Journal of Urban Economics 101 (2017) 27–44

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

Sorting and agglomeration economies in French economics departments

Clément Bosquet[b,1,2], Pierre-Philippe Combes[a,1,3,*]

[a] University of Lyon and Sciences Po, France
[b] ThEMA (University of Cergy-Pontoise) and SERC (LSE), France

ARTICLE INFO

Article history:
Received 13 November 2015
Revised 12 May 2017
Available online 1 June 2017

JEL classification:
R23
J24
I23

Keywords:
Research productivity
Local externalities
Skill sorting
Peer effects
Knowledge spillovers
Co-author networks
Economics of science

ABSTRACT

Are agglomeration and peer effects at stake in academic research? To tackle this question, we study how departments' characteristics affect the quantity and quality of academics' publications in economics in France, controlling for individual time-varying characteristics and individual fixed effects. Department characteristics have an explanatory power at least equal to a quarter of that of individual characteristics and possibly as high as theirs. The quantity and quality of an academic's publications in a field increase with the presence of other academics specialised in that field and with the share of the department's publications output in that field. In contrast, department size, proximity to other large departments, homogeneity in terms of publication performance, presence of colleagues with connections abroad, and composition in terms of positions and age matter for some publication measures but only if not controlling for individual fixed effects.

1. Introduction

Every academic has an opinion about what makes a good department. However, there are surprisingly few econometric studies that quantify this precisely, despite possible implications for the design of education and research institutions, an always-topical concern (see for instance Aghion et al., 2010). Indeed, a large literature documents both the gains from spatial concentration (see Rosenthal and Strange, 2004; Combes and Gobillon, 2015) and the effects of local peers and networks (see Sacerdote, 2011; Jackson, 2011), which all could be at stake in academic departments. Here, we focus on the role on individual publication records in economics in France of both individual characteristics and a large set of departments' characteristics. We develop a careful strategy that controls for possible spatial selection of academics and missing variables.

Both the urban economics and the local peer effects literatures have emphasised the importance and difficulty of disentangling the role of individual sorting from the causal impact of the local environment. What makes individuals productive? Is it their own abilities, or the location (firm, city, school, etc.) where they operate? In the context of universities, do academics publish more because of their higher ability (based on gender, age or some other possibly unobserved characteristics) and a publication strategy that brings higher rewards (e.g. research field, number and location of co-authors)? Or because they are located in departments that provide better local environments and stronger externalities, which

* Corresponding author.
E-mail addresses: clement.bosquet@u-cergy.fr (C. Bosquet), ppcombes@gmail.com (P.-P. Combes).

1 We are very grateful to Philippe Donnay and Charles Laitong for excellent research assistance. We also thank Yann Bramoullé, Gilles Duranton, Laurent Gobillon, Jean-Marc Robin, Maurizio Iacopetta, Frédéric Robert-Nicoud, Elisabet Viladecans-Marsal, Cecilia García-Peñalosa, Alain Trannoy, Alexander Lembcke, Stuart Rosenthal, two anonymous referees, and participants at various seminars and conferences for useful suggestions. This study was made possible and was financed in part by the French Ministry of Research and Higher Education. Pierre-Philippe Combes is a researcher at CNRS, whose financial support is gratefully acknowledged.

2 ThEMA, UMR 8184, Université Cergy-Pontoise, CNRS, F-95000 Cergy-Pontoise. 33 boulevard du Port, 95011 Cergy-Pontoise cedex, France.

3 Univ Lyon, CNRS, GATE Lyon Saint-Etienne UMR 5824, 93 chemin des Mouilles, F-69131 Ecully, France; Sciences Po, Department of Economics, 28, Rue des Saints-Pères, 75007 Paris, France. Also, research fellow at CEPR.

http://dx.doi.org/10.1016/j.jue.2017.05.003
0094-1190/

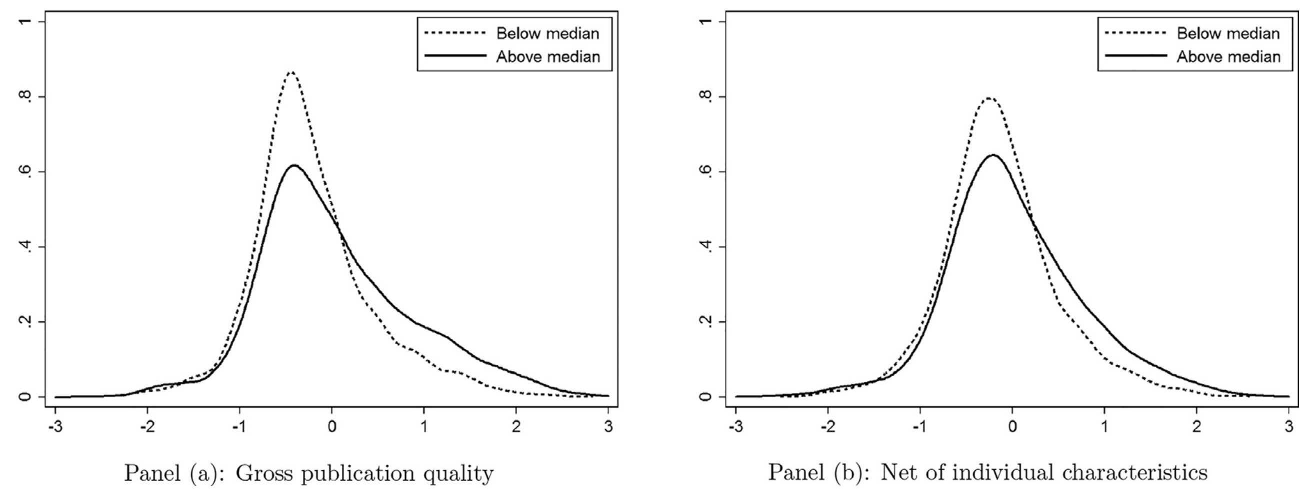

Fig. 1. Distribution of the (detrended logarithm of) individual publication quality in departments above and below median total number of publications.
Notes: Panel (a): Net of time and field fixed effects only. Panel (b): Net of time and field fixed effects and observed individual characteristics (gender, age and age squared, position, number of authors per publication, overall field diversity, co-authors located abroad). Publication measures and individual characteristics are defined in Section 2.

