# COMMUNICATION EXTERNALITIES IN CITIES

First Draft

February 2003

Sylvie CHARLOT
INRA Dijon
26, Bd Docteur Petitjean
BP 87999
21079 Dijon Cedex, France
charlot@enesad.inra.fr

Gilles DURANTON
London School of economics and CEPR
Houghton Street
London WC2A 2AE, UK
g.duranton@lse.ac.uk
http://cep.lse.ac.uk/~duranton/

**Abstract.** To identify communication externalities in French cities, we exploit a unique survey recording workplace communication of individual workers. Our identifying assumption is that in larger and/or more educated cities, workers should communicate more. In turn, more communication should have a positive effect on individual wages. By estimating both an earnings and a communication equation, we find evidence of communication externalities. Being in a larger and more educated city makes workers communicate more and in turn this has a positive effects on wages. However, only a small part of the overall effects of a more educated and larger city on wages percolates through this channel. We also find evidence of spatial sorting of workers: communication externalities affect 'stayers' much more than 'movers'.

Keywords: human capital, cities, communication externalities.

JEL numbers: J31, R19, R29

We thank Nathalie Greenan for providing us with the COI data and Virginie Piguet for kindly matching it with location and human capital data. We are also grateful to Cécile Détang-Dessendre, Laurent Gobillon, David Rigby, Bertrand Schmitt, participants at the 2002 NARSA meeting and the 2003 CEPR geography workshop in Villars for comments and suggestions. Financial support from INRA-ENESAD is gratefully acknowledged.

## 1. Introduction

The strength of human capital externalities is a fundamental determinant of the optimal subsidy to education. Furthermore, as argued by Lucas (1988) and his followers, human capital externalities could constitute a fundamental engine of growth and development. Local human capital externalities are also accepted as one of the main reasons to justify the existence of cities since Marshall (1890). Hence, obtaining reliable estimates for the strength of human capital externalities is widely acknowledged to be of fundamental importance.

It is also *essential to know how these externalities percolate*. There are three main reasons for this. First, the literature typically infers the existence of human capital externalities indirectly by estimating a wedge between private and social returns to education.[1] It speaks of human capital externality when an aggregate measure of human capital has a positive effect on individual earnings over and above that of individual characteristics. Such findings however might be driven by some missing variables and not by human capital externalities. Second, getting the optimal education subsidy right is one thing but there could be other corrective policies. If, as suggested by Marshall (1890), human capital externalities take place mostly between workers in the same city and industry city through face-to-face interactions, regular meetings within local professional associations may be a good way to improve economic efficiency. Third, without knowing how these externalities percolate, nothing prevents the theorists from assuming whatever they like. Theoretical progress is thus hampered by our lack of knowledge about the precise nature of human capital effects.[2] In short, knowing how human capital externalities percolate is of considerable importance for both theory and policy.

When elaborating on "human capital externalities", the literature almost inevitably alludes to some form of technological externalities and mentions face-to-face meetings, word-of-mouth communication, direct interactions between skilled workers, and the like. This quasi-exclusive focus on a particular sub-set of human capital externalities, which we call *communication externalities*, may not be warranted. Human capital could have some external

[1] See for instance Rauch (1993), Acemoglu and Angrist (2000), Adserà (2000), Moretti (2000), or Simon and Nardinelli (2002). This literature is discussed more in-depth below.

[2] There is a long tradition in urban economics that justifies the existence of cities by an appeal to technological externalities involving direct interactions between agents (see Duranton and Puga, 2003, for a review). However, this literature offers few micro-economic foundations for these externalities. This difficult theoretical undertaking is made all the more difficult by the absence of evidence about how these externalities percolate, whom they concern, and at which spatial scale they take place.

effects through a variety of other channels. More human capital in a city could foster the supply of intermediate goods and in turn improve the productivity of final producers – a pecuniary externality unrelated to communication externalities. More human capital could also lead to better matches between employers and employees. One could also invoke a more extensive division of labour within a more educated workforce, etc.

In this paper, we propose a novel attempt to identify communication externalities and distinguish them from the other external effects of human capital. To do this, we exploit a unique survey recording workplace communication practices for around 6,000 French workers in 1997. Because of its careful design and implementation, we believe this survey contains very valuable information about workplace communication.[3] Our identifying assumption is that larger and/or more educated cities should favour communication as postulated for instance by Beckmann (1976), Borukhov and Hochman (1977), Fujita and Ogawa (1982), Black and Henderson (1999), Glaeser (1999), Berliant, Reed and Wang (2001) or Lucas (2001) among others. Then it should also be the case that more communication should have a positive impact on individual earnings (Jovanovic and Rob, 1989, and the references above). The strength of communication externalities can then be computed as the effects of cities size and average urban schooling on communication times that of communication on earnings.[4]

Consistent with our two hypotheses, we find that workplace communication is positively associated with earnings. Furthermore, city size and average urban schooling are positively associated with workplace communication. However, in total only about a tenth of the effects of city size and average urban schooling appear to percolate through communication externalities. This suggests that the bulk of the effects of city size and average urban schooling on earnings, which are sizeable and highly significant, must percolate through channels other than communication externalities.

With respect to the strength of communication externalities, these findings may be biased downwards because we can only observe the amount of communication and not its quality. If

[3] Communication externalities is one of those topics like social capital for which finding 'market data' that leave reliable paper trails appears hopeless. For the time being, progress on this type of topics may have to rely on survey data and experiments.

larger and more educated cities allow for better rather than more communication, we may underestimate the strength of communication externalities. However, these unobservable quality effects would need to be very large to justify communication externalities being of primary importance.[5] On the other hand, our findings about communication externalities may be biased upwards because of workers' unobserved heterogeneity. If high-wage cities attract good communicators, we may overestimate the strength of communication externalities. To investigate this issue, we use information about the worker's birthplace to distinguish between 'movers' and 'stayers'. We find some evidence of spatial sorting but the effects are subtle. Movers' workplace communication is only weakly influenced by their environment. Wherever they are, movers communicate more than stayers and their returns to communication are higher than those of stayers. At the same time, movers also take greater advantage of the other benefits of city size and average schooling.

As a further concern, workplace communication could be suspected of being determined simultaneously with earnings. For instance, when a worker is promoted to a higher position, this is likely to involve both a higher wage and more workplace communication. To investigate this, one would like to have variables that determine workplace communication but remain un-correlated with the residual in our earnings equation. Fortunately our survey data contains information about the working environment and working conditions. Some of the variables related to these issues provide good instruments for workplace communication. Using these instruments leaves our results mostly unchanged.

Two final limitations must be acknowledged. Firstly, we only measure static communication externalities and not the kind of long-run learning benefits provided by the cities highlighted by Glaeser and Maré (2001). Secondly, we measure communication externalities only at the city level and not at the country level.[6] It is however natural to expect communication

[4] Because we are able to observe not only human capital inputs and wages but also an important channel through which externalities should percolate, we call our approach *direct* as opposed to the *indirect* approaches that prevail in the literature.

[5] It is not fully clear how large cities could offer massively better communication without workers communicating more to take advantage of these quality effects. In other words, large quality effects should be captured by equilibrium volumes of communication.

[6] See Cohen and Soto (2001) for a recent cross-country attempt to measure social returns to education and a discussion of the macro literature. In this literature, the pendulum is fast moving between two extremes that view the role of human capital as either fundamental or negligible. Cohen and Soto (2001) argue that much of the disagreement finds its source in bad data. In their work, they use better data and claim that social and private returns to education are essentially equal. However, in their regression, they use the rate of urbanisation as control to proxy for total factor productivity. This variable may however also capture human capital externalities taking place in cities.

externalities to be strongly localised. Bearing these caveats in mind, our general conclusion is that communication externalities are present in cities but that only a small fraction of the external effects of human capital percolate through communication.

The strategy of the rest of the paper is the following. The next section reviews the relevant empirical literature. Section 3 presents a theoretical model to illustrate the identification issues discussed above. Section 4 provides some estimation results for our model. Section 5 performs a series of robustness test. Section 6 draws some conclusions.

## 2. Related empirical literature

Our paper is related to three main strands of literature. First, using Roback (1982)'s equilibrium location approach, Rauch (1993) estimates hedonic earnings equations where individual earnings are regressed on a set of individual variables together with city level variables. Despite numerous controls for both individual and city characteristics, he finds strong effects of average schooling on individual earnings within US cities. This finding has been replicated many times (e.g., Adserà, 2000, Simon and Nardinelli, 2002, etc). According to this type of estimation, external returns to education in cities could be very large, between 50 and 100% of the private returns.

Rauch's seminal approach has been criticised on several grounds. Cross-section estimations make it difficult to distinguish human capital externalities from the effects of unobserved city heterogeneity whereby high-wage cities might attract high-education workers. A second concern regards individual unobserved heterogeneity. If workers with good unobserved characteristics sort themselves in high-education cities, the estimates for external returns to human capital obtained in a Rauch-style regression are again biased upwards.[7]

To deal with these issues, instrumental variables have been considered. A good instrument for average schooling would affect the schooling of the majority of workers in a given location without being correlated with local wages. Acemoglu and Angrist (2000) argue that differences in school compulsory attendance laws and child labour laws in US states over the 20$^{th}$ century provide such variation, at least for secondary education. In their preferred estimation, they obtain very small external returns to education. In the same spirit but using different instruments for schooling, Moretti (2000) shows that the share of college graduates in a city has a strong effect on individual earnings. These results are confirmed when using the longitudinal dimension of a panel of workers.

