative supply of skills is $-\frac{1}{\sigma} = -0.534$. This is the classical substitution effect produced by an increase in the relative supply. Directed technical change comes through the impact of k on z_K. By using (8) the effect of directed technical change on w is $\frac{(\sigma - 1)^2}{\sigma} = 0.405$. Thus, directed technical change offsets the substitution effect, even though it does not completely eliminate the negative impact of an increase in k. Clearly, the sum of the two effects is 0.129, which is $(\sigma - 2)$. As we did in the previous Section, the full effect of a unit shock in the relative supply k has to take into account that changes in w also affect the relative supply k from (5). Again this amounts to multiplying the demand effect $(\sigma - 2)$ by $[1 - (\sigma - 2)\cdot\chi_w]^{-1}$. Given the estimated χ_w in Table 4 this produces an overall effect of a unit shock to k on the skill premium of –0.104. Table 5 reports this estimated decomposition, along with the statistcal significance of each component. As the Table shows, the overall effect of –0.104, which is smaller in absolute value than in the previous Section, is statistically significant at $p < 10\%$.

TABLE 5 ABOUT HERE

4.4 Endogenous Per Capita Income

Even when we take into account the endogeneity of directed technical change, the ultimate effect of an increase in the relative supply of skills is still potentially at odd with the observed trends in the US, which have exhibited an increase in the college premium over the past two decades, after a decrease during the 1970s. The latter has been interpreted as the short-run consequence of an increase in the relative supply of skills, which initially reduced its relative price. The subsequent increase has instead been taken as evidence of the longer-run effects of investments aimed at improving the relative productivity of the more abundant input. Our empirical analysis is a cross-section across cities around the end of the 1990s, and not a longitudinal analysis over three decades. It could be that the cross-section results are different from the time series, or that the end of the 1990s has stabilized the increase in the skill premium. Yet, to some extent the same forces might be at work across geographical areas and over time. In this respect, if the skill premium has steadily increased in the US over two decades together with increases in the relative supply of skills, we might be able to observe such an increase in the cross-section as well if we could properly account for all the factors that may have moved along the time series.

In this respect, another potential source of endogeneity in our equations is income per capita itself. After all, the rise in the relative supply of skills means more educated labor, and as one observes a higher share of the labor force with college education it is natural to expect that this affects the overall productivity of the system. We then estimated (7), (8), and (5) along with the following equation explaining income per capita,

$$LOG(GDPPOP) = const + \theta_i \xi_{1i} + \theta_{z_L} z_L + \theta_k k \qquad (9)$$

where ξ_1 is the set of variables that also affects (7) and (8), the θs are parameters to be estimated, and z_L is our index in (6) for unskill-biased factors. We basically assume that income per capita is a function of the exogenous variables of the model and of the relative supply k. Moreover, in estimating (7), (8), (9), and (5), we changed slightly our variables in ξ_1 to avoid the endogeneity of GDP still appearing as a denominator of PATGDP. We used as regressors the logs of trademarks (TMARKS), patents (PATENTS), and added FIRMS as a separate regressor.

This specification allows us to estimate the full effect of an increase in the relative supply on the skill premium, taking also into account potential effects coming from GDPPOP. Consider first the effects on the relative demand w. From (7), the impact of k on w is composed of three main effects. First, there is the direct effect $-\frac{1}{\sigma}$. Second, there is the effect coming from z_K, which is in turn composed of two effects, viz. the direct effect of k on z_K and the indirect effect coming through income per capita, as implied by (8). Third, because income per capita directly affects the skill premium as well, one has to take into account that k affects income per capita, and this affects w. With some algebra, from (7) and (8) one obtains

$$\frac{\partial w}{\partial k} = -\frac{1}{\sigma} + \left[\frac{(\sigma-1)^2}{\sigma} + \frac{(\sigma-1)}{\sigma}(\gamma_{GDPPOP} + \sigma \cdot \eta_{GDPPOP}) \cdot \theta_k \right] + \frac{\gamma_{GDPPOP}}{\sigma} \cdot \theta_k \qquad (10)$$

where the three terms of this expression correspond to our three effects. Note that if there was no effect of k on income per capita (i.e. $\theta_k = 0$), we would be back to the previous situation with no GDPPOP effects. In addition, the expression inside the square brackets, i.e. the impact of k on z_K, is in turn composed of various effects. The first term is the direct effect of k on z_K. The second term is composed of two other terms. The γ_{GDPPOP} effect is the one produced by income per capita on skill-intensive goods. Inside the square brackets we have the effect of the demand for skill-intensive goods on the incentives to undertake skill-biased productivity improvements. By contrast, the η_{GDPPOP} effect is the direct effect on z_K. It measures whether income per capita encourages or discourages skill- or unskill-biased productivity changes. The last term in (10) is the direct effect of income per capita on the skill-premium because of its effect on productions characterized by higher or fewer skill-intensity. Finally, as in the previous sections we can take into account the effects on w coming from the fact that w affects k which in turn affects w again. The multiplier of (10) is $[1-w_k \cdot \chi_w]^{-1}$ where w_k is expression (10).