include both standard agglomeration economies due to specialisation, size and proximity to other departments, and the composition of local peers and their connections to foreign co-authors? Using an exhaustive panel of French academics in economics, over 19 years (1990–2008), including their quality-adjusted publication records in EconLit[4] and the location of the French economics departments employing them, we find that both individual skills and location matter for publications.

One the one hand, this contrasts with a few recent papers that consider a subset of the effects identified here. For instance, Waldinger (2012) concludes for Germany that there were no localised peer effects among physicists, mathematicians and chemists under the Nazi regime. Somewhat similarly, Kim et al. (2009) conclude that affiliation to one of the top 25 US universities in the 1990s, unlike in the 1970s and 1980s, no longer had an effect on individual academic outcomes in economics and finance. This is confirmed for mathematics all over the world by Dubois et al. (2014), who show that the best departments do not necessarily generate positive externalities even if they are the most successful at hiring the most promising academics. Oyer (2006) shows that top placements for new PhD graduate economists have long-term benefits for their careers, but no benefits related to enhanced productivity (in the 1990s). Our somewhat discordant conclusion might be explained by either the different context under study, which would mean that European institutions currently generate more local externalities than modern-day US universities, or German universities under the Nazi regime, or by the fact that our data set allows us to consider more local effects and to develop a more complete econometric strategy.

On the other hand, our finding clearly matches the agglomeration effects literature, which concludes that gains from spatial concentration exist in market activities even if individual characteristics and spatial sorting explain much of productivity differentials. This is illustrated for instance by Combes et al. (2008a) who use a reduced form approach similar to the one considered here or by Baum-Snow and Pavan (2012) in a structural approach. The conclusion is also in line with the local peer effect literature, which emphasises a significant (although not always large) role of peers and networks, either on labour markets (see recent examples by Damm, 2014; Hellerstein et al., 2014), at school (e.g. Lefgren, 2004; Kang, 2007; Lavy et al., 2012) or in criminal activities (e.g. Zenou, 2003; Bayer et al., 2009). This is also consistent with the role of proximity found for innovative activities. Indeed, studying academic publications also bears the advantage, compared to general labour market outcomes for instance, to better isolate a specific agglomeration mechanism, innovation and knowledge spillovers, which the literature usually does by using data on patents and innovation, as surveyed by Carlino and Kerr (2015).

Fig. 1, inspired by Combes et al. (2012), shows both the higher quality of the publications of academics located in departments that produce a larger number of publications, and the fact that individual observed characteristics explain only part of this difference. In panel (a) in Fig. 1, the distribution of (the detrended logarithm of) individual average publication quality in a given field (for precise definitions see Section 2) is plotted for two groups of academics from departments above and below the median for total number of publications. Clearly, the former distribution is shifted and dilated, to the right of the latter. Academics in departments with more publications have higher average publications quality. This can be seen at any point in the distribution (the shift) and is especially obvious for higher quality (the dilation). Interestingly, in panel (b), which uses the same department grouping, the conclusion still emerges when individual average publication quality is netted out of the role of some individual observed characteristics. However, it holds to a lesser extent due to the positive sorting of academics with better characteristics into better departments. More generally, we show that, when not controlling for individual fixed effects, location explains as much as do observed individual characteristics. When controlling for individual fixed effects, location still represents at least a quarter of the explanatory power of all individual characteristics.

Beyond the respective roles of individual and local effects, it is crucial for the optimal policy design to shed some light on the mechanisms underlying local effects. This goal is shared for instance by Hellerstein et al. (2014) who try to assess whether neighbourhood effects on labour markets are stronger between or within groups, in terms of race or ethnicity, by Lavy et al. (2012) who evaluate the impact of the presence of low-ability peers in classrooms on teachers' pedagogical practice and the quality of inter-student and student-teacher relationships, or by Agrawal et al. (2008) who study the relative role of spatial and social proximity for knowledge flows. In urban economics, the initial focus on

[4] EconLit is the electronic bibliography of the American Economic Association (see http://www.aeaweb.org/econlit/index.php). It is one of the largest publication data sets, listing more than 560,000 articles published between 1969 and 2008 in more than 1200 journals.