A more fundamental critique to Rauch's approach has been recently put forward by Ciccone and Peri (2002). It builds on a well-known fact: workers are imperfect substitutes in production (see Topel, 1997, for a recent overview). When workers are paid at their social marginal product, the unskilled wage is typically expected to be higher in cities where the proportion of skilled workers is larger. In other words, without accounting for imperfect substitutability between different types of workers, human capital externalities may be confused with complementarities in production.[8] Ciccone and Peri (2002) develop a novel approach to assess the effects of an increase in human capital in a city keeping the skill composition of the workforce constant. Applying this 'constant composition' methodology to US cities, they find small and insignificant human capital externalities.

We differ from this literature in our use of workplace communication data to directly identify communication externalities, a sub-set of human capital externalities. Unlike the aforementioned papers, our primary interest is not a precise measure of total external returns to education. We instead focus on communication, that is one particular channel for human capital externalities that figures prominently in our thinking about cities. In this regard, our approach is related to the small literature attempting to identify the sources of urban increasing returns.

Only few papers attempt to disentangle empirically between different micro-foundations of urban increasing returns.[9] Exemplary in this literature, Holmes (2002) uses the differences in the location patterns of sales offices of small vs. large firms. This allows him to separate the effects of local market size from those of cost-reducing externalities and comparative advantage. Building on Jaffee, Trajtenberg and Henderson (1993), Almeida and Kogut (1999) show that the citation trail for patents coincides with the movement of key scientific personnel. This suggests that "spillovers" may be channelled through the labour market rather

[7] In the same vein, it could be the case that cities with higher wages provide more schooling.
[8] However many such complementarities may be nothing else but pecuniary externalities from human capital. The fundamental issue is whether the urban production function exhibits constant or increasing returns to scale.
[9] See Rosenthal and Strange (2003) for a more complete survey of this literature.

than word-of-mouth communication between scientists. In a different vein, Dumais, Ellison and Glaeser (1997) use carefully constructed proxies to distinguish between the three Marshallian motives for agglomeration. They also find support for thick local labour market effects.

Finally, although the terminology may differ, communication externalities enjoy widespread popularity and attention among other social scientists. Saxenian (1994) is a good example of this type of work. She forcefully argues that the root of Silicon Valley's success are to be found in a unique culture that favours frequent and open face-to-face contacts, which in turn lead ideas to flow freely across workers and firms. This literature is discussed at length in Storper (1997) who reviews a large body of work offering suggestive evidence about communication externalities. Unfortunately this literature relies mostly on qualitative evidence and communication externalities are never formally tested against other channels.[10]

## 3. Communication externalities vs. other human capital effects in cities

The main objective of the model that follows is to show that *not all human capital externalities can be interpreted as communication externalities.* Instead more human capital in a city can a priori generate both stronger communication externalities and stronger other human capital externalities unrelated to communication which we model as an input sharing mechanism.[11] This model also shows how communication externalities can be empirically distinguished from these other human capital externalities.

### 3.1. Urban scale effects unrelated to communication externalities

The first part of our model follows Ethier's (1982) production-side version of Dixit and

[10] Of particular interest in this body of work, Goddard (1973) and Goddard and Morris (1976) compile very detailed communication data about a large group of workers in London. They show a strong link between the intensity of communication and central locations. They also document a wealth of interesting features about workplace communication. Unfortunately they do not explore the links between communication and productive efficiency.

[11] We build on a specific model wherein input sharing between final producers implies increasing returns at the city level. Similar results can be obtained with any alternative source of local increasing returns that does *not* rely on communication externalities. See Duranton and Puga (2003) for a survey of the different microeconomic foundations of urban increasing returns. Our preference for input-sharing as opposed to say matching is that the former mechanism naturally benefits all workers symmetrically whereas it is more difficult to conceive how a large fraction of skilled workers in a city could help the matching of unskilled workers to jobs.

Stiglitz's (1977) model of monopolistic competition. This model was first embedded in an urban framework by Abdel-Rahman (1988), Fujita (1988) and Rivera-Batiz (1988).

Final goods producers use intermediate goods produced by differentiated suppliers to produce under constant returns to scale a homogenous consumption good. This final good, which also serves as numéraire, can be traded across cities at no cost. By contrast, intermediates cannot be traded across cities so that final good producers can only buy from intermediate producers located in the same city. Final producer *k* in city *i* produces according to:

$$y_k = \left[ \int_{z \in i} q_k(z)^{(\sigma-1)/\sigma} \mathrm{d}z \right]^{\sigma/(\sigma-1)}, \ \sigma > 1 \tag{1}$$

where $q_k(z)$ is the quantity of intermediate *z* bought by *k*, $\sigma(>1)$ is the elasticity of substitution across intermediates, and the notation $z \in i$ denotes any (intermediate producer) *z* located in city *i*. After denoting by $p(z)$ the price of intermediate *z*, final producer *k*'s profit is given by:

$$\pi_k = y_k - \int_{z \in i} p(z) q_k(z) \mathrm{d}z . \tag{2}$$

As in Ethier (1982), intermediate goods are produced by monopolistically competitive firms. To produce any variety of intermediates, there is a fixed labour overhead to start production and a constant quantity of labour for each marginal unit. Employment in firm *z*, expressed in *effective* units of labour, is thus:

$$l(z) = \beta q(z) + \alpha \ . \tag{3}$$

After denoting by $w_i$ the wage rate in city *i*, the profit function of intermediate producer *z* in city *i* is:

$$\pi(z) = p(z) q(z) - w_i [\beta q(z) + \alpha] . \tag{4}$$

To solve the model, note first that the maximisation of (2) by final producer *k* implies:

$$p(z)=\frac{q_k(z)^{-1/\sigma}y_k}{\int\limits_{z\in i} q_k(z)\mathrm{d}z}. \tag{5}$$

Since, intermediates cannot be traded across cities, summing over all final producers in the city yields the inverse-demand faced by intermediate producer *z*. It can be inserted into the profit of *z* given by (4). Profit maximization by intermediate producers then implies that the price of intermediates, $p(z)=\sigma\beta w_i/(\sigma-1)$, is a mark-up over marginal cost and is independent of total market size. Then under free entry in the production of intermediates, the output of any intermediate producer is independent of market size: $q(z)=\alpha(\sigma-1)/\beta$. After denoting by $L_i$ total effective labour supply in city *i*, total output is then given by:

$$Y_i=\frac{(\sigma-1)}{\alpha^{1/(\sigma-1)}\beta\sigma^{\sigma/(\sigma-1)}}L_i^{\sigma/(\sigma-1)}\equiv\Phi L_i^{\sigma/(\sigma-1)}. \tag{6}$$

Finally clearing on the labour market and free entry for final producers imply that final output is fully dissipated in the wage bill. This yields the following wage rate:

$$w_i=\Phi L_i^{1/(\sigma-1)}. \tag{7}$$

This wage increases with city size. A larger workforce in a city leads to a wider range of intermediates being produced for final good production. Since these intermediates enter the production function of final good producers with the same constant elasticity of substitution $\sigma$, a wider range of intermediates results in final output rising more than proportionately. Hence, despite constant returns to scale at the firm level in final production, there are aggregate increasing returns working through this pecuniary externality. The strength of these aggregate increasing returns decreases with $\sigma$, the elasticity of substitution between intermediates. To summarise, in this model, which is arguably the canonical model of agglomeration economies, local increasing returns arise from sharing a greater variety of intermediates by final producers as the local market expands.

### 3.2. Human capital in cities

The effective labour supply of worker *j*, $l_j$, is a function of her human capital, $h_j$. Then, aggregate effective labour supply in city *i*, $L_i$ is the sum of the effective labour supply of all workers living in the city. Specifically:

$$L_i = \sum_{j \in i} l_j \text{ , where } \quad l_j = e^{h_j} . \tag{8}$$

In turn, the human capital of worker *j* living in city *i*, $h_j$, is a function of her measured skills (such as those accruing from education and labour market experience), her communication and some unobserved productivity shock:

$$h_j = s_j + x_j + \varepsilon_j , \tag{9}$$

where $s_j$ is the worker's measured skills, $x_j$ her volume of communication, and $\varepsilon_j$ an idiosyncratic shock which is assumed to be normal and iid.[12] The basic justification for this type of assumption is that a worker's productivity depends not only on her observed skills (education, experience, etc) but also on the informal knowledge and information she receives when communicating and some unobserved random component (such as a temporary productivity shock or some unobserved abilities).[13]

The labour market earnings of worker *j* living in city *i*, $W_j$ are equal to $w_i l_j$. After inserting (8) and (9) into (7), we obtain:

$$W_j = \Phi \left( \sum_{g \in i} e^{s_g + x_g + \varepsilon_g} \right)^{1/(\sigma - 1)} e^{s_j + x_j + \varepsilon_j} . \tag{10}$$

[12] Note that skills and communication are perfect substitutes in the production of human capital but they complement each other when it comes to effective labour supply. Using (8) and (9), it is easy to verify that: $\partial^2 l_j / \partial s_j \partial x_j > 0$.

[13] Observation and imitation may also provide a channel for increasing human capital. For most skilled occupations, we expect this channel in its pure form to be of trivial importance. Instead we expect observation and imitation to be combined with communication.

When $s$ and $x$ are small, a Taylor expansion implies

$$\log W_j \approx \log\Phi + \frac{1}{\sigma - 1}\left(\log N_i + \bar{s}_i + \bar{x}_i + \bar{\varepsilon}_i\right) + s_j + x_j + \varepsilon_j\,, \tag{11}$$

where $\bar{s}_i + \bar{x}_i$ is the average human capital in city *i*, $N_i$ its population, and by the law of large numbers: $\bar{\varepsilon}_i = 0$. According to (11), individual earnings increase with the individual effective labour supply. In turn, the latter is determined by individual characteristics (i.e., skills, $s_j$, communication, $x_j$, and the unobserved random component, $\varepsilon_j$). At the same time, earnings are also determined by city aggregates. Because of input sharing, earnings are higher in cities where aggregate effective labour supply is higher, that is where workers are more numerous, more skilled, and communicate more (i.e., higher $N_i$, $\bar{s}_i$, and $\bar{x}_i$ respectively).[14]

### 3.3. Communication externalities in cities

The theoretical literature on externalities in cities, as summarised in Duranton and Puga (2003), proposes a variety of mechanisms to explain why cities may foster communication externalities. Glaeser (1999) and Berliant, Reed and Wang (2001) explicitly assume that larger cities offer more opportunities for face-to-face meetings (a population size effect). Furthermore Jovanovic and Rob (1989) and many others suggest that during face-to-face interactions, skilled workers learn more from other skilled workers than from unskilled workers. Hence holding the number of meetings fixed, a more skilled workforce provides better learning opportunities (an average human capital effect). Combining the two effects, larger and more educated cities are widely assumed to increase workplace communication. This in turn implies higher wages.