The empirical results in Table 6 confirm the results of Table 4. There is a strong impact of z_K through the estimated μ_K compared to the impact of z_L through μ_L. The weights λ are similar as well. We still estimate σ to be equal to 1.874, which we found not to be significantly different from 2. The similarity of these estimates with their equivalent parameters in Table 4 suggests that there is no new effect here when the income per capita is kept constant.

TABLES 6 AND 7 ABOUT HERE

The novel finding is that we find an effect of demand on the rise of the skill premium. This is consistent with the market size effect theoretically showed by Acemoglu (1998 and 2003), that is the effect of skill-biased technical change is reinforced by the workers using new technology goods that create new demand and increasing returns. To see this, first note that the relative supply of skills k has an important effect on income per capita (θ_k in Table 6). Second, Table 7 presents the estimates of the different components of (10). As noted, the effects unrelated to income per capita (i.e. the first two lines in Table 7) almost cancel out, as implied by our estimated σ being close to 2. The effect of k on z_K coming through income per capita (third line in Table 7) is negative (–0.310).[13] However, the direct effect of income per capita on the skill premium, which is in the fourth line of Table 7 is quite high (0.531). The net effect of income per capita on the skill premium is then 0.531-0.310 = 0.221, which we found to be significantly different from zero at $p < 1\%$ according to the Wald test. Since we now obtained $(\sigma - 2) = -0.126$, then if we did not take into account the impact of the relative supply of skills on income per capita, the long-run effect of an increase in the relative supply of income would produce a slightly negative effect on the skill premium. But the effect of k on the skill premium coming through income per capita, viz. 0.221, now makes the overall

[13] The total effect on an increase in k on z_K is however positive. By summing up all the estimates of the terms inside the square brackets of (10) we obtain 0.408-0.310=0.098, which is higher than zero with $p < 10\%$.

impact of the relative supply of skills positive, i.e. 0.221 – 0.126 = 0.095, as in Table 7. This point estimate is not statistically significant. Yet, aggregate demand factors have now increased the skill premium by 0.224 (i.e. 0.095 + 0.129), where –0.129 is the total effect produced by *k* on *w* in Table 5.[14] Finally, if we take into account the feedback effect on the supply of skills, in Table 7 the total effect of a unit shock increase in *k* is 0.105, which is still insignificantly different from zero, but 2 decimal points higher than –0.104 in Table 5, and statistically different from it at $p < 5\%$ according to the Wald test.

In sum, this discussion reconciles our analysis with the observed trend towards increases in the skill premium in the US as the relative supply of skills increases. The structure of our empirical model suggests that increases in the supply of skills increase income per capita. This encourages a higher relative demand for skill-intensive goods, which explains the increase in the skill premium. By itself directed technological change may not totally offset the reduced skill premium due to the higher supply of skills. Ultimately, the skill premium is explained by a combination of demand, organizational and technology factors.[15]

5. Conclusions

This paper has combined two traditions. First, the literature has documented and explained the growing inequality of earnings between skills in the US, and attributed inequality in good part to skill-biased technical change. Second, other authors have highlighted that the growing skill-intensity of modern productions often takes the form of new models of organizing

[14] We also found that the differences between the two parameters is statistically significant at $p < 5\%$ according to the Wald test.

[15] We also estimated our final full model in Table 6 by allowing for state-dependence following the structure in Acemoglu (2002b), pp.793-5. This amounts to introducing a parameter δ that makes the elasticity of *k* equal to $(\sigma - 2 + \delta)/(1 - \delta \cdot \sigma)$ rather than just $\sigma - 2$. Since δ cannot be separately estimated, we conducted a grid search of the parameter across the space *[0, 1]* (details are available upon request). The basic results do not change and we find that the objective function is maximized for δ close to zero. This suggests that the condition $\sigma > 2$ which we used throughout this paper is close to being necessary and sufficient for the upward effect of a rise in the relative supply of skills on inequality.

industrial activities, based on knowledge spillovers, human capital externalities, and the creation of more entrepreneurial firm settings.

Our empirical analysis has tried to disentagle the factors that contribute to skill-biased productivity improvements, and then to the formation of skill premia. In this respect, an important contribution of this research stems from the collection of a unique and quite articulated set of data for a sample of US cities that host a good deal of industrial and technological activities. Moreover, to our knowledge this is one of the first attempts to examine empirically the importance of "directed technical change" in explaining the inequality of earnings across skills. We combine this approach with the second tradition above. We assume that different organization structures imply different ways of exploiting production factors and we find that variables that account for knowledge spillovers across firms and more generally for externalities in the production of ideas can explain a good part of the skill premia. This compares to the traditional industrial settings based on large integrated firms that ground their production activities on an extensive internal division of labor, economies of scale and scope, and the use of information generated within its organizational boundaries.

Our main result is that the skill premium is explained by directed technical and organizational change, and by demand factors. In our story, technical and organizational changes arise because the larger relative supply of educated people induces them to create models of industrial organization based on technology- and human-capital intensive productions, knowledge spillovers, entrepreneurial firms, start-ups, spin-offs, and the like. At the same time, the increase in income per capita produced by the increase in the skill-level of the workforce spawns a greater demand for skill-intensive goods, which further raises the skill premium. Ultimately, we estimated that the the sum of the effects produced by directed technical or organizational change and by demand offset the reduction in the skill premium due to an increase in the relative supply, and can explain the growing inequality in earnings between skills.