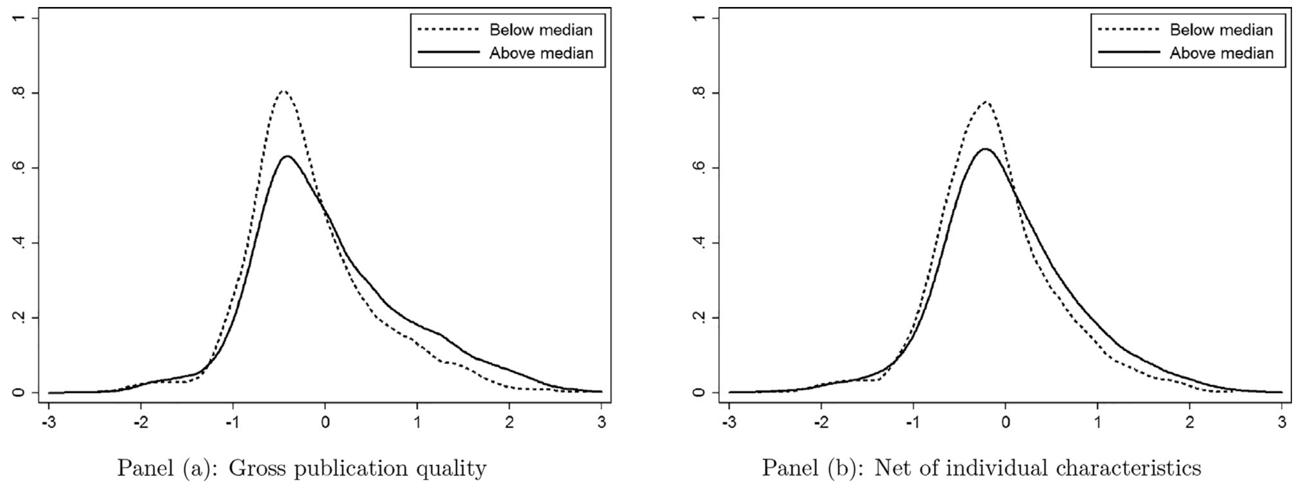

Fig. 2. Distribution of the (detrended logarithm of) individual publication quality in departments above and below median field presence.
Notes: Panel (a): Net of time and field fixed effects only. Panel (b): Net of time and field fixed effects and observed individual characteristics (gender, age and its square, position, number of authors per publication, overall field diversity, co-authors located abroad). Publication measures and individual characteristics are defined in Section 2.

city size and specialisation has been extended to a large number of city characteristics (see Combes and Gobillon, 2015). We try to open the black box of department effects and assess the relative magnitude of the channels through which these effects operate. However, standard urban economics variables usually correlated with the strength of local externalities have little explanatory power when controlling for individual fixed effects. This holds for both the variables capturing within-field externalities ('localisation economies') and overall department observed characteristics ('urbanisation economies'). This also holds for the local composition of peers and the network effects that we consider.

However, some department variables exert a significant impact and, especially, those related to localisation economies. Being specialised in a field significantly and largely increases the quantity and quality of the publications in this field even when controlling for individual fixed effects. This is illustrated by Fig. 2, which plots the distribution of individual average publication quality in a given field as in Fig. 1, but now for departments with field presence (presence of other academics publishing in the same field) above (respectively below) the median. The right shift and dilation are present even when individual characteristics are controlled for in panel (b). More generally, we show that the mere presence in the department of other academics publishing in the field generates large positive externalities, even when controlling for individual fixed effects. This increases the average quantity and quality of other academics' publications in this field by 40%. This effect is reinforced if the department's field share increases. For instance, the number of an academic's publications in a field increases by 6% if other academics' share of publications in that field doubles. In contrast, department size, proximity to other large departments, homogeneity in terms of publication performance, presence of colleagues with connections abroad, and composition in terms of positions and age matter for some publication measures but only if not controlling for individual fixed effects.

We also study the role of a number of time-varying individual characteristics. Controlling for department characteristics, publishing with a high number of co-authors per paper reduces the number of publications per individual but is positively correlated with higher publication quality, which suggests the presence of increasing returns to scale at the co-author team level. Having two co-authors rather than one for instance increases the average quality by between 8% and 25%. The average quality of an academic's publications increases also with the number of his/her publications, suggesting the presence of increasing returns to scale also at the individual level. Consistently with the literature on network effects, we also find that being connected to foreign co-authors increases both the quantity and quality of publications.

Concern over possibly endogenous location choices and spatial sorting of talents that might influence the measurement of department effects are taken seriously. We cannot use a natural experiment to remove endogenous selection to departments, as in Waldinger (2012) who uses the dismissal of scientists in Nazi Germany, or Azoulay et al. (2010) who employs premature death among superstar academics. The inflow of Soviet mathematicians to the US after 1992 is employed in Borjas and Doran (2012), who show that Soviet mathematicians substituted mainly for local mathematicians, whose publications fell sharply while overall publications slightly increased. However, they do not consider the effect of location within the US, and of departments' characteristics. Natural experiments are also widely used in the peer effect literature, as for instance in Kang (2007) and Damm (2014) who use quasi-random assignments of peers to individual students in middle schools of South Korea and of refugee immigrants to Danish municipalities.

Still, availability of an individual panel allows us to estimate specifications that consider both individual and department variables and both individual and department-time fixed effects. Therefore, local effects are net of possible academic spatial sorting, whether based on time-varying observed or time-constant unobserved individual effects, which corresponds to a fairly general setting. The French context we use helps also to reduce selection concerns. While initial affiliation, which is captured by the individual fixed effect, certainly is related to individual characteristics in France, most subsequent moves clearly are driven not by publication performance, but rather by friendship connections or personal/family reasons. This is due to the features of the French academic system. For instance, moving does not affect salary since academics are civil servants, who receive the same remuneration in any department even if there might be non-monetary benefits of being in a better department, as higher social status, access to better students, more interesting colleagues, etc. In addition, the most frequent transition from assistant professor to full professor, the largest source of movements in France, involves success in the 'Agrégation' contest, following which spatial allocation is largely random for most candidates (see Bosquet et al., 2014, for more details). Thus, although our experiment is not completely random, we do not think that individual time-varying publication shocks conditional on individual and department-time fixed effects are affecting location choices and biasing our evaluation of local effects to any great extent. Also, not using a natural experiment has the advantage that the results obtained are more general and, thus, have greater external validity. For instance, the characteristics of the coauthors of superstars, or the scientists dismissed by the Nazis, may differ from the characteristics of the average current academic.[5]

In addition to the endogeneity of individual location choices, reverse causality biases can also affect the role of department characteristics. For instance, Combes et al. (2010) show that the impact of city size in market activities decreases by up to 20% when local variables are instrumented. Given the large number of department characteristics we consider here, it would be difficult (and would not make much sense) to instrument all of them and, in addition, would introduce possible weak instrument issues. We are not aware of other studies of agglomeration and peer effects in academia that propose instrumentation of department characteristics. We leave this issue to future contributions.