However, we cannot expect workplace communication to be influenced only by city level

[14] Previewing some estimation issues, note that we assume that cities differ only by their average human capital and their population size. The effects of exogenous productivity differences between cities (i.e., differences in $\Phi$) and spatial sorting of workers with respect to their unobserved characteristics (i.e., $\bar{\varepsilon}_i \neq 0$) are discussed in Section 5. We also bear in mind that in small cities, the number of observations may be small so that the approximation $\bar{\varepsilon}_i = 0$ given by the law of large numbers is likely to be invalid and thus generate some heteroscedasticity in our results.

variables. First, some workers may be better at communicating because of their better measured skills (viz. reading writing, etc). Furthermore unobserved social skills may foster one's abilities to extract information during face-to-face communication, to put forward ideas in small-group meetings or to generate loyalty from colleagues, subordinates, employers, etc. The simplest specification is to assume that personal workplace communication of worker *j* in city *i* is given by:

$$x_j = A + B.\log N_i + C.\bar{s}_i + D.s_j + \mu_j . \quad (12)$$

where *A* is a constant, *B* is a coefficient measuring the elasticity of individual workplace communication with respect to city size, *C* measures the effect of average skills in city *i*, *D* is the effect of individual (measured) skills and $\mu_j$ captures individual unobserved social skills, which are assumed to be iid and normally distributed.[15] Workers with better skills (measured, $s_j$, or social and unobserved, $\mu_j$) are expected to communicate more. Workers located in larger and/or more educated cities are also expected to communicate more because of more people to meet and higher quality meetings.[16] In this sense, we speak of communication externalities.

Inserting (12) into (11), we obtain,

$$\log W_j \approx \text{constant} + \frac{1+\sigma B}{\sigma - 1}\log N_i + \frac{1+\sigma C}{\sigma - 1}\bar{s}_i + (1+D)s_j + \varepsilon_j + \mu_j . \quad (13)$$

This reduced form is such that the earnings of worker *j* increase when she has better measured skills, $s_j$, when she has better social skills $\mu_j$ (and better unobserved productive skills, $\varepsilon_j$), when she lives in a larger city and when she lives in a city with more skilled workers. The last two effects percolate through two different channels (partly nested into each other). First, a larger and/or more educated city implies stronger communication externalities, which in turn

[15] We model the effects of the environment on workplace communication as a pure externality. More subtle (and realistic) mechanisms may be considered. In particular, workers may play a less passive role and decide how much time they want to communicate depending on where they are. With the data we have, unfortunately we cannot directly explore these issues.

[16] Again, previewing some estimation issues, note that we still consider that cities only differ by their skill composition and their population size. Our specification also assumes that the unobserved social skills are iid across cities. These two assumptions are discussed further in Section 5.

lead to more efficient workers and thus higher earnings. Second, a larger and/or more educated city also implies stronger externalities to unrelated communication like those caused by input sharing. They also lead to higher earnings.

Estimating directly (13), which is similar to what us estimated by Rauch (1993) and his followers, would not allow us to identify the effects of communication externalities. To identify such effects and distinguish them from the other external effects of human capital, our empirical strategy is to estimate (11) and (12) instead of estimating only the reduced form (13). To do this, data about workplace communication is obviously needed.

## 4. Data and results

### 4.1. Data

In what follows, we exploit data from a detailed survey conducted in 1997 in France. 8812 workers were randomly drawn from the labour force employed in manufacturing, services (accounting only) and retail (DIY chains only). Selected workers were individually interviewed and the information was matched with corresponding firm level data, other individual data, and location data. For each worker who responded to the survey and whose data could be matched with other data, we know their responses to around 80 questions covering a wide range of topics: working conditions, organisational change, workplace communication and information technologies, workplace location (rural, suburban, or urban with the city population), earnings, industry and employment in the relevant establishment.

This data, further described in Appendix, is critically analysed in Greenan and Hamon-Cholet (2001a, b) and Greenan and Mairesse (1999) who offer various checks regarding its quality. A key part of the questionnaire regards 11 questions related to the workplace communication of the surveyed workers. A first set of questions is about communication within the firm. A second set of questions regards communication external to the firm. Finally the last set of questions is concerned with the media used by the surveyed workers. In this paper, we aggregate individual answers to all these questions into synthetic communication indices whose construction is explained in Appendix.[17]

### 4.2. Econometric specification

To estimate our earnings equation (11), we need to proxy workers' skills by a set of observable characteristics, which are measured in different units. This implies the following econometric specification:

$$\begin{aligned}\log W_j &= a + b_1.Gender_j + b_2.Educ_j + b_3.Age_j + b_4.Age^2{}_j + b_5.Com_j \\ &+ c_1.Urban_j \text{ X } \log Pop_i + c_{1'}.Suburb_j + c_{1''}.Rural_j + c_2.shareGraduates_i + c_3.meanCom_i + \varepsilon_j .\end{aligned} \tag{14}$$

In equation (14), the (natural) log of the earnings of each worker is regressed on a set of personal characteristics together with a set of city characteristics corresponding to our theoretical specification. The variables *Gender*, *Educ*, *Age* and *Age2* are standard proxies for individual skills in earnings equations. Education is measured by 6 dummies for educational attainment.[18] *Age* and its square proxy for labour market experience, which is unknown.[19] *Com* is the index of individual workplace communication that ranges from 0 to 100. We also use a set of location characteristics. Because population is unknown for rural and (remote) suburban location, we introduce 3 dummy variables *Urban*, *Suburb,* and *Rural* for workers located in urban, suburban and rural areas respectively. At the city level, we proxy the city workforce, $N_i$, by total city population, $Pop_i$ (which is interacted with the *Urban* dummy). The average skill level in a city or area is measured by the share of college and university graduates, our two highest educational attainments. Finally we construct an index of average city communication, *meanCom*, from our individual communication data. It also ranges from 0 to 100. See the Appendix for more details on this index.

The novelty with respect to the standard canonical human capital framework is twofold. First we introduce workplace communication (*Com*) as another individual determinant of earnings.

[17] In a companion paper (Charlot and Duranton, 2003), we explore the details of the answers to these questions.

[18] *Educ1* corresponds to university graduates (with at least three years of higher education), *Educ2* denotes college graduates (two years of higher education), *Educ3* is for high-school graduates, *Educ4* is for graduates from vocational schools, *Educ5* is for junior high-school graduates, and *Educ6* (our reference) corresponds to the absence of degree. See the Appendix for more details.

This is because, as argued above, the human capital of workers is a function of their skills (determined primarily by their education and their age) and of their workplace communication. To capture the external effects of aggregate labour supply highlighted in (11), equation (14) also estimate three city specific coefficients $c_1$, $c_2$ and $c_3$ relating to population, the share of graduates, and aggregate communication.

Turning to the communication equation (12), our econometric specification is:

$$\begin{aligned} Com_j &= A + B_1.Gender_j + B_2.Educ_j + B_3.Age_j + B_4.Age^2{}_j \\ &+ C_1.Urban_j \text{ X } \log Pop_i + C_{1'}.Suburb_j + C_{1''}.Rural_j + C_2.shareGraduates_i + \mu_j. \end{aligned} \quad (15)$$

In this equation, we proxy individual skills, the city workforce and average schooling in the same way as in the earnings equation. The two estimations (14) and (15) have a number of pitfalls, which are discussed in the next section. Before discussing them, we first present some basic OLS results.

### 4.3. Earnings equation: OLS results

[Insert Table 1 about here]

Column 1 of Table 1 is our preferred specification for the earnings equation (14) since it is the one that matches most closely our theoretical specification (equation 11). The effects on earnings (or wages for that matter since we consider annualised earnings for full-time employees) of the standard individual variables (gender, education, age and its square) are in line with what is usually obtained in this type of exercise. The first novel feature in this estimation regards the magnitude of the coefficient on individual communication. A one-point increase in the communication index on scale from 0 to 100 corresponds to a wage increase of 0.46%. More tellingly perhaps, a one-standard-deviation increase in the communication index (21.3 points) corresponds to a wage increase of about 10%. This effect is large and highly significant.

[19] Since we know only the educational attainment and not the number of years of education, we cannot proxy labour market experience in the usual way (i.e., age - number years of education – 5). Note that this variable will capture not only the effects of experience but also cohort effects, etc. This need no worry us because we are not primarily interested in the structural interpretation of the coefficient on this variable.