References

Acemoglu, D. (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality", *The Quarterly Journal of Economics*, 113(3): 1055-1089

Acemoglu, D. (2002a) "Technical Change, Inequality, and the Labor Market", *Journal of Economic Literature* **40** (1), 7-72.

Acemoglu, D. (2002b) "Directed Technical Change" *Review of Economic Studies* **69**, 781-809.

Acemoglu, D. (2003) "Patterns of Skill Premia", *Review of Economic Studies* **70** (243), 199-230

Acemoglu, D. and Angrist, J. (2000) "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroeconomics Annual 2000* **15**, 9-59.

Aghion, P., Caroli, E. and Garcia-Peñalosa, C. (1999) "Inequality and Economic Growth: The Perspective of the New Growth Theories", *Journal of Economic Literature* **37** (4), 1615-1660.

Agrawal, A. and Cockburn, I. (2003) "The Anchor Tenant Hypothesis: Exploring the Role of Large, Local, R&D-Intensive Firms in Regional Innovation Systems", *International Journal of Industrial Organization* **21**, 1227-1253.

Alba, R. and Moore, G. (1982) "Ethnicity in the American Elite", *American Sociological Review* **47** (3), 373-383.

Arora, A. and Gambardella, A. (eds.) (2005) *From Underdogs to Tigers: The Rise and Growth of the Software Industry in Brazil, China, India, Ireland, and Israel*, Oxford Universiy Press, Oxford UK.

Audretsch D. and Feldman, M. (1996) "Knowledge Spillovers and the Geography of Innovation and Production", *American Economic Review* **86**, 630-640.

Berman, E., Bound, J., and Machin, S. (1998) "Implications of Skill Biased Technological Change: International Evidence", *Quarterly Journal of Economics* **113** (4), 1245-1279.

Bhidè, A. (2000) *The Origin and Evolution of New Businesses*, Oxford University Press, New York.

Bresnahan, T. (1999) "Computerisation and Wage Dispersion: An Analytical Reinterpretation", *Economic Journal* **109**, F390-F415.

Bresnahan, T. and Gambardella, A. (eds.) (2004) *Building High-Tech Clusters: Silicon Valley and Beyond*, Cambridge University Press, Cambridge UK.

Bresnahan, T., Brynjolfsson, E. and Hitt, L. (2002) "Information Technology, Workplace Organization and the Demand for Skilled Labour: Firm-level Evidence", *Quarterly Journal of Economics* **117** (1), 339-376.

Chandler, A. (1990) *Scale and Scope. The Dynamics of Industrial Capitalism*, The Belknap Press of Harvard University Press, Cambridge MA.

Christensen, C. 1997. *The Innovator's Dilemma*. Harvard Business School Press, Boston MA.

David, P. (1990) “The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox”, *American Economic Review* **80** (2), Papers and Proceedings, 355-361.

Florida, R. (2002) *The Rise of the Creative Class*, Basic Books, New York

Galor, O., and Moav, O. (2000) “Ability Biased Technological Transition, Wage Inequality and Economic Growth”, *Quarterly Journal of Economics* **115**, 469-498.

Goldin, C. and Katz, L. (1998) “The Origins of Technology-Skill Complementarity”, *Quarterly Journal of Economics* **113**, 693-732.

Gottschalk, P. (1997) “Inequality, Income Growth, and Mobility: The Basic Facts”, *Journal of Economic Perspectives* **11** (2), 21-40.

Gottschalk, P. and Smeeding, T. (1997) “Cross-National Comparisons of Earnings and Income Inequality”, *Journal of Economic Literature* **35** (2), 633-681.

Hall, B., Jaffe, A. and Trajtenberg, M. (2001) “The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools”, NBER Working Paper 8498.

Henderson, R., Cockburn, I. (1996). “Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery”. *Rand Journal of Economics*, 27 (1): 32-59

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993) “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations”, *Quarterly Journal of Economics* **10**, 577-598.

Kalmijn, M and Kraaykamp, G. (1996) “Race, Cultural Capital, and Schooling: An Analysis of Trends in the United States”, *Sociology of Education* **69** (1), 22-34.

Langlois, R.N. (2003) “The Vanishing Hand: The Changing Dynamics of Industrial Capitalism”, *Industrial and Corporate Change* **12**, 351-385.

Levy, F., Murnane R.J. (1992), “US Earnings Levels and Earnings Inequality: a Review of Recent Trends and Proposed Explanations”, *Journal of Economic Literature*, 30(3): 1333-1381

Machin, S. and Van Reenen, J. (1998) “Technology and Changes in Skill Structure: Evidence from Seven OECD Countries”, *Quarterly Journal of Economics* **113** (4), 1215-1244.

Moore, G. and Davis, K. (2004) “Learning the Silicon Valley Way”, in Bresnahan, T. and Gambardella, A. (eds.) (2004) *Building High-Tech Clusters: Silicon Valley and Beyond*, Cambridge University Press, Cambridge UK.