As Combes and Gobillon (2015) detail, simultaneous identification of individual and location-time fixed effects is demanding and requires sufficient mobility of individuals between locations. Exact identification conditions are difficult to check empirically and, in the literature, this is never attempted in practice. The mobility of French academic economists across departments is not very high and the sample size is much lower than that of standard employer-employee data sets. Conversely, the affiliation of some academics to more than one department at the same time increases identification power. Also, information on age, gender, position, publication fields and author connections, may make it less important to control for individual fixed effects at the cost of considering possibly endogenous individual control variables. Overall, it is difficult to assess whether individual and local effects are well identified here. We present the two sets of estimations, with and without individual fixed effects, and comment if the conclusions differ between the two. A slight positive sorting of academics with the best observed characteristics into departments that generate higher positive externalities is observed if individual fixed effects are not considered. When considering individual fixed effects, sorting on unobservables is slightly negative, which, in the literature, is considered a possible sign of lack of identification power (see Abowd et al., 2004; Andrews et al., 2012). Therefore, it is not possible to be sure that the model with both individual and department-time fixed effects is correctly identified. This concern becomes more evident when we observe that removing 15% of the academics located in the largest departments (alternatively those among the 15% most publishing academics), who, thus, contribute the most to identification, reinforces the presence of negative sorting. Conversely, our conclusions about the relative importance of individual and department effects for publication, and the positive role of department's specialisation, are robust to these checks.

Finally, we decompose individual productivity into three components: the probability to publish in a given period, the number of publications, and the average quality of those publications. We study the determinants of these three dimensions separately, at the detailed sub-field level within economics. Most previous works consider broader fields with only the quality adjusted number of publications as the dependent variable, which could blur the respective effects of publication quantity and quality. We show that the effect of some variables differs from one productivity dimension to another, which means that the optimal strategy for an individual or a department depends on the dimension targeted. In addition, we perform our estimations on two different, more or less selective indexes of publication quality. Typically, the (individual and department) determinants of publication in top journals might differ from the determinants of publication in field journals. Our use of a quality index for all 1200 EconLit journals enables us to study such differences, whereas most of the studies in the literature focus on a small number of journals: 23 in Waldinger (2012), 41 in Kim et al. (2009), 98 in Dubois et al. (2014).

Our data set of all academics in economics located in France is both exhaustive and provides the other important advantage that it includes non-publishing academics. Studies based only on bibliometric sources necessarily exclude this group. This means that department characteristics, computed only on publishers, can be affected by potentially large measurement error since non-publishing academics can account for 30% of a department's membership (see Combes and Linnemer, 2003, for both European and US departments). That is, department size is based not on the actual number of academics used here, but the number of academics in the department who published over the period analysed (which is usually a short time period). Last, we have information on more individual characteristics, such as age, gender and academic position, which might affect publication output and, usually, are not considered by the data sets used in other studies. All of these aspects could influence the results obtained and may explain our new findings, although we acknowledge that these are estimated only on French economics departments.

Section 2 presents the data and econometric strategy. The results for the relative roles of individual characteristics and department effects for individual output are presented in Sections 3 and 4, and the results for channels of department effects are presented in Section 5. Section 6 concludes.

2. Data and estimation strategy

2.1. Academics, departments, publications

The French Ministry of Higher Education and Research, CNRS and INRA[6] provided us with lists of academics in economics in France during the period 1990 to 2008. Each academic is affiliated to at least one university department or to a CNRS or INRA research centre. In this study, 'department' refers either to the only affiliation of the economist in a university (the majority of cases) or the aggregation of all the university departments or research centres that include economists. We believe that this aggregation better matches the French reality of academic research compared to using detailed affiliations.

The French system allows for multiple affiliations and 9% of academics are affiliated to two or three departments. In those cases, each department is weighted equally. For the few cases of academics with positions in both France and abroad, we use their CVs to evaluate the share attributable to the French department. Finally, we want the analysis to focus only on academics who can really be considered as forming a local group of academics who work together. Therefore, we retained only those departments with

[5] Another strategy would first specify a model for the academic's department choice and then estimate our empirical model conditional on that choice. However, this requires exclusion restrictions to be satisfied, i.e. variables that explain department choice, but not publication record. We cannot envisage suitable variables since even family characteristics might explain publication record. Alternatively, some authors, such as Gould (2007), Baum-Snow and Pavan (2012) or Beaudry et al. (2014) for instance, have proposed structural estimations where the full set of local opportunities is specified for any individual, and optimal inter-temporal location choices are made. We chose not to follow this suggestion which imposes a great deal of structure on the estimation.

[6] Respectively, Ministère de l'Enseignement Supérieur et de la Recherche - Direction Générale de la Recherche et de l'Innovation, Centre National de la Recherche Scientifique, and Institut National de la Recherche Agronomique.

at least four full-time equivalent academics and excluding isolated economists in universities with no formal economics department.[7]

The data set includes a number of individual characteristics such as gender, age and position. We merge this with EconLit by last name and first initial of the first name because the recording of full first names in EconLit is not sufficiently systematic. Using the first initial increases the number of academics with identical names but very slightly. Their publication records were dealt with manually. For all academics and for each year between 1990 and 2008, we have data on academics' individual characteristics, departments of affiliation and publication record, in economic journals only, which excludes even top journals such as *Nature* and *Science*.