The effect of city size is in line with previous results in the literature (albeit in the lower tier of existing estimates): an increase of 1% in city size corresponds to a wage increase of 1.019%. This corresponds to scale economies of around 2%. The coefficient on mean city communication is only weakly significant and its magnitude is around one fourth of that on individual communication. Finally, the coefficient on the share of college and university graduates is large like in all previous estimates using this variable (Rauch 1993, Adserà 2000, Moretti, 2000 and 2002, etc). A one percentage point increase in the share of university and college graduates in a city of corresponds to a 0.5% increase in wages for all workers in this city. Stated differently, the measured external returns on college and university education are roughly of the same magnitude as the measured private returns.[20] Note however that the correlation between the log of population and the share of graduates across urban areas is highly significant and quite high at 0.53. Hence the coefficient on the share of graduates may be capturing some of the effects of city size (and conversely).[21]

Columns 2-5 of Table 1 report results for similar estimations where one or more city level regressors are left aside. They confirm the basic findings of Column 1 regarding the importance of individual communication, that of city size, that of the share of graduates and the mixed results for average communication. The very large coefficient on the share of graduates in Column 4 where city size is omitted confirms the strong collinearity between these two variables. Column 6 of Table 1 is a standard earnings equation with no city level variable and no individual communication. Comparing the coefficients on education and experience between Columns 5 and 6 shows that between 20 and 50% of the effects of skills are captured through individual communication. This is consistent with our theoretical specification where education affects earnings directly through a higher effective labour supply and indirectly through more communication (see equations 11 and 12). At the same time however, individual communication seems to have an important independent effect: the $R^2$ goes up by five percentage points when individual communication is added to a standard earnings equation (see Columns 5 and 6).

[20] We do not know in which proportions innate abilities and educational inputs determine educational attainment. This issue is of secondary importance in our analysis because our model takes skills as given and tries to estimate their external effects going through communication. However, any fully thought through policy attempting to internalise human capital externalities will require a precise answer to this question.

[Insert Table 2 about here]

Table 2 reports the results of another series of earnings regressions this time adding more controls. There are two reasons for doing this: robustness and comparison with the results of Rauch (1993) and his followers who use extensive sets of controls for individual and city characteristics. Adding further controls only marginally increases the explanatory power of the regressions. The $R^2$ increases from 49% in our preferred specification to 51% when introducing family, sectoral and occupational controls. With respect to the four coefficients of interest (on individual communication, city population, mean communication and mean education), the changes are small. The effect of individual communication remains strong and very significant. Adding family, sectoral and occupational controls reduces the coefficient on individual communication by at most 15%. The coefficient on city population is reduced by around 20% and it remains highly significant. Interestingly, these extra controls increase the coefficient on mean city communication and improve its significance. With all the extra controls, this coefficient is now significant at 5% whereas in Table 1 it is significant at 10% in only one specification. The coefficient on the share of graduates is mostly unchanged. Overall these results appear to reinforce our previous findings.

In regressions not reported here, we also used city population instead of its log as is common in the literature. In this case, the coefficient on city size is very small but significant and the impact on the other coefficients (apart from the share of graduates) is minimal. We also ran the regressions reported in Table 1 and 2 using two alternative communication indices.[22] Again the differences were minimal. Finally, we also ran our regressions using different classes of city size with no significant change (to the extent that the coefficients can be compared).

To conclude on the earnings equation, it is worth noting that despite the differences in interpretation and the presence of communication variables, our results here are not very different from those obtained in the literature. Because he is also using the share of graduate to capture the external effects of human capital, we can compare directly our results with those of Moretti (2000) who uses US wage data and Moretti (2002) who is doing a similar

[21] This fact is usually ignored in the human capital literature where city size is measured in levels and not in logs if at all.
[22] Again see the Appendix for details.

exercise with production functions. We find that a one point increase in college graduates increases wages between 0.5% and 0.9% whereas Moretti (2000) obtains estimates ranging between 0.4% and 1.9% and Moretti (2002) is between 0.4% and 0.9% for output per worker. Because they are using different variables, a direct comparison with Rauch (1993), Adserà (2000), or Simon and Nardinelli (2002) is more difficult. Nonetheless, they also find external returns to education being between 30% and 100% of the private returns.

**4.4. Communication equation: OLS results**

To explore the determinants of individual workplace communication, the latter is regressed on a set of individual characteristics and a set of city level variables.[23] Note first that because the residuals in our two equations are not a priori correlated, it is legitimate to use another OLS. Results for multi-stage regressions are reported in Section 5.

[Insert Table 3 about here]

Column 1 of Table 3 is our preferred specification for the communication equation (15) because it offers the closest match with our theoretical specification (equation 12). Looking first at the coefficients on the individual characteristics, we find evidence of a communication gender gap. The communication score of women is about 3.9 points below that of men. Given that the mean communication score is around 37, the relative communication gap between men and women is around 10%. This is less than the earnings gender gap, which is around 20%. In view of the results of Columns 4 and 5, this gender coefficient does not proxy for sectoral or occupational differences. Quite the contrary, the gender communication gap increases when these controls are introduced (see Columns 3 and 4). The coefficients on educational attainments show that education is strong determinant of workplace communication. The difference between university and high-school graduates (i.e., *Educ1* and *Educ3*) is 14 points whereas that between high-school graduates and dropouts with no degree is even larger at 24 points (i.e., more than one standard deviation). As shown by columns 3 and 4, only a small part of this gap is accounted for by sectoral and occupational dummies. There is also evidence of a small age effect. An extra year of age at 20 increases communication by 0.6 points. At age 40, the increase is smaller: 0.15 points. Communication

[23] We are not aware of any similar analysis in the literature. The only exception is Gaspar and Glaeser (1998) who regress a few city level communication variables on city characteristics.

peaks at around 50 years, slightly before earnings.

Turning to city level variables, we find first a significant but small effect of city population. Moving from a city with a population of 10,000 to one with 5 million corresponds to an increase in the individual communication score of 2.4 points. Regarding the share of college and university graduates, we again find a significant but small effect. Workers in a city with twice the national share of graduates (i.e., 30% instead of 15%) see their communication score increase by 2 points.

Column 2 of Table 3 reports results when mean city communication is added as regressor. This variable appears highly significant and the coefficient on city size becomes insignificant. This could be because workers communicate more in environments where there is more communication, a feature ignored by our theoretical specification. However, we must remain cautious here because the *meanCom* is constructed by aggregating individual observations. As a result, this variable is correlated with individual communication in small cities where the number of observations is low.

Adding sectoral dummies (Column 3) to the preferred estimation changes close to nothing. By contrast, adding occupational dummies (Column 4) makes the city level variables less significant or not significant at all. This was to be expected because managerial occupations in which workers communicate a lot are overwhelmingly located in large and highly educated cities. However, occupation and location could well be jointly determined making the interpretation of Column 4 problematic. Even when accepting the results in this column at face value, one would need to explain why high-communication positions tend to be disproportionately located in larger and more educated cities.

The main conclusions we draw from Table 3 are the following. The local environment matters to determine individual communication but the effects are small – much smaller than in the earnings equation. Also, individual communication is fairly well explained by individual characteristics (gender, age, and education). However, the $R^2$ are on average below those obtained in the earnings equations, hinting at the importance of unobserved social skills ($\mu_j$).

What about the earnings effects of communication externalities then? Using the Columns 1 of

Tables 1 and 3, a one log point increase in city population increases individual communication by 0.385. This in turn implies an increase of $0.385 \times 0.00458 = 0.00176$ log point for the wage. This corresponds to a wage increase of about 0.18%. At the same time, the direct coefficient on city size in the earnings equation is about ten times as large at 1.9%. Turning to the share of university and college graduates, a percentage point increase in the local fraction of graduates increases communication by 0.131, which implies a wage effect of +0.06%. The direct coefficient of the share of graduates in the earnings equation is about 8 times as large. When calculating the standard errors for the effects of communication externalities on earnings, they are both significant at 5%. Hence communication externalities appear to take place in cities but they are small. About one tenth of the benefits of city size percolate through communication.[24] Turning to the external returns to education, again around 12% only seem to pass through communication. Stated differently, the external returns to education mediated by communication correspond to around 10% of the private returns.

## 5. Estimation issues and robustness tests

Estimating equations (14) and (15) with simple OLS raises a series of issues that need to be discussed. They regard the spatial sorting of workers, the possible endogeneity of communication, city size, and education. Let us discuss them in turn.

### 5.1. Spatial sorting

The model above assumes that earnings differences between cities are explained only by exogenous differences in population and human capital. Furthermore, the disturbance term on individual labour supply was assumed to be iid across cities, just like the social skills (i.e., the disturbance term in the communication equation). Instead, large cities could attract workers with high levels of human capital and social skills because of better amenities for instance. This alternative explanation could easily explain the strong correlation between average

[24] There is also an indirect effect of these communication externalities. A greater share of graduates and/or a larger population increase the human capital of all workers through communication externalities. In turn a larger human capital has a positive effect on output through the other externalities. This indirect (or combined) effect is however very small because the effects of average communication are only about a quarter of those of individual communication. Furthermore, this indirect effect is only weakly significant.

human capital and earnings.[25] This is a long-standing worry. As already noted by Alfred Marshall (1890 - p.199): “the large towns and especially London absorb the very best blood of all the rest of England; the most enterprising, the most highly gifted, those with the highest physique and strongest character go there to find scope for their abilities”. Of course what applies to London could certainly apply to Paris.

To make a first cut at this issue, note that our data contains some information about the birthplace of workers. This information is at the level of the 95 French 'départements' which cover the country. This is certainly not ideal but we should be able to detect spatial sorting (if any) through different outcomes for those who still work close to their birthplace (the 'stayers') and the others (the 'movers'). We started by running our preferred earnings and communication equations with a full set of dummies. We gradually got rid of those that were not significant to reach the specifications that appear in Table 4. Column 1 reports the earnings equation with the relevant dummies for the movers. Interestingly we find that the returns to communication for movers are higher even after controlling for age, gender, education and city characteristics. We also find that movers benefits more from being in larger cities and in cities with more communication. Column 2 reports the communication equation. Consistent with Column 1, it shows that movers communicate more but that their communication is influenced by city size less than that of movers.

These findings support the idea that workers sort themselves across cities. Workers with better unobserved abilities or social skills tend to go to larger cities and are able to benefit more from them. At the same time, these workers are less sensitive to their environment in their communication behaviour. Hence, for movers only 7% of the effects of city size appear to percolate through communication externalities instead of 10% in the basic estimates. For stayers, about 25% of the effects of city size percolate through communication externalities.