New York Times (2005a) “Shadowy Lines That Still Divide”, May 15.

New York Times (2005b) “The College Dropout Boom”, May 24.

Ottaviano, G. and Peri, G. (2004). “The Economic Value of Cultural Diversity: Evidence from US Cities”, NBER Working Paper 10904

Saxenian, A. (1994) *Regional Advantage*, Harvard University Press, Cambridge MA.

Saxenian, A. (2002) *Local and Global Networks of Immigrant Professionals in Silicon Valley*, Public Policy Institute of California, San Francisco.

Topel, R.H., (1997), “Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality”, *The Journal of Economic Perspectives*, 11(2): 55-74

Appendix 1

Data and Their Sources

City variables. Income per capita (GDPOP) and population (POP) for our cities in 2000 were obtained from www.economagic.com and www.epodunk.com. From the same sources we obtained our proxy for the relative supply of skilled workers *K/L*. The 1998-2000 average wages by MA for "management occupations" and "production workers" (INEQ), as well as RANGE, are from the BLS. The share of trade union members by States (SMEMB) are also from the BLS, while the percentage of votes for Bush in the city in the 2000 Presidential election (BUSH) are from CNN website, which we reaggregated for our cities (www.cnn.com/ELECTION/2000/results/president/). We obtained TECHIX, the Milken High-Tech Index, from Richard Florida's research group for their list of cities (www.creativeclass.org). Not all their cities coincided with ours, in which cases we used the index for the nearest city. Information about firms was obtained from the database Icarus of Bureau Van Dijk, which contains the profiles of 1.4 million public and private US companies (www.icarus.bvdep.com). Using Icarus' address of the firm establishments, we collected the total number of firms (FIRMS) in our cities in 2000, the total number of firms with more than 1500 employees (SHARE1500), and the average number of firm employees in the city (AVSIZE).

Trademarks (TMFIRMS). We downloaded from the US Patent & Trademark Office (USPTO) database all the trademarks that were still alive (not abandoned or cancelled) in 2000 whose owner's address was in one of our cities (http://tess.uspto.gov). We assigned the trademarks to the cities in which the companies have their headquarters or legal representation. This may not be the place where the action takes place. However, there are many companies with just one location. Even for the large multi-location companies the problem did not appear to be that severe. We randomly checked a sample of large companies with the Mergent Industrial Manual (www.mergent.com) which provides data on plants, offices and other facilities for more than 2,000 top industrial corporations. In the vast majority of cases, the address in the trademark document corresponded to quite a few establishments and offices in the city.

Patents (PATGDP). We downloaded all the patents granted in 1998-2000 in which the address of the assignee (indicated in the patent document) was one of the cities of our sample. Like with trademarks, the patent may report the address of the headquarter or legal offices of the company, even if the research was carried out elsewhere. Our conservative approach attributed the patent to the city only if at least one of the inventors' addresses was in the same US State of the city. This does not rule out that there may be cities that host inventive activities by multi-location companies whose address as assignees is elsewhere. We inspected our data to check how serious this problem is, and found again that it is not crucial. Note that we could not just select all the patents whose inventors were located in the city. This is because many inventors give their home address in the patent, and this can well be in a nearby city or area. Searching for all the inventors located in nearby cities of all our patents would be quite hard. We were then forced to pre-select our patents by the address of the assignee, and then use the criterion suggested above.

Sector dummies. The front page of the trademark reports a sector code out of 48 product and service categories in which the USPTO classifies them, which we used to single out the sector with the highest number of trademarks in each of our cities. Only 11 of the 48 trademark sectors were leading sectors in one or more of our cities.

Citation data. We matched our city-level patent data with the NBER US Patent Citation Dataset (Hall, Jaffe and Tratjenberg 2001). Our LOCITE variable is the ratio between the citations made by patents in the city to patents by unaffiliated assignees whose address is in the same city, and the total citations of the patents in the city. We employed the same conservative rule that at least one inventor's address had to be in the same city. We also constructed SELF as the ratio between the total number of self-citations (i.e. citations to the same assignee) to the total citations made by the patents in the city.

Instruments. Acemoglu and Angrist (2000) provide the sources of the CL variable. For any given year, CL is the maximum between the years of schooling required for receiving a work permit and the difference between the minimum work age and the maximum enrollment age. The share of the population with English ancenstry (ENGANC) and the weather variables (AVTEMP, CLEARSKY) in the city are from www.city-data.com. We obtained the tolerance index for the US cities (TOLIX) by Richard Florida and his research group. It combines four indices: the Gay Index (ratio of gay families over total families in the city); the Bohemian Index (ratio between artists, dancers, musicians, and other "bohemian" professions, over total employees of the city); the Melting Pot Index (diversification index of population with different origins); and the Integration Index (ratio between foreign born population and total population in the city). The tolerance index is the average value of the four indices. Since the original indices are between 0-1, the tolerance index is also comprised between 0-1. When not available for one of our cities, we used the index of the nearest city in Florida's list.

APPENDIX 2

ROBUSTNESS CHECKS

See Table I-III at the end of the manuscript.