We measure the publications output of academics in field f at date t as the weighted sum of their publications in field f listed in EconLit over the period τ. In most cases, τ corresponds to years $t+1,\ t+2,\ t+3$ and output at date t is a moving average over these three years. This choice is in line with the literature and was adopted recently by Ductor et al. (2014). It seems to be a reasonable estimation of the time needed to write a paper and the publication delays.[8]

Each publication is weighted first by the quality of the journal in which it is published. We use two of the Combes and Linnemer (2010) journal weighting schemes, a very selective one, which we describe as 'Top quality', and a less selective one, which we call 'Quality'. Second, in line with common practice in the literature, each publication is weighted also by the inverse of its number of authors. Third, the output measure takes account also of the article's relative number of pages (within a journal). Last, output is measured at field level by attributing $1/J$ of the corresponding publication output to each of the J Jel codes (aggregated in 18 different categories) mentioned in the publication.

Finally, instead of studying only the determinants of the number of an academic's quality-adjusted publications, we decompose this into three variables whose determinants are studied successively. These are: the probability to publish in a given field over the period, the number of publications in the field over the period, and the average quality (or top quality) of the publications over the period. Appendix A provides more details on these measures.

2.2. Econometric specifications

To separate agglomeration effects from the role played by individual characteristics, we follow the econometric strategy proposed in Combes et al. (2008a), whose advantages are discussed in Combes and Gobillon (2015). This strategy comprises a two-step procedure. In the first step, the logarithm of academic i output in field f at date t, y_{ift} is regressed on individual characteristics that vary or not over time (and possibly an individual fixed effect), a department-time fixed effect (β_{dt}) and the department characteristics that depend on the field:

$$\begin{aligned}\log y_{ift} = {} & \theta_i + \text{Individual's Characteristics}_{it}\,\varphi \\ & + \text{Department's Field-Specific Characteristics}_{dft}\,\eta \\ & + \beta_{dt} + \mu_{ft} + \varepsilon_{ift}, \qquad (1)\end{aligned}$$

where θ_i and μ_{ft} are individual and field-time fixed effects, respectively, and ε_{ift} is an individual random productivity component assumed to be independent and identically distributed (iid) across individuals, fields and periods.[9]

The first step allows us to evaluate the respective explanatory power of individual characteristics, department field-specific characteristics, and department-time fixed effects. These last ones capture not only the department's observed overall characteristics but also any unobserved local effect. The second-step estimation allows us to identify the separate roles of each department's overall characteristic on the estimated department-time fixed effect, net of individual effects, $\hat{\beta}_{dt}$:

$$\begin{aligned}\hat{\beta}_{dt} = {} & \text{Department's Overall Characteristics}_{dt}\,\gamma \\ & + \delta_t + \upsilon_{dt}, \qquad (2)\end{aligned}$$

where δ_t is a time fixed effect and υ_{dt} is a random department level component, which is assumed to be iid across departments and periods. Since the dependent variable in the second-step is affected by measurement error due to its estimation in the first step, we use Feasible Generalised Least Squares (FGLS) in the second step. We assume that the specifications (1) and (2) hold for each component of the individual's publication record: the probability to publish, the number of publications and the publication's average quality (or top quality).

Finally, the first-step estimations need to weight the individual observations for two reasons. First, an academic can belong to more than one department. For each academic, date and field, we have as many observations as the academic's number of affiliations, and each has a weight α_{idt}, which is the share attributed to each affiliation. Second, since each publication is split between its Jel codes, academics may publish in many fields each year. To take account of both of these effects, the first-step estimations are weighted by α_{idt} for the probability to publish, and by $\alpha_{idt} S_{ift}$ for the publication quantity and quality, where S_{ift} is the share of field f in academic i's output at date t. This means that in all the regressions each academic receives the same weight. Ordinary Least Squares (OLS) estimations are employed for the probability to publish due to the presence of many fixed effects. We checked that a probit estimation leads to similar results.

2.3. Department characteristics

'Department's Overall Characteristics' include a first set of variables for the department's demographic structure: the logarithm of department size, measured by its number of full-time equivalent academics (referred to in the tables as *'Size'*), the share of women in the department ('*% women*'), the average age of the academics ('*Average a*ge'), the share of upper positions (full professor rather than assistant professor, for instance, described as '*% rank A*'), and the share of a number of specific academic positions ('*Positions*' shares').

Department size corresponds to the variable for total employment density in standard estimations of agglomeration economies.

[7] We conducted robustness checks using detailed affiliations (i.e. not aggregated affiliations within each university) and on departments with more than 9 full-time equivalent academics. The results are fully consistent with present ones, as presented in Bosquet and Combes (2017).

[8] As a robustness check and because this choice is both somewhat arbitrary and could result in some autocorrelation of residuals, in Bosquet and Combes (2017) we present the regressions with τ reduced to year $t+2$. The results are very similar to those obtained using a three-year moving average.

[9] Introducing field-time fixed effects responds to an interpretation issue. If it is assumed that differences in publication records between fields at the France level in a given year are a matter of "fashion" rather than talent and a real difference in productivity among academics and departments specialised in different fields, then the differences at the France level should be removed by the introduction of field-time fixed effects, allowing a focus on spatial variability independent of specialisation choices. Conversely, if it is assumed that a higher number of publications in a field at the France level in a given year genuinely corresponds to higher productivity in that field, then field-time fixed effects should not be introduced into the specification. Here, we adopt the former assumption and introduce field-time fixed effects. This is the more conservative strategy since it removes the role of possible correlation between local effects and field specialisation choices. It is also the assumption adopted in urban economics, which considers systematic industry fixed effects in wage or productivity equations. It estimates local effects once direct composition effects are removed. Importantly, this does not prevent us from identifying the local externality role of the department's field-specific characteristics.