In conclusion, we seem to have, on the one hand, good communicators who go to large cities and benefit strongly from them. These good communicators also have high returns to their own communication but they do not gain much from communication externalities. On the other hand, we have stayers with less favourable characteristics who are able to benefit relatively more from communication externalities. However, this evidence of spatial sorting

[25] However it may be problematic with respect to communication and wages. See below.

does not modify greatly our previous conclusions.[26] Communication externalities account for only a small share of the external effects of city size and average human capital.

[[More to be done here with the estimation of a model instrumenting city size using the place of birth and the work-status of the mother and the profession of the father at birth.]]

### 5.2. Endogeneity of communication

Our communication variable has a lot of intuitive appeal but may be suspected of being simultaneously determined with the wage. Indeed, workers with higher (unobserved) abilities are likely to occupy positions implying both a higher wage and more communication. More generally, we may have ignored a crucial variable, which determines both earnings and communication. This estimation problem can be investigated using instrumental variables (Davidson and Mackinnon, 1993). Ideally, we need exogenous variables, which determine workplace communication, but remain un-correlated with the residual in our earnings equation.

In this respect, we can use the part of our data related to working conditions and changes in the work environment. Time spent daily in front of computer and having to do repetitive movements at work can be used as predictors for communication. Important recent changes in the work environment or in the organisation of work also predict the communication score of workers since such changes are typically correlated with a more important role for information technologies. These variables are not correlated with the residual in the earnings equation.

A Hausman specification test (Green, 2000) shows that indeed endogeneity is an issue. However instrumenting communication by organisational change and working conditions does not change dramatically our results. Comparing the IV results reported in Table 5 with those of Table 1, it appears first that the coefficient on individual communication is mostly

[26] Even if sorting accounted for all our results, one would need to explain why 'good communicators' cluster in large cities where more communication takes place. If this is because they want to communicate more and these places offer opportunities to do so (or a lower cost), the spirit of our results would not be modified. More serious would be the case of an omitted ability bias such that (i) it makes workers more efficient, (ii) it is correlated with large city location, and (iii) it leads workers to communicate more in a world where (iv) communication plays no productive role. However, it is unclear to us which plausible alternative model could satisfy these four conditions.

unchanged just like all the other coefficients related to individual variables. The coefficient on city population is now 10% higher at 2.2% whereas that of the share of graduates decreases by around 15%. Overall, it seems indeed that there is some simultaneity between communication and earnings. But after instrumenting for communication, the overall results with respect to the importance of communication externalities are mostly the same.

### 5.3. Other issues

Further to these considerations, one may assume that the provision of education is biased with high-wage cities offering more and better education.[27] This channel could explain a positive correlation between earnings and average human capital. In France, primary and secondary education are managed mostly by the central government. The French education system is also very concerned about equity within the country. Consequently, programmes are the same everywhere. The central government also aims to equalize class size, resources spent per pupil, etc. Hence a strong bias at the level of primary or secondary education is unlikely. The case of higher education is more subtle. Paris hosts most of the best universities and 'grandes écoles'.[28] The recruitment of most grandes écoles is national in scope so that any bias should be part of the spatial sorting bias already highlighted. Universities have a local catchment area at the bachelor level and for Parisian universities, a national recruitment at the master and doctoral level. Hence, at the master and doctoral level, we are back again to spatial sorting. The unexplored bias regards Parisian students, who during their first two or three years at university, are able to study in more prestigious institutions like La Sorbonne instead of a local university. However, funding per student across universities is also equalised. Whether the more prestigious professors in Parisian universities are enough to create a large bias at the bachelor level is somewhat doubtful.[29]

Finally, one may argue that city size is endogenous to wages. This again seems very unlikely. As shown by Eaton and Eckstein (1997), French cities have experienced parallel population growth over the last 200 years. Hence it is difficult to argue that causation runs from high wages to population growth. To corroborate this, Ciccone (2002) instruments the population

[27] These issues are only given a cursory treatment here. They are successfully dealt with by Moretti (2000).
[28] France has a two-tier system for higher education. 'Grandes écoles' (engineering, business and literature) have competitive entrance examinations whereas there is no selection to enter university.
[29] In his careful instrumentation of the share of graduates in the US, Moretti (2000) found that accounting for endogeneity here did not modify his conclusions, nor would the use of the longitudinal dimension of his data.

of French départements in his analysis of agglomeration effects in Europe and finds only very weak endogeneity problems.

## 6. Conclusions

This paper addresses the issue of the external returns to human capital in cities. Rather than trying to estimate the total external returns to education, it focuses on one particular channel through which human capital externalities are often alleged to percolate: communication externalities at the workplace. To estimate such externalities, we use a unique French data set which surveys workplace communication for around 6000 workers in 1997. This allows us to estimate both a communication equation and an earnings equation using individual workplace communication as independent variable along with other personal and city characteristics.

We find that only around 10% of the effects of a larger and more educated city percolate through communication. Even acknowledging unobserved quality effects, it seems highly unlikely that communication externalities account for the bulk of the positive effects of larger and more educated cities on earnings. At the same time however, it is possible to interpret these findings in a more positive light and argue that communication externalities appear to be indeed present in cities.

These conclusions seem to indicate two directions for future research. First, we are yet to explore the full richness of our data. These findings based on an aggregate communication index warrant further research into the details of the different media being used, the workers involved, the type of communication taking place and how location matters with respect to these issues. The analysis of workplace communication could also be enriched by looking at the level of the firm rather than individual workers. Second, channels other than communication should be explored in more details. In particular and with respect to city size and average urban schooling, the benefits of labour market pooling and input-output linkages should receive more attention.

## References

ABDEL-RAHMAN H. M., 1988, “Product Differentiation, Monopolistic Competition and City Size“, *Regional Science and Urban Economics* 18, 69-86.

ACEMOGLU D. and ANGRIST J., 2000, “How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws“, *NBER Macroannual*, pp 9-59.

ADSERÀ A., 2000, “Sectoral Spillovers and the Price of Land: A Cost Analysis across U.S. States“, *Regional Science and Urban Economics* 30, 565-585.

ALMEIDA P. and KOGUT B., 1999, “Localization of Knowledge and the Mobility of Engineers in Regional Networks”, *Management Science* 45, 905-917.

BECKMANN M. J., 1976, “Spatial Equilibrium in the Dispersed City”, in Y.Y. Papageorgiou (ed.), *Mathematical Land Use Theory*, Lexington Books, Lexington (Mass.), 117-125.

BERLIANT M., REED J. and WANG P., 2001, “Knowledge Exchange, Matching and Agglomeration“, processed, University of Washington, St Louis.

BLACK D. and HENDERSON V., 1999, “A Theory of Urban Growth“, *Journal of Political Economy* 107, 252-284.

BORUKHOV E., and HOCHMAN O., 1977, “Optimum and the Market in a Model of City without a Predetermined Center”, *Environment and Planning A* 9, 849-856.

CHARLOT S. and DURANTON G., 2003, “Cities and Workplace Communication: Some Descriptive Evidence“, mimeo in progress.

CICCONE A., 2002, “Agglomeration Effects in Europe“, *European Economic Review* 46, 213-227.

CICCONE A. and PERI G., 2002, “Identifying Human Capital and Externalities: Theory with an Application to US Cities“, CEPR discussion paper 3350.

COHEN D. and SOTO M., 2001, “Growth and Education : Good Data, Good Results“, CEPR discussion paper 3025.

DAVIDSON R. and MACKINSON J. G., 1993, *Estimation and Inference in Econometrics*, Oxford University Press.

DIXIT A. K. and STIGLITZ J. E., 1977, “Monopolistic Competition and Optimum Product Diversity“, *American Economic Review* 67, 297-308.

DUMAIS G., ELLISON G. and GLAESER E., 1997, “Geographic Concentration as a Dynamic Process”, NBER Working Paper 6270.

DURANTON G. and PUGA D., 2003, “Microeconomic Foundations of Urban Increasing Returns“, in J. F. Thisse and J. V. Henderson (eds), *Handbook of Regional and Urban*

*Economics*, vol 4, North-Holland, Amsterdam, forthcoming.
EATON J. and ECKSTEIN Z., 1997, "Cities and Growth: Theory and Evidence from France and Japan", *Regional Science and Urban Economics* 27, 443-474.
ETHIER W. J., 1982, "National and International Returns to Scale in the Modern Theory of International Trade", *American Economic Review* 72, 389-405.
FUJITA M., 1988, "A Monopolistic Competition Model of Spatial Agglomeration: A Differentiated Product Approach", *Regional Science and Urban Economics* 18, 87-124.
FUJITA M., and OGAWA H., 1982, "Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations", *Regional Science and Urban Economics* 12, 161-196.
GASPAR J., and GLAESER E. L., 1998, "Information Technologies and the Future of Cities", *Journal of Urban Economics* 43, 136-156.
GLAESER E. L., 1999, "Learning in Cities", *Journal of Urban Economics* 46, 254-277.
GLAESER E. L., and MARÉ D., 2001, "Cities and Skills", *Journal of Labor Economics*, 19, 316-342.
GODDARD J. B., 1973, "Office Linkages and Location: A Study of Communications and Spatial Patterns in Central London", *Progress in Planning* 1, 111-232.
GODDARD J. B. and MORRIS D., 1976, "The Communication Factors in Office Decentralisation", *Progress in Planning* 6, 1-80.
GREENE W. H., 2000, *Econometric Analysis*, Prentice Hall International, London.
GREENAN N., and HAMON-CHOLET S., 2001a, "COI un Dispositif d'Enquêtes Couplées Employeurs / Employés sur les Changements Organisationnels et l'Informatisation. Tome 1 : Présentation et Questionnaires", processed Centre d'Etude de l'Emploi, Paris.
GREENAN N., and HAMON-CHOLET S., 2001b, "COI un Dispositif d'Enquêtes Couplées Employeurs / Employés sur les Changements Organisationnels et l'Informatisation. Tome 2 : Dictionnaire", processed Centre d'Etude de l'Emploi, Paris.
GREENAN N., and MAIRESSE J., 1999, "Organisational Change in Manufacturing: What Do We Learn From Firm Representatives and From Their Employees ?", NBER Working Paper 7285.
HOLMES T. J., 2002, "The Role of Cities: Evidence from the Placement of Sales Offices", Federal Reserve Bank of Minneapolis Staff Report #298
JAFFE A. B, TRAJTENBERG M., and HENDERSON R., 1993, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics* 108, 577-598.
JOVANOVIC B., and ROB R., 1989, "The Growth and Diffusion of Knowledge", *Review of*