Table 1: Definition of Variables

Variables	Definition
INEQ	Ratio between the 1998-2000 average annual salaries of the BLS classes "managerial occupations" and "production workers" in the Metropolitan Area of the city
K/L	Share of population with a 4-year degree in the city in 2000 over total population without a 4-year degree
Sector Dummies	Dummies = 1 for the trademark sector with the largest number of trademarks in the city. (See Table 2 for the names of the sectors.)
TECHIX	Milken Tech Pole Index = High-tech output of the city over total GDP of the city times the high-tech output of the city over the US high-tech ouput in 1998
TMFIRMS	Number of live trademarks in 2000 in the city on the number of firms (FIRMS) in the city
PATGDP	Number of city patents on city GDP in million $. (Patents granted in 1998-2000 assigned to the city as indicated above; GDP in 2000)
FIRMS	Total number of firms in the city in 2000
GDPPOP	Annual GDP over Population in 2000 (in $)
SMEMB	Share of trade union members on total employees in the State of the city in 2003
BUSH	Percentage of votes to Bush in the city in the 2000 Presidential election
SHARE1500	Share of firms with more than 1500 employees over the total number of firms in the cities in 2000
LOCITE	Share of citations to patents granted to unaffiliated entities located in the city over total citations made by the patents. (1998-2000 patents as above.)
CL	Child Labor Laws in the State in 1959 = Max(years of school required for work permit; min work age – max enrollment age). [CL7-9 = Dummies for CL=7, 8, ≥9]
ENGANC	Share of population of the city with English ancestors
TOLIX	Florida's "tolerance index" (0-1) in the city. Combines four indices: Gay index, melting-pot index, bohemian index, integration index
AVTEMP	Average daily temperature in the city in 2004 (Farenheit)
CLEARSKY	Average monthly days of clear sky in the city in 2004
AVSIZE	Average number of employees of the firms in the city in 2000
SELF	Share of self citations over total citations made by the patents. (Patents granted in 1998-2000 to assignees whose address is in the city and at least one inventor's address is in the city)
RANGE	Difference between the 75 and 25th percentile of the salaries of the occupational class "Managerial Occupations" in the MA in 2000 (in $)

Table 2: Descriptive Statistics

Variable	*Mean*	*Stand.Dev.*	*Min*	*Max*
INEQ	2.768	0.382	1.966	3.810
K/L	0.325	0.123	0.092	0.605
SECTOR DUMMIES				
Chemicals	0.014	0.116	0	1
Cosmetics and Cleanings	0.021	0.142	0	1
Pharmaceuticals	0.021	0.142	0	1
Electrical and Scientific Apparatus	0.582	0.494	0	1
Paper Goods and Printed Matters	0.048	0.214	0	1
Clothing	0.048	0.214	0	1
Toys and Sporting Goods	0.021	0.142	0	1
Staple Foods	0.034	0.182	0	1
Advertising and Business	0.034	0.182	0	1
Insurance and Financial	0.041	0.199	0	1
Computer, Scientific and Legal	0.137	0.345	0	1
TECHIX	0.616	0.143	0.215	0.844
TMPLANT	2.105	12.815	0.008	147.101
PATGDP	$8.454 \cdot 10^{-8}$	$1.779 \cdot 10^{-7}$	$7.012 \cdot 10^{-10}$	$1.261 \cdot 10^{-6}$
FIRMS	10490.263	13719.533	79	106137
GDPPOP	28190.548	10985.275	12438	76668
SMEMB	13.200	5.643	3.100	24.600
BUSH	0.415	0.129	0.090	0.731
SHARE1500	0.028	0.214	0.000	2.582
LOCITE	0.091	0.088	0	0.414
CL	8.452	1.232	6	12
ENGANC	0.088	0.043	0	0.210
TOLIX	0.664	0.113	0.316	0.850
AVTEMP	56.657	7.047	44.992	77.342
CLEARSKY	9.594	3.226	4.417	17.667
AVSIZE	21.865	6.150	11.136	67.956
SELF	0.063	0.063	0	0.304
RANGE	53813.308	7727.316	38720	74850

Table 3: OLS and Instrumental Variable (IV) Estimation of (4) and (5)