It reflects possible local externalities emerging from the overall size of the local economy. The list of possible positive effects from department size is long, but includes, for example, the fact that academics in larger departments may benefit from larger and more intense scientific interactions with their peers, from a larger (more numerous) administrative and/or research assistance staff, or shared computing facilities, from greater bargaining power within the university or greater likelihood at the national level to receive more research funds, or from greater overall visibility which reinforces external network effects. We cannot exclude the possibility of congestion effects causing size also to generate some negative effects. In line with most of this literature, only the total net effect of size is identified, which is the case for most of the local variables in these settings.

For a given department size, departments may have younger or older academics, more or less women, or a higher ratio of full professors to assistant professors, for instance. As Hellerstein et al. (1999) suggest, these composition effects are introduced into the specification as their proportions in the department's total number of academics. This allows us to assess whether different types of academics generate stronger local externalities.[10] For instance, older academics may contribute their experience, women may generate more externalities than men, and similarly for the different types of positions. A strand of work in the industrial organisation literature (see, for instance, Besley and Ghatak, 2008; Auriol et al., 2012) studies the role of status incentives and the implications for the optimal shares of the different positions within firms, which is another interpretation of these variables. The French academic system is rather complex in terms of possible academic status. We distinguish first between lower and upper (Rank A) positions (assistant professor versus full professor). Also, some academics have teaching obligations while others do not, some are attached to the local university while others depend on national research institutes (CNRS, INRA, EHESS, etc.), and a few are linked to domains other than economics (business, mathematics, etc.). Each type of position can generate more or fewer externalities since both the time devoted to research and the incentives to cooperate locally will differ. We distinguish among ten different positions.

In order to identify different channels of department externalities, we consider a second set of variables, which are more related to the research characteristics of the department. First, we evaluate the role of 'Department's Field-Specific Characteristics'. Marshall (1890) initially proposed the idea that the relative size of an industry within the local economy can generate stronger local externalities for that industry, for instance, if it uses specific local public goods, or specific inputs or labour types, which urban economics describe as 'localisation economies'. The same intuition can be developed for an academic research field, for instance, because not all fields within economics are internationalised to the same extent, or because they do not need the same research mix in terms of research assistance, computing facilities, or access to data. Benefiting from a publications measure at the field level allows us to test whether academics in departments specialised in a particular field publish more in that field.

Since many fields are absent from many departments, we consider a non-linear specification for localisation effects, using two variables. First, we consider a dummy variable ('*Dep. field presence*') that takes the value 1 if at least one academic in the department other than the one considered has published once in the field. Second, we use a specialisation variable ('*Dep. specialisation*') - the share of department d's output in field f at date t (other than the academic's output) - to assess the role of the field's relative size in the department. Importantly, because the academic's field choice in a given year may not be an accurate reflection of the average field of specialisation, we compute the individual variables that rely on field over a longer time span than the one used for the dependent variable. All publications until τ are taken into account, but the more recent ones count more than the older ones (see Appendix A for details). Then, the department-field level explanatory variables are based on publication output at the field level for all academics affiliated to the department at date t.

Jacobs (1969) popularised the idea that the overall diversity of the local activity can be beneficial for local productivity, especially in research-intensive sectors. According to this viewpoint, diversity encourages the cross-fertilisation of ideas between industries, which strengthens innovation and growth. There is a large literature which tests this idea by introducing a diversity index into the estimated specifications. This, typically, is a Herfindahl index on the shares of each industry in the local economy. We adopt a similar procedure using the shares of each Jel code in the department's publications to obtain the department's overall field diversity ('*Diversity*').

Beyond department size, the physical proximity to other departments with which academics might either interact or compete, can matter. We capture this effect using an external research access variable ('*Research access*'), which is the spatially-discounted sum of the sizes of all other departments. It provides information on whether externalities emerge between different, but proximate departments, as highlighted by urban economics over the last twenty years for the case of market activities.

Whether or not hiring top academics is a good strategy and benefits the other academics in the department is a debated question. We test, more generally, whether the department's heterogeneity in terms of its academics' publication records, measured by the within-department coefficient of variation of individual output, has an impact on individual publication output ('*Heterogeneity*').

Departments also may differ in terms of the academics' patterns of co-authorship. Having academics connected to foreign academic institutions can generate positive externalities via network effects, for instance, an aspect that has been emphasised in the case of both market activities and research (see Ductor et al., 2014, for a recent example in economics). We compute the share of the department's academics connected to (at least one) co-author located outside of France, but not in the USA ('*Non-USA connections*'), and the share of the department's academics connected to co-authors located in the USA ('*USA connections*').

We next describe the individual variables. It is crucial to control for individual characteristics in order not to attribute a simple sorting of different academics in different departments to an externality effect. The data set we use allows us to identify both the impact of individual characteristics and, therefore, to control for the possible non-random selection of academics across departments, and the externality impact of those characteristics, simultaneously. For instance, older academics might publish less individually, but exert a positive externality on the other academics in the department. Therefore, we consider the role, at the individual level, of all the variables for which a possible department level externality is tested. This includes academics' age (and its square), gender, position held and rank A status, and dummy variables for connection to at least one co-author abroad, but not in the USA, and connection to a co-author in the USA.

We also include variables that reflect an academic's individual research characteristics. To test for the presence of economies of scale within co-author teams, we introduce academics' average number of authors per publication ('*Authors per publication*'). This variable is central in many studies of the determinants of publication records that ignore the role of location, but evaluate the returns from co-authorship, following Sauer (1988). We also consider

[10] As Ciccone and Peri (2006) emphasise, only a combination of the externality and of some possibly negative substitution effects is identified.

academics' field diversity ('*Individual diversity*'), to assess whether academics benefit from knowledge acquired in other fields, to publish in a given field. This tests the presence of complementarities between fields at the individual level. The variable is the logarithm of the number of fields in which the academic has published.