*Economic Studies* 56, 569-582.
LUCAS R. E. JR, 1988, "On The Mechanics of Economic Development", *Journal of Monetary Economics* 22, 3-42.
LUCAS R. E. JR, 2001, "Externalities and Cities", *Review of Economic Dynamics* 4, 245-274.
MARSHALL A., 1890, *Principles of Economics*, Macmillan, London.
MORETTI E., 2000, "Estimating the Social Return to Education: Evidence From Longitudinal and Cross-Sectional Data", processed UCLA.
MORETTI E., 2002, "Human Capital Spillovers in Manufacturing: Evidence from Plant-Level Production Function", NBER Working Paper 9316.
RAUCH J. E., 1993, "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics* 34, 380-400.
RIVERA-BATIZ F., 1988, "Increasing Returns, Monopolistic Competition, and Agglomeration Economies in Consumption and Production", *Regional Science and Urban Economics* 18, 125-153.
ROBACK J., 1982, "Wages, Rents, and the Quality of Life," *Journal of Political Economy* 90, 1257-1278.
ROSENTHAL S.S. and STRANGE W.C., 2003, "Microfoundations of Urban Increasing Returns: Empirical Evidence", in J. F. Thisse and J. V. Henderson (eds), *Handbook of Regional and Urban Economics*, vol 4, North-Holland, Amsterdam, forthcoming.
SAXENIAN A., 1994, *Regional Advantage*, Harvard University Press, Cambridge (Mass.).
SIMON C. J. and NARDINELLI C., 2002, "Human Capital and the Rise of American Cities, 1900-1990", *Regional Science and Urban Economics* 32, 59-96.
STORPER M., 1997, *The Regional World: Territorial Development in a Global Economy*, The Guilford Press, New York.
TOPEL R., 1997, "Factor Proportions and Relative Wages," *Journal of Economic Perspectives* 11, 55-74.

## Appendix 1. Data description

### The enriched COI data

The basic data is from the 1997 “Changement Organisationnel et Informatisation” (COI) survey. This data is composed of four different business surveys matched with one labour force survey. The first business survey (manufacturing) was conducted by the French Ministry of Industry. The second (food industry) was conducted by the French Ministry of Agriculture while the last two (DIY chains and accountants) were carried out by INSEE (French National Institute for Statistics and Economics Studies). The conception of the business survey in manufacturing, that of the labour survey, and the coordination of the 4 surveys were directed by Nathalie Greenan at the Centre d’Etudes de l’Emploi at the French Ministry of Labour.

This firm/employee matched survey is mostly concerned by organisational change and information technologies. This data was later on matched with the Déclaration Annuelle Des Salaires (DADS) and with Enquête Annuelle d'Entreprises (EAE) also from INSEE. The DADS data is collected for fiscal purpose. It is exhaustive on all French salaried workers and contains information about employment and earnings. Furthermore, for all workers born in October of even years (those selected for COI), it also contains a wealth of personal characteristics. The EAE survey is also exhaustive for all firms with more than 20 employees. It contains a wealth of firm level data.

Initially 4025 representative firms were selected from general manufacturing (2541), the Food industry (478), accounting firms (734), and DIY shops (272). Within each group (general manufacturing, food industry, accounting, DIY), firms were randomly drawn among those with 50 or more employees. This sectoral heterogeneity in the nature of the surveyed firms need not worry us because in most cases, the data is broken down by sector.

Interviewers went to interview directly 1, 2 or 3 randomly chosen employees in each selected firm. When it was impossible to meet face-to-face with an employee, the interviewers did the survey on the phone. A total of 8812 employees were initially drawn. In total, 6177 employee questionnaires were obtained from 3153 firms. The 30% of non-respondents include employees who refused to respond (about 9%), those who could not be found by the interviewers (11 %) and those who had left their firm by the end of the year (9%) and could not subsequently be matched with the DADS data. Further details on the data can be found in Greenan and Mairesse (1999) and Greenan and Hamon-Cholet (2001a,b).[30] The match with firm level data led to a loss of another 594 observations.

**Location data**

Using postcode data at the establishment level, this survey was matched to a set of spatial units. We lost another 101 observations in the process. Metropolitan France contains 361 urban areas where employment is at least 5,000. The rest of the country is classified into different levels of 'peri-urban' (i.e., remote suburban) and rural areas. For simplicity, outside urban areas we only distinguish suburban from rural areas.

Note that the French definition of urban areas in this typology is rather broad and matches rather closely that of metropolitan areas in the US except that the threshold is much below (5,000 jobs instead of 100,000 inhabitants). To be consistent with the breadth of the definition of urban areas, the definition of suburban area is rather restrictive and narrow. A 'suburban' area in this typology is usually a rather remote ring around an urban area (a.k.a. exurban or peri-urban). Unfortunately, because a significant fraction of these remote suburban areas are functionally linked with two or more cities, they cannot be matched to particular adjacent urban areas. Urban areas contain about 60% of the French population and 70% of French employment. Around 65% of our observations are located in urban areas. This slight urban under-representation is due to the large number of establishments surveyed in the food industry whose location is often rural. Suburban areas account for around 10% of our observations. Finally, 25% of the observations are located in rural areas.

**The communication questions**

The employee questionnaire contains around 80 questions covering a wide range of topics about working conditions, organisational change and information technologies. Regarding communication, the most relevant questions are:

[30] Among many quality checks, they compared the COI data on education with similar data in the DADS and found only minimal discrepancies. They also checked answers to questions about organisational change in the employee questionnaire against similar questions in the employer questionnaire. Unsurprisingly, the differences were larger.

30. Are you in contact (face-to-face or telephone) with customers? (All the time/Regularly/At times/ Never)

31. Except for your subordinates (if any), do you give instructions to the following persons about their work?

a/ Colleagues with whom you are usually working: Yes/No/Not Applicable.

b/ Others, working for the same firm: Y/N/NA.

c/ Others, working for another firm (customer, supplier, etc.): Y/N/NA.

32. Except for your superior(s), do you receive instructions from the following persons about your work?

a/ Colleagues with whom you are usually working: Y/N/NA.

b/ Others, working for the same firm: Y/N/NA.

c/ Others, working for another firm (customer, supplier, etc.): Y/N/NA.

34. How do you receive important instructions about your work?

a/ Face-to-face communication: Y/N.

b/ Telephone: Y/N.

c/ Paper (including fax, telex, etc.): Y/N.

d/ Computer (electronic mail, etc.): Y/N.

40. Do you do some of your work in a team? Y/N.

40b. If yes, are you involved with the following?

a/ Colleagues from the same unit: Y/N.

b/ Others, from the same firm: Y/N .

c/ Others, external from your firm: Y/N .

40c. If yes, what type of work is concerned? Conception (or design, or research) / Production

52. Do you use ever a PC or a workstation at work? Y/N.

55. Do you use information technologies to search for information? Y/N.

68. Do you use internet at work? Y/N.

69. Do you use an intranet at work? Y/N.

Note that these questions refer to different aspects of workplace communication. Question 30 regards the intensity of external communication with customers. Unfortunately this is the only question in this direction. Questions 31, 32, 40a and b refer to whom the worker communicate with. Note also that questions 31 and 32 relate to "instructions" received or given by the worker. The questionnaire defines instructions as "important information received or given by the worker on a regular basis to conduct his or her work". Gossiping around the coffee machine does not constitute an instruction whereas asking for help to carry out a task does. Question 40c is of particular interest because it relates to creative activities upon which the literature puts special emphasis. The remaining questions refer to the media used by the worker for his or her workplace communication. The main weakness of this data is that we know where the surveyed workers are located but we do not know where the workers they communicate with are.

**Our communication indices**

The answers to all binary questions were coded 1 for yes and 0 for no. Answers to questions 30 were coded 4 for the highest level of communication, 3, 2, and 1 for the lowest. We then aggregated this into a variety of communication indices. Our main index, *Com*, is a weighted average. We gave equal an weight to the following five dimensions:

- Communication internal to the firm (sum of 31a, 31b, 32a, 32b, 40b a, and 40b-b).
- Communication external to the firm (sum of 31c, 32c, and 40b-c).
- Intensity of communication (30).
- Media (sum of 34a, b, c, d, e, 52, 55, 68, and 69).
- Involvement in creative activities with others (40c).

This index is normalised to be between 0 and 100. The mean score for our sample is 37.1 and the standard error is 21.3. We also constructed two alternative indices: *Com_b* and *Com_c*. *Com_b* is simply the sum of the score to all questions normalised to take a maximum of 100. *Com_c* is a slightly more complex index where external communication together with involvement in creative activities is given the same weight as internal communication. The sum of the two is then multiplied by the intensity of communication as measured by question

30. This product counts for 75%. The remaining 25% is given by the sum to the media questions. Again this index is normalised to take a maximum of 100.