Variable	*OLS*	*IV*	*IV*
Dependent Variable: LOG(INEQ)			
Const	-0.159 (0.810)	-3.477 ** (0.042)	-4.148 ** (0.019)
LOG(TECHIX)	0.125 *** (0.007)	0.156 *** (0.001)	0.151 *** (0.002)
LOG(TMFIRMS)	0.017 (0.437)	0.050 ** (0.042)	0.056 ** (0.026)
LOG(PATGDP)	-0.003 (0.725)	0.000 (0.976)	0.001 (0.902)
LOG(GDPPOP)	0.077 (0.173)	0.352 ** (0.014)	0.411 *** (0.006)
LOG(SMEMB)	0.033 (0.295)	-0.031 (0.455)	-0.053 (0.220)
LOG(BUSH)	-0.017 (0.669)	-0.001 (0.892)	-0.017 (0.722)
LOG(SHARE1500)	-0.020 (0.330)	-0.055 ** (0.017)	-0.062 *** (0.009)
LOG(1+LOCITE)	0.324 * (0.080)	0.582 *** (0.005)	0.650 *** (0.002)
σ	1.960 *** (0.000)	1.631 *** (0.000)	1.594 *** (0.000)
Dependent Variable: LOG(K/L)			
Const.	0.369 (0.746)	--	-0.231 (0.856)
LOG(ENGANC)	5.202 *** (0.000)	--	5.973 *** (0.000)
CL7	0.061 (0.656)	--	0.112 (0.434)
CL8	0.228 * (0.089)	--	0.174 (0.214)
CL9	0.285 ** (0.043)	--	0.253 * (0.080)
LOG(TOLIX)	0.277 (0.127)	--	-0.027 (0.915)
LOG(AVTEMP)	-0.725 ** (0.018)	--	-0.865 *** (0.006)
LOG(CLEARSKY)	0.264 ** (0.028)	--	0.109 (0.495)
LOG(BUSH)	0.111 (0.259)	--	0.159 (0.132)
LOG(INEQ) (χ_w)	0.325 (0.189)	--	1.708 ** (0.016)
N. obs.	146	146	146

Hetheroskedastic consistent p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01. In the OLS equations R^2=0.25 and 0.37. In the third column the inequality and supply equations are jointly estimated. The inequality equations include the dummies for the leading sector in the city (Table 2). The instruments are all the variables in the two equations (including the sector dummies), but LOG(INEQ) and LOG(K/L).

Table 4: Maximum Likelihood (ML) Estimation of (7), (8), and (5)

γ, η, σ, μ, λ parameters in (7) and (8)		*Dependent Variable: LOG(K/L)* *(Relative Supply Equation (5))*	
γ_{TECHIX}	0.367 ** (0.026)	*Const.*	-0.750 (0.531)
$\gamma_{TMFIRMS}$	0.099 * (0.093)	χ_{ENGANC}	5.422 *** (0.000)
γ_{PATGDP}	-0.042 (0.181)	χ_{CL7}	0.063 (0.647)
γ_{GDPPOP}	0.875 *** (0.000)	χ_{CL8}	0.165 (0.221)
η_{TECHIX}	-0.297 * (0.087)	χ_{CL9}	0.226 (0.107)
$\eta_{TMFIRMS}$	-0.114 * (0.082)	χ_{TOLIX}	0.012 (0.962)
η_{PATGDP}	0.052 (0.111)	χ_{AVTEMP}	-0.771 *** (0.008)
η_{GDPPOP}	-0.702 *** (0.000)	$\chi_{CLEARSKY}$	0.131 (0.397)
$\lambda_{SHARE1500}$	0.250 ** (0.016)	χ_{BUSH}	0.128 ** (0.176)
λ_{SMEMB}	0.602 *** (0.000)	χ_{w}	1.843 ** (0.010)
σ	1.871 *** (0.000)		
μ_K	6.681 *** (0.000)		
μ_L	0.653 *** (0.001)		

Hetheroskedastic consistent standard errors, p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01. The three equations are estimated jointly. Both (7) and (8) include dummies for the leading sector in the city (Table 2). All the variables in the three equations, but LOG(INEQ), LOG(K/L), and LOG(1+LOCITE), are employed as instruments (including the sector dummies).

Table 5: Estimated Components of the Skill Premium, equations (7), (8), (5)

Parameters	*Effect*	*Estimate*
$-\dfrac{1}{\sigma}$	Direct effect of *k* on *w*	-0.534 *** (0.000)
$\dfrac{(\sigma-1)^2}{\sigma}$	Indirect effect through the direct effect of *k* on z_K, $\dfrac{\partial w}{\partial z_k}\cdot\dfrac{\partial z_k}{\partial k}$	0.405 *** (0.000)
$\dfrac{\partial w}{\partial k}$ (Total demand-side effect)	Sum of the above	-0.129 (0.164)
$[1-(\sigma-2)\cdot\chi_w]^{-1}$	Multiplier produced by the effect of changes in *w* on *k*	1.237 *** (0.000)
$\dfrac{\partial w}{\partial k}\cdot[1-(\sigma-2)\cdot\chi_w]^{-1}$	Total effect on *w* of a unit shock on *k*	-0.104 * (0.095)

Wald tests based on parameter and variance covariance estimates of Table 4; p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01.

Table 6: Maximum Likelihood (ML) Estimation of (7), (8), (5), and (9)