Finally, some estimations also consider individual fixed effects, which capture the role of any individual characteristic that is constant over time. The definitions of all the variables are provided in Appendix A.

2.4. Samples and mobility

Using a three-year moving average for publication output prevents us from considering years 2006, 2007 and 2008. Therefore, the time observations in our data set, span 1990 to 2005. Both the number of academics and the number of departments are monotonically increasing over time, from 1753 academics and 69 departments in 1990, to 2914 academics and 81 departments in 2005. Over the 16 years of our panel, this leads to 39,266 academic-year observations and 1267 department-year observations, 1208 with at least one academic who publishes.

The regressions are performed at the field level. Since there are 18 possible fields, the 39,266 academic-year observations translate into 706,788 academic-field-year observations, which then is increased because some observations are duplicated in the case of academics with multiple affiliations. As a result, the number of academic-field-year observations in our first-step estimation for the probability to publish is 770,202. This reduces to 758,790 when department-year fixed effects are considered because some department-years have no one that publishes. This reduces further to 424,044 when we include individual fixed effects because some academics have no publications. Finally, because academics do not publish in all fields, many of these observations correspond to zero outcomes in a given field. There are 'only' 38,836 non-zero academic-field-year observations, which are the observations used for the first-step quantity and quality estimations for which we take logs. In the second step, we have 1208 department-year observations.

Appendix Table B.1 presents some descriptive statistics for individuals and departments. Overall, there is substantial variation in the data in relation to individual publication output and department characteristics. As emphasised in the introduction, identification relies on the mobility of some academics between departments, and on the fact that some have multiple department affiliation at the same date. On average, each year, 2.1% of academics move to another department, and 14.8% have moved at least once over the whole period. 9.0% of the individual-year observations correspond to multi-affiliated academics, and 18.6% of the academics have been multi-affiliated at least once. Mobility rates between cities in comparable estimations of agglomeration effects in market activities, in general, also are around 2% per year. There would appear to be room for mobility to be sufficiently high to allow for the identification of both individual and department-time fixed effects. However, what matters is the number observations relative to the number of fixed effects to be identified, which is higher when matched employer-employee wage or productivity data are used. Overall, these figures, on their own, make it difficult to assess whether or not our estimations are correctly identified. We return to this issue in the discussion of the results.

3. Productive academics: Individual abilities versus department effects

This section studies the determinants of individual productivity and assesses the relative weights of individual and department effects. We regress individual productivity in a specific field on individual characteristics related to both individual abilities and individual research features (including field-time fixed effects), department field-specific variables (field presence and specialisation) and department-time fixed effects. Table 1 Columns (1) and (2) ('Publishing') refer to a linear probability model where the dependent variable is 1 if academic i produces in field f at date t and 0 otherwise. Columns (3) and (4) ('Quantity') concern the log of the number of publications, and Columns (5) and (6) ('Quality') and Columns (7) and (8) ('Top quality') regard the log of the average publication quality using the standard and top journal quality indexes respectively. For each output measure, the first column excludes individual fixed effects, which are included in second column.[11]

Before turning to the effect of each variable, we start with some variance analysis. A first remark regards the large increase in the R^2, of 17% for the probability to publish, of 22% for the number of publications, of 32% for publication quality and of 28% for top quality when department effects are considered, compared to the estimations (not reported here) that do not include department characteristics in the specification (and when individual fixed effects are not included in either case). The explanatory power of the model increases even more if individual fixed effects are introduced. It becomes 50% to 80% higher compared to if only individual variables and department effects are included, and reaches levels that are comparable although slightly lower, than those obtained for the standard individual wage or productivity equations, with R^2 between 0.54 for quantity and 0.72 for top quality.[12] A last conclusion is that the model explains the average quality better than the number of publications and, especially, if we consider the index for top quality. This would seem to make sense from an academic perspective, since publication in a top journal requires more specific skills, which are captured by the model, than just publishing in general.

Obtaining more precise insights into the sources of output variation requires a more detailed variance analysis. It is provided in Table 2 for the determinants of individual publication quality, without (left panel) and with (right panel) individual fixed effects. First, the 'Std. dev.' columns report the standard deviation of the effect of a variable or a group of variables for the quality estimations presented in Table 1 columns (5) and (6) respectively. The higher this standard deviation relative to the standard deviation of the dependent variable to be explained (reported in the first line), the larger the explanatory power of this variable or group of variables. However, a variable or group of variables has a large explanatory power if its effect is largely correlated with the dependent variable. This is reported in the 'Corr.' columns.[13]

[11] The effects of women and age are not displayed in columns (2), (4), (6), and (8) since it is not possible to identify them separately from individual and year fixed effects. There are some variations for the rank A and position variables within individuals, but, since they are rare, we prefer also to remove these variables when considering individual fixed effects. The results are hardly affected if they are included.

[12] The R^2 is lower for the probability to publish, but we do not have a benchmark value for this case, and it obviously arises, at least in part, from the large number of zero observations.

[13] Imagine a model with only two explanatory variables, $y_i = \alpha x_i + \beta z_i + \varepsilon_i$. The first row in the table would report in column 'Std. dev.' the variance of y_i, to be explained. The second row in Column 'Std. dev.' would report the variance of αx_i, i.e., of the effect of x_i on y_i, all else being equal. The 'Corr.' column would report the correlation between y_i and again αx_i. Considering the effect of x_i, αx_i, and not x_i only, takes account of whether the impact of x_i is large, with everything else controlled for or not (assuming that α is deterministic). This provides a way to assess the relative explanatory power of the variable. This can be applied nicely to a group of variables. For instance, the variance of $\alpha x_i + \beta z_i$ and then its correlation with y_i would be reported. See Abowd et al. (1999) for more details on this type of variance analysis. The last column, 'Sorting', reports the correlation between the effect of a variable or group of variables, with all department effects included (hence,

Table 1
Determinants of individual publications.