We prefer *Com* because it gives more weights to the intensity of communication with persons external to the firm and the involvement in creative activities, which we view as potentially important channels for communication externalities. *Com_b* is much simpler but internal communication may be over-represented because of the wealth of questions on this issue. *Com_c* is possibly more realistic since communication is weighted by its intensity but it relies to a large extent on the answer to only one question (30). The pair-wise correlation between these indices is between 0.93 and 0.96 across workers. We replicated all our regressions using *Com* with *Com_b* and *Com_c*. Unsurprisingly, the results were never very different.

**Our average communication index**

The variable measuring average communication is constructed from our individual data. We computed a weighted mean city communication $meanCom_i$. We summed across educational attainments (denoted *g*), the product of the 1999-census share of shares of workers ($s_i(g)$) by their empirical mean level of communication ($Com_i(g)$):

$$meanCom_i = \sum_{g=1}^{6} s_i(g) Com_i(g), \qquad \text{(A1)}$$

The problem is that in small cities, we have no observation for some educational attainments. To fill the blanks and using all the observations we have for the city, we computed the following communication ratio:

$$q_i = \sum_{g=1}^{6} \left[ \frac{n_i(g)}{n_i} \frac{Com_i(g)}{Com(g)} \right], \qquad \text{(A2)}$$

where $n_i(g)$ is the number of workers in the sample with educational attainment *g* in city *i*, $n_i$ is the sampled population in city *i*, and $Com(g)$ is the national mean of workers of educational attainment *j*. We then proxied $Com_i(g)$ by $q_i Com(g)$.

Because we use the real proportion of each educational attainment, this index avoids sample sorting biases (in this respect) for small urban areas for which we have no more than a handful of observations.

**Working condition variables**

To instrument for the possible endogeneity of workplace communication, we use data regarding working conditions:

- *Computer use*: number of hours spent daily in front of a computer.
  (The information is coming from Question 79: On average how many hours do you spend in front of your computer if at all? Hours per week on average.)
- *Repetitive movements*: whether the position occupied by the worker involves the repetition of the same movement(s).
  (Question 43. Does your work involve the continuous repetition of the same series of movements and/or operations? Y/N)
- *Organisational change:* whether the working environment and the organisation of work has changed significantly in the last three years.
  (Questions 8 and 8b. Has your working environment changed dramatically in the last three years? Y/N. If yes, is it because of a change of position? Y/N. A change of technique? Y/N. Restructuring of the firm? Y/N. A change in the organisation of your work? Y/N. Other?)

**Other variables**

The COI data contains 14 possible levels of educational attainment. Given the small population in some cells and the lack of discernible wage differences between some cells, we aggregated these 14 categories into the six which are used by the French census.

*Educ1*: university graduates (a degree involving at least three years of higher education).
*Educ2*: college graduates (two years of higher education).
*Educ3*: high-school graduates.
*Educ4*: vocational school graduates.
*Educ5*: junior high-school graduates.

*Educ6* (our reference): no degree (early school drop-out).

The variable *shareGraduates* was constructed from the 1999 census. It reports the share of the population in each area with our two highest levels of educational attainments (Educ1 or Educ2).

To distinguish retail (DIY) and accounting firms from manufacturing, we created two dummy variables:
*Accounting* for workers employed in accounting firms.
*Retail* for workers employed in DIY firms.
Workers in manufacturing are used as the reference.

The *Wage* variable refers to the net annualised earnings received by the employee. It comes from the DADS data. This data is collected from all employers (and self-employed) in France for pension, benefits and tax purposes. A report must be filled by every establishment for each of its employees so that there is a unique record for each employee-establishment-year combination. The mandatory aspect of this data is a guarantee of its quality.

The occupational dummy *Non-routine* was created from the detailed information in the COI data about the occupation. We classified as “non-routine” all jobs which *a priori* involve some autonomy for the employee. All other occupations were classified as “routine”. Among non-routine jobs, we have functions such as: sales, management, accounting, research, teaching, etc whereas we classified as routines functions like cleaning, production, domestic work, handling of goods, etc. This distinction attempts to generalise the usual white collar / blue collar opposition in manufacturing. We tried more detailed breakdowns for the occupations. This only led to minor changes in our results.

Finally we exploited the information about the place of birth (at the level of the 95 French départements --- these units are comparable to counties in the US or the UK) and the place of work. If someone works in the département he or she was born, this person is classified as a stayer. Otherwise, this person is a mover. In our sample, we have 43% of movers and 57% of stayers.

[[Mother's work status at birth and father's profession at birth to be described]]

## Appendix 2. Results

Table 1: Earnings equations

logW

| Regressors | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Intercept | 1.618 | 1.600 | 1.670 | 1.767 | 1.936 | 1.911 | 1.629 |
| | (0.0956) | (0.0957) | (0.0892) | (0.0903) | (0.0826) | (0.0856) | (0.0919) |
| Gender | -0.209 | -0.207 | -0.209 | -0.210 | -0.1978 | -0.217 | -0.228 |
| | (0.00932) | (0.00932) | (0.00932) | (0.00932) | (0.00934) | (0.00962) | (0.00955) |
| Educ1 | 0.573 | 0.578 | 0.570 | 0.576 | 0.624 | 0.827 | 0.746 |
| | (0.0218) | (0.0218) | (0.0217) | (0.0218) | (0.0212) | (0.0193) | (0.0201) |
| Educ2 | 0.332 | 0.331 | 0.329 | 0.336 | 0.351 | 0.513 | 0.473 |
| | (0.0197) | (0.0198) | (0.0197) | (0.0198) | (0.0197) | (0.0186) | (0.0186) |
| Educ3 | 0.211 | 0.210 | 0.209 | 0.214 | 0.221 | 0.346 | 0.0320 |
| | (0.0188) | (0.0189) | (0.0188) | (0.0188) | (0.0189) | (0.0185) | (0.0184) |
| Educ4 | 0.109 | 0.107 | 0.107 | 0.109 | 0.109 | 0.176 | 0.167 |
| | (0.0145) | (0.0145) | (0.0144) | (0.0145) | (0.0146) | (0.0147) | (0.0145) |
| Educ5 | 0.0217 [ns] | 0.0194 [ns] | 0.0204 [ns] | 0.0223 [ns] | 0.0175 [ns] | 0.0434 | 0.0434 [a] |
| | (0.0185) | (0.0185) | (0.0185) | (0.0185) | (0.0187) | (0.0194) | (0.0190) |
| Age | 0.0705 | 0.0702 | 0.0704 | 0.0706 | 0.0701 | 0.0759 | 0.0756 |
| | (0.00406) | (0.00407) | (0.00406) | (0.00407) | (0.00408) | (0.00422) | (0.00418) |
| $Age^2$ | -0.000616 | -0.000612 | -0.000614 | -0.000616 | -0.000600 | -0.000660 | -0.000668 |
| | (0.0000491) | (0.0000492) | (0.0000491) | (0.0000492) | (0.0000492) | (0.0000560) | (0.0000505) |
| **Com** | **0.00458** | **0.00466** | **0.00467** | **0.00461** | **0.00507** | | |
| | (0.000261) | (0.000261) | (0.000254) | (0.000261) | (0.000253) | | |
| Rural | 0.205 | 0.305 | 0.208 | | | | 0.214 |
| | (0.0402) | (0.0338) | (0.0401) | | | | (0.0414) |
| Suburb | 0.224 | 0.345 | 0.227 | | | | 0.244 |
| | (0.0443) | (0.0355) | (0.0442) | | | | (0.0456) |
| **logPop** | **0.0192** | **0.0297** | **0.0195** | | | | **0.0217** |
| | (0.00349) | (0.00263) | (0.00348) | | | | (0.00359) |
| **meanCom** | **0.00116** [c] | **0.000581** [ns] | | **0.00142** [b] | | | |
| | (0.000771) | (0.000762) | | (0.000771) | | | |
| **shareGraduates** | **0.490** | | **0.463** | **0.891** | | | **0.525** |
| | (0.107) | | (0.105) | (0.0740) | | | (0.108) |
| Adj. $R^2$ | 0.489 | 0.487 | 0.489 | 0.486 | 0.483 | 0.446 | 0.456 |
| N. obs | 5329 | 5329 | 5329 | 5329 | 5329 | 5329 | 5329 |

Note: Standard errors are shown in parentheses. All coefficient significant at the 1% level except [a] significantly different from zero at the 5% level, [b] significantly different from zero at the 10% level and [c] significantly different from zero at the 15% level, [ns] not significant.