γ, η, σ, μ, λ parameters in (7) and (8)			
γ_{TECHIX}	0.251 ** (0.039)	η_{FIRMS}	0.058 (0.317)
γ_{TMARKS}	0.118 ** (0.036)	η_{GDPPOP}	-1.291 *** (0.001)
$\gamma_{PATENTS}$	-0.070 ** (0.038)	$\lambda_{SHARE1500}$	0.130 ** (0.137)
γ_{FIRMS}	-0.040 (0.457)	λ_{SMEMB}	0.739 *** (0.000)
γ_{GDPPOP}	1.450 *** (0.000)	σ	1.874 *** (0.000)
η_{TECHIX}	-0.178 (0.137)	μ_K	5.047 *** (0.000)
η_{TMARKS}	-0.120 ** (0.035)	μ_L	0.539 *** (0.002
$\eta_{PATENTS}$	0.088 ** (0.012)		
Dependent Variable: LOG(K/L) *(Relative Supply Equation (5))*		*Dependent Variable: LOG(GDPPOP)* *(Income Per Capita Equation (10))*	
Const.	-0.747 (0.497)	*Const.*	10.987 *** (0.000)
χ_{ENGANC}	5.286 *** (0.000)	θ_{TECHIX}	0.047 (0.537)
χ_{CL7}	0.047 (0.689)	θ_{TMARKS}	-0.033 (0.169)
χ_{CL8}	0.181 (0.126)	$\theta_{PATENTS}$	-0.022 (0.201)
χ_{CL9}	0.269 ** (0.025)	θ_{FIRMS}	0.038 (0.183)
χ_{TOLIX}	0.158 (0.509)	θ_k	0.687 *** (0.000)
χ_{AVTEMP}	-0.629 ** (0.021)	θ_{z_L}	0.250 *** (0.005)
$\chi_{CLEARSKY}$	0.262 * (0.079)		
χ_{BUSH}	0.095 (0.289)		
χ_w	1.011 (0.158)		

Hetheroskedastic consistent standard errors, p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01. To avoid endogeneity (because GDP appears as the denominator of PATGDP) we employed as regressors LOG(TMARKS), LOG(PATENTS), and added log(FIRMS) as a separate regressor. The four equations are estimated jointly. Equations (7), (8), and (9) include dummies for the leading sector in the city (Table 2). All the variables in the four equations, but LOG(INEQ), LOG(K/L), LOG(1+LOCITE), and LOG(GDPPOP) are employed as instruments (including the sector dummies).

Table 7: Estimated Components of the Skill Premium, equations (7), (8), (5), (9)

Parameters	*Effect*	*Estimate*
$-\dfrac{1}{\sigma}$	Direct effect of *k* on *w*	-0.534 *** (0.000)
$\dfrac{(\sigma-1)^2}{\sigma}$	Indirect effect through the direct effect of *k* on z_K, $\dfrac{\partial w}{\partial z_k}\cdot\dfrac{\partial z_k}{\partial k}$	0.408 *** (0.000)
$\dfrac{(\sigma-1)}{\sigma}\left(\gamma_{GDPPOP}+\sigma\cdot\eta_{GDPPOP}\right)\cdot\theta_k$	Indirect effect through the effect of *GDPPOP* on z_K, $\dfrac{\partial w}{\partial z_k}\cdot\dfrac{\partial z_k}{\partial GDPPOP}\cdot\dfrac{\partial GDPPOP}{\partial k}$	-0.310 *** (0.000)
$\dfrac{\gamma_{GDPPOP}}{\sigma}\cdot\theta_k$	Indirect effect through the direct effect of *GDPPOP* on *w*, $\dfrac{\partial w}{\partial GDPPOP}\cdot\dfrac{\partial GDPPOP}{\partial k}$	0.531 *** (0.000)
$\dfrac{\partial w}{\partial k}$ (Total demand-side effect)	Sum of all the above	0.095 (0.122)
$[1-\dfrac{\partial w}{\partial k}\cdot\chi_w]^{-1}$	Multiplier produced by the effect of changes in *w* on *k*	0.904 *** (0.000)
$\dfrac{\partial w}{\partial k}\cdot[1-\dfrac{\partial w}{\partial k}\cdot\chi_w]^{-1}$	Total effect on *w* of a unit shock on *k*	0.105 (0.162)

Wald tests based on parameter and variance covariance estimates of Table 6; p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01.

Table I: IV Estimation of (4) and (5), Alternative Specifications

Variable	*IV*	*IV*
Dependent Variable: LOG(INEQ)		
Const	-2.766 *	-3.249 **
	(0.077)	(0.037)
LOG(TECHIX)	0.167 ***	0.146 ***
	(0.001)	(0.003)
LOG(TMFIRMS)	-0.006	-0.007
	(0.523)	(0.402)
LOG(PATGDP)	0.000	-0.002
	(0.982)	(0.856)
LOG(GDPPOP)	0.337 **	0.355 ***
	(0.012)	(0.008)
LOG(SMEMB)	-0.019	-0.029
	(0.603)	(0.435)
LOG(BUSH)	-0.023	-0.043
	(0.626)	(0.354)
LOG(AVSIZE)	-0.093	--
	(0.127)	
LOG(1+SELF)	--	-0.087
		(0.757)
LOG(1+LOCITE)	0.566 ***	0.600 ***
	(0.005)	(0.005)
σ	1.674 ***	1.681 ***
	(0.000)	(0.000)
Dependent Variable: LOG(K/L)		
Const.	-0.019	-0.254
	(0.987)	(0.839)
LOG(ENGANC)	5.741 ***	5.869 ***
	(0.000)	(0.000)
CL7	0.106	0.086
	(0.436)	(0.547)
CL8	0.187	0.173
	(0.165)	(0.219)
CL9	0.258 *	0.250 *
	(0.062)	(0.084)
LOG(TOLIX)	0.059	0.001
	(0.794)	(0.976)
LOG(AVTEMP)	-0.823 ***	-0.841 ***
	(0.006)	(0.007)
LOG(CLEARSKY)	0.150	0.130
	(0.304)	(0.408)
LOG(BUSH)	0.143	0.156
	(0.144)	(0.132)
LOG(INEQ) (χ_w)	1.273 **	1.616 **
	(0.030)	(0.019)
N. obs.	146	146

Hetheroskedastic consistent standard errors, p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01. The two equations are estimated jointly. The inequality equations include dummies for the leading sector in the city (Table 2). Same instruments as in Table 3.