# Table 2: Augmented earnings equations

logW

| Regressors | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 1.750 | 1.776 | 1.726 | 1.747 |
| | (0.119) | (0.0984) | (0.0990) | (0.0993) |
| Gender | -0.274 [c] | -0.225 | -0.229 | -0.197 |
| | (0.172) | (0.00997) | (0.00992) | (0.00972) |
| Educ1 | 0.502 | 0.505 | 0.541 | 0.559 |
| | (0.0222) | (0.0222) | (0.0223) | (0.0219) |
| Educ2 | 0.259 | 0.262 | 0.301 | 0.320 |
| | (0.0203) | (0.0203) | (0.0204) | (0.0198) |
| Educ3 | 0.163 | 0.165 | 0.183 | 0.212 |
| | (0.0192) | (0.0192) | (0.0194) | (0.0188) |
| Educ4 | 0.0887 | 0.0890 | 0.0972 | 0.111 |
| | (0.0143) | (0.0143) | (0.0146) | (0.0143) |
| Educ5 | 0.0202 [ns] | 0.0194 [ns] | 0.0251 [ns] | 0.0210 [ns] |
| | (0.0181) | (0.0181) | (0.0184) | (0.0183) |
| Age | 0.0614 | 0.0615 | 0.0656 | 0.0624 |
| | (0.00547) | (0.00431) | (0.00437) | (0.00435) |
| $Age^2$ | -0.000508 | -0.000524 | -0.000563 | -0.000525 |
| | (0.0000652) | (0.0000519) | (0.0000526) | (0.0000524) |
| Children | 0.0170 | 0.0129 [a] | 0.0128 | 0.0116 |
| | (0.00486) | (0.00408) | (0.00415) | (0.00412) |
| Gender X Age | 0.000775 [ns] | | | |
| | (0.00889) | | | |
| Gender X $Age^2$ | -0.0000141 [ns] | | | |
| | (0.000108) | | | |
| Gender X Children | -0.0194 [a] | | | |
| | (0.00895) | | | |
| Non-routine occup. | 0.135 | 0.138 | 0.0794 | |
| | (0.0131) | (0.0131) | (0.0123) | |
| Accounting | 0.0959 | 0.0952 | | 0.0470 |
| | (0.0140) | (0.0139) | | (0.0133) |
| Retail | -0.160 | -0.160 | | -0.156 |
| | (0.0217) | (0.0217) | | (0.0219) |
| **Com** | **0.00392** | **0.00392** | **0.00386** | **0.00502** |
| | (0.000279) | (0.000279) | (0.000284) | (0.000262) |
| Rural | 0.145 | 0.150 | 0.206 | 0.164 |
| | (0.0396) | (0.0396) | (0.0400) | (0.0400) |
| Suburb | 0.165 | 0.168 | 0.223 | 0.184 |
| | (0.0436) | (0.0437) | (0.0441) | (0.0441) |
| **logPop** | **0.0149** | **0.0153** | **0.0187** | **0.0168** |
| | (0.00342) | (0.00342) | (0.00347) | (0.00346) |
| **meanCom** | **0.00161 [a]** | **0.00162 [a]** | **0.00123 [b]** | **0.00147 [b]** |
| | (0.000754) | (0.000755) | (0.000767) | (0.000763) |
| **shareGraduates** | **0.496** | **0.480** | **0.468** | **0.514** |
| | (0.104) | (0.104) | (0.106) | (0.105) |
| Adj. $R^2$ | 0.512 | 0.510 | 0.493 | 0.500 |
| N. obs | 5329 | 5329 | 5329 | 5329 |

Note: Standard errors are shown in parentheses. All coefficient significant at the 1% level except [a] significantly different from zero at the 5% level, [b] significantly different from zero at the 10% level and [c] significantly different from zero at the 15% level, [ns] non significant.

## Table 3: Communication equations

Com

| Regressors | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | -8.814 [b] | -39.611 | -8.677 [b] | -3.453 [ns] |
| | (4.819) | (4.990) | (4.741) | (4.396) |
| Gender | -3.998 | -3.855 | -5.129 | -8.118 |
| | (0.501) | (0.487) | (0.510) | (0.483) |
| Educ1 | 37.834 | 37.480 | 36.934 | 24.363 |
| | (1.052) | (1.022) | (1.025) | (1.037) |
| Educ2 | 30.781 | 30.780 | 30.143 | 18.100 |
| | (0.975) | (0.947) | (0.963) | (0.978) |
| Educ3 | 23.819 | 23.652 | 22.433 | 12.968 |
| | (0.962) | (0.935) | (0.956) | (0.940) |
| Educ4 | 12.828 | 13.002 | 12.310 | 7.575 |
| | (0.760) | (0.738) | (0.748) | (0.711) |
| Educ5 | 4.933 | 5.437 | 5.330 | 4.178 |
| | (0.997) | (0.969) | (0.977) | (0.906) |
| Age | 1.114 | 1.133 | 1.258 | 1.019 |
| | (0.219) | (0.213) | (0.214) | (0.198) |
| Age² | -0.0120 | -0.0120 | -0.0131 | -0.012 |
| | (0.00265) | (0.00257) | (0.00258) | (0.00239) |
| Non-routine occup. | | | | 18.115 |
| | | | | (0.603) |
| Accounting | | | -6.053 | 0.982 [ns] |
| | | | (0.696) | (0.686) |
| Retail | | | 4.790 | 3.397 |
| | | | (1.173) | (1.088) |
| Rural | 1.393 [ns] | -0.314 [ns] | 4.261 [a] | 1.776 [ns] |
| | (2.170) | (2.110) | (2.037) | (1.889) |
| Suburb | 3.578 [c] | 1.156 [ns] | 6.297 | 3.381 [c] |
| | (2.391) | (2.327) | (2.244) | (2.081) |
| **logPop** | **0.385 [a]** | **0.178 [ns]** | **0.528** | **0.253 [c]** |
| | (0.188) | (0.183) | (0.173) | (0.161) |
| meanCom | | 0.701 | | |
| | | (0.0393) | | |
| **shareGraduates** | **13.092 [a]** | **28.238** | **11.458 [a]** | **5.232 [ns]** |
| | (5.687) | (5.590) | (5.410) | (5.016) |
| Adj. R² | 0.318 | 0.357 | 0.345 | 0.438 |
| N. obs | 5329 | 5329 | 5482 | 5482 |

Note: Standard errors are shown in parentheses. All coefficient significant at the 1% level except [a] significantly different from zero at the 5% level, [b] significantly different from zero at the 10% level and [c] significantly different from zero at the 15% level, [ns] non significant.

Table 4: Earnings and communication equations with controls for movers and stayers

| logW | | Com | |
|---|---|---|---|
| Regressors | (1) | Regressors | (2) |
| Intercept | 1.822 | Intercept | -7.520[c] |
| | (0.109) | | (4.808) |
| Mover | -0.345 | Stayer | -2.708 |
| | (0.100) | | (0.831) |
| Gender | -0.209 | Gender | -3.943 |
| | (0.00930) | | (0.494) |
| Educ1 | 0.560 | Educ1 | 37.886 |
| | (0.0219) | | (1.018) |
| Educ2 | 0.329 | Educ2 | 30.696 |
| | (0.0197) | | (0.964) |
| Educ3 | 0.209 | Educ3 | 23.713 |
| | (0.0188) | | (0.955) |
| Educ4 | 0.108 | Educ4 | 12.988 |
| | (0.0144) | | (0.757) |
| Educ5 | 0.0209 [ns] | Educ5 | 5.323 |
| | (0.0185) | | (0.992) |
| Age | 0.0703 | Age | 1.113 |
| | (0.00406) | | (0.216) |
| Age² | -0.000615 | Age² | -0.0117 |
| | (0.0000490) | | (0.00261) |
| **Com** | **0.00392** | | |
| | (0.000333) | | |
| **Mover X Com** | **0.00138** | | |
| | (0.000442) | | |
| Rural | 0.0677 [ns] | Rural | 2.103 [ns] |
| | (0.0602) | | (2.140) |
| Suburb | 0.0998 [c] | Suburb | 4.078 [b] |
| | (0.0635) | | (2.329) |
| Mover X Rural | 0.206 | | |
| | (0.0741) | | |
| Mover X Suburb | 0.171 [a] | | |
| | (0.0771) | | |
| **logPop** | **0.00822 [c]** | **logPop** | **0.325 [b]** |
| | (0.00520) | | (0.177) |
| **Mover X logPop** | **0.0151** | **Stayer X logPop** | **0.197 [a]** |
| | (0.00586) | | (0.0825) |
| **shareGraduates** | **0.449** | | **11.946 [a]** |
| | (0.107) | | (5.536) |
| **meanCom** | **0.0000630 [ns]** | | |
| | (0.000968) | | |
| **Mover X meanCom** | **0.00307** | | |
| | (0.00156) | | |
| Adj. R² | 0.492 | Adj. R² | 0.326 |
| N. obs | 5329 | N. obs | 5482 |

Note: Standard errors are shown in parentheses. All coefficient significant at the 1% level except [a] significantly different from zero at the 5% level, [b] significantly different from zero at the 10% level and [c] significantly different from zero at the 15% level, [ns] not significant.

## Table 5: IV earnings equations

logW

| Regressors | (1) | (2) |
|---|---|---|
| Intercept | 1.635 | 1.646 |
| | (0.0879) | (0.0883) |
| Gender | -0.204 | -0.204 |
| | (0.00922) | (0.00925) |
| Educ1 | 0.584 | 0.583 |
| | (0.0211) | (0.0212) |
| Educ2 | 0.333 | 0.331 |
| | (0.0195) | (0.0195) |
| Educ3 | 0.211 | 0.210 |
| | (0.0187) | (0.0187) |
| Educ4 | 0.110 | 0.109 |
| | (0.0144) | (0.0144) |
| Educ5 | 0.0234 [ns] | 0.0223 [ns] |
| | (0.0184) | (0.0185) |
| Age | 0.0703 | 0.0700 |
| | (0.00402) | (0.00403) |
| Age² | -0.000614 | -0.000614 |
| | (0.0000487) | (0.0000491) |
| **Com** | **0.00472** | **0.00472** |
| | (0.000250) | (0.000251) |
| Rural | 0.237 | 0.237 |
| | (0.0381) | (0.0382) |
| Suburb | 0.260 | 0.260 |
| | (0.0420) | (0.0421) |
| **logPop** | **0.0222** | **0.0222** |
| | (0.00325) | (0.00326) |
| **shareGraduates** | **0.400** | **0.398** |
| | (0.102) | (0.102) |
| Adj. $R^2$ | 0.500 | 0.500 |
| N. obs | 5482 | 5459 |

Note: Standard errors are shown in parentheses. All coefficients significant at the 1% level except [a] significantly different from zero at the 5% level, [b] significantly different from zero at the 10% level and [c] significantly different from zero at the 15% level, [ns] not significant.

Instruments for communication: (1) *Computer use* (see Appendix 1 for a description) and *Repetitive movements*. These variables are significant at the 1% level in the instrumental regression and are not correlated with the error term of the augmented regression.
(2): same instruments in (1) plus *Organisational change*. All instruments are significant at the 1% level in the instrumental regression and are not correlated with the error term of the augmented regression.