Table II: IV Estimation of (4) and (5), 75 Metropolitan Areas

Dependent Variable: LOG(INEQ)		
Const	-5.632 * (0.051)	-5.614 * (0.052)
LOG(TECHIX)	0.115 ** (0.046)	0.114 ** (0.046)
LOG(TMFIRMS)	0.074 ** (0.047)	0.073 ** (0.047)
LOG(PATGDP)	-0.021 (0.140)	-0.021 (0.140)
LOG(GDPPOP)	0.497 ** (0.040)	0.495 ** (0.041)
LOG(SMEMB)	-0.099 (0.193)	-0.099 (0.194)
LOG(BUSH)	-0.059 (0.283)	-0.059 (0.284)
LOG(SHARE15000)	-0.096 ** (0.014)	-0.096 ** (0.014)
LOG(1+LOCITE)	0.882** (0.010)	0.881 ** (0.010)
σ	1.529 *** (0.000)	1.530 *** (0.000)
Dependent Variable: LOG(K/L)		
Const.	--	-0.903 (0.579)
LOG(ENGANC)	--	9.368 *** (0.000)
CL7	--	0.085 (0.619)
CL8	--	0.187 (0.263)
CL9	--	0.214 (0.216)
LOG(TOLIX)	--	-0.178 (0.457)
LOG(AVTEMP)	--	-0.713 (0.125)
LOG(CLEARSKY)	--	0.041 (0.768)
LOG(BUSH)	--	-0.024 (0.843)
LOG(INEQ) (χ_w)	--	1.420 ** (0.055)
N. obs.	75	75

Hetheroskedastic consistent standard errors, p-values in parenthesis: *p<0.1; **p< 0.05; ***p<0.01. Cities in the same Metropolitan Areas (from Census, see text) were aggregated, and all the other variables were taken at their averages across them. In the second column the two equations were estimated jointly. The inequality equation includes dummies for the leading sector in the city (Table 2). Same instruments as in Table 3 (averaged for cities in the same MA).

Table III: "Within" Inequality Estimation

Variable	*OLS*	*OLS*	*OLS*	*IV*
Dependent Variable: LOG(RANGE)				
Const	9.757 ***	9.105 ***	9.223 ***	5.564 ***
	(0.000)	(0.000)	(0.000)	(0.001)
LOG(TECHIX)	0.074 *	0.080 **	-0.040	0.114 **
	(0.072)	(0.043)	(0.331)	(0.010)
LOG(TMFIRMS)	-0.009	-0.003	-0.001	0.032
	(0.669)	(0.895)	(0.935)	(0.200)
LOG(PATGDP)	0.006	0.007	0.014	0.011
	(0.506)	(0.457)	(0.108)	(0.257)
LOG(GDPPOP)	0.099 ***	0.152 ***	0.134 ***	0.446 ***
	(0.001)	(0.007)	(0.000)	(0.002)
LOG(SMEMB)	0.004	-0.009	-0.006	-0.077 *
	(0.849)	(0.696)	(0.830)	(0.058)
LOG(BUSH)	-0.121 ***	-0.119 ***	-0.072 **	-0.108 **
	(0.000)	(0.001)	(0.016)	(0.011)
LOG(SHARE1500)	0.008	0.001	0.004	-0.037 *
	(0.688)	(0.967)	(0.781)	(0.099)
LOG(1+LOCCIT)	0.544 ***	0.594 ***	0.293 *	0.870 ***
	(0.004)	(0.001)	(0.064)	(0.000)
LOG(K/L)	--	-0.064	--	-0.415 ***
		(0.196)		(0.007)
LOG(ENGANC)	--	--	-1.118 ***	--
			(0.003)	
CL7	--	--	0.048	--
			(0.433)	
CL8	--	--	0.034	--
			(0.538)	
CL9	--	--	0.057	--
			(0.328)	
LOG(TOLIX)	--	--	0.224 ***	--
			(0.000)	
LOG(AVTEMP)	--	--	0.065	--
			(0.532)	
LOG(CLEARSKY)	--	--	0.129 ***	--
			(0.002)	
N. obs.	146	146	146	146

Hetheroskedastic consistent standard errors. p-values in parenthesis. *p<0.1; **p< 0.05; ***p<0.01. The equations also include dummies for the leading sector in the city (Table 2). The R^2 of the OLS equations are 0.39, 0.40, and 0.54.The IV equation uses all the variables in the third OLS equation (including the sector dummies) as instruments for LOG(K/L).