	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Individual characteristics								
Women	−0.016[a]		−0.119[a]		−0.067[a]		−0.270[a]	
	(0.001)		(0.009)		(0.009)		(0.024)	
Age	−0.005[a]		−0.035[a]		−0.022[a]		−0.093[a]	
	(0.000)		(0.003)		(0.003)		(0.009)	
Age square	0.000[a]	−0.000[a]	0.000[a]	−0.000	0.000[b]	−0.000[a]	0.000[a]	−0.001[a]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rank A	0.044[a]		0.218[a]		0.136[a]		0.542[a]	
	(0.001)		(0.009)		(0.009)		(0.024)	
Authors per publication			−0.948[a]	−0.925[a]	0.186[a]	0.192[a]	0.508[a]	0.539[a]
			(0.011)	(0.015)	(0.010)	(0.013)	(0.027)	(0.034)
Individual diversity			−0.096[a]	−0.130[a]	0.013[c]	0.003	0.109[a]	0.024
			(0.007)	(0.009)	(0.007)	(0.008)	(0.018)	(0.020)
Non-USA connection			0.376[a]	0.193[a]	0.307[a]	0.084[a]	1.128[a]	0.342[a]
			(0.011)	(0.014)	(0.011)	(0.012)	(0.029)	(0.031)
USA connection			0.408[a]	0.223[a]	0.509[a]	0.209[a]	1.604[a]	0.611[a]
			(0.015)	(0.019)	(0.014)	(0.016)	(0.038)	(0.042)
Dep.-field characteristics								
Dep. field presence	0.063[a]	0.122[a]	0.345[a]	0.334[a]	0.114[a]	0.087[a]	0.359[a]	0.318[a]
	(0.001)	(0.002)	(0.022)	(0.022)	(0.021)	(0.019)	(0.055)	(0.048)
Dep. specialisation	0.014[a]	0.024[a]	0.098[a]	0.088[a]	0.036[a]	0.020[a]	0.132[a]	0.084[a]
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.003)	(0.009)	(0.008)
Fixed effects								
Field-time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department-time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position	Yes	No	Yes	No	Yes	No	Yes	No
Individual	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.07	0.13	0.33	0.54	0.37	0.65	0.46	0.72
Observations	758,790	424,044	38,836	38,836	38,836	38,836	38,836	38,836

Notes: Standard error between brackets. [a, b, c]: significant at the 1%, 5% and 10% levels respectively. OLS estimates. Variables are defined in Section 2.2.

Table 2
Variance analysis of the individual publication quality.

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Corr.	Sorting	Std. dev.	Corr.	Sorting
Explained: Quality	0.453	1.000		0.453	1.000	
Individual effects	0.159	0.452	0.193	0.340	0.753	−0.078
Indiv. fixed effect	-	-	-	0.331	0.636	−0.097
Obs. indiv. effects	0.159	0.452	0.193	0.178	0.254	0.031
Women	0.015	0.048	0.017	-	-	-
Age	0.082	0.132	−0.026	0.162	0.134	0.016
Position	0.046	0.151	0.060	-	-	-
Rank A	0.038	0.102	0.136	-	-	-
Authors per pub.	0.040	0.215	0.092	0.041	0.215	0.038
Individual diversity	0.004	0.141	0.120	0.001	0.141	0.032
Non-USA connection	0.061	0.276	0.123	0.017	0.276	0.008
USA connection	0.075	0.317	0.175	0.031	0.317	0.039
Department effects	0.154	0.434	1.000	0.111	0.189	1.000
Dep.-time fixed eff.	0.152	0.418	0.988	0.110	0.173	0.992
Dep.-field-time eff.	0.024	0.134	0.169	0.014	0.142	0.133
Field presence	0.016	0.066	0.116	0.012	0.066	0.033
Specialisation	0.027	0.081	0.082	0.015	0.081	0.099
Field-time fixed effect	0.091	0.302	0.121	0.063	0.257	0.024
Residuals	0.360	0.795	0	0.268	0.593	0

Notes: The table presents the variance analysis of the estimations reported in Table 1 columns (5) and (6). First, all variables are centred with respect to their annual mean. Therefore, all the variables are detrended and the variance analysis is performed along the within-time dimension. The “Individual effects” row corresponds to the simultaneous roles of individual observed and fixed effects if any. The “Obs. indiv. effects” row corresponds to the role of all the observed individual effects taken together. “Department effects” corresponds to the role of both department-time fixed effect and all department-field-time effects. “Dep.-field-time eff.” reports the simultaneous roles of “Field presence” and “Specialisation’. See footnote 13 for details of what is reported in each column.

The variance analyses, considering or not individual fixed effects, respectively the right- and left-hand sides of the table, differ greatly. Without individual fixed effects, the explanatory power of individual and department effects is very similar in terms of both the standard deviation of the effects (each around one-third of the standard deviation of the dependent variable) and the correlation to the dependent variable (slightly less than 0.5). The two groups of variables contribute to the same extent to explaining individual publication output.

In contrast, individual effects have much larger explanatory power if individual fixed effects are considered. The right-hand side of Table 2 shows that the standard deviation of individual effects is two-thirds that of the dependent variable, and is three times bigger than the one of department effects. The correlation to the dependent variable is four times larger for individual effects. This means that some unobserved individual effects are significantly influencing publications quality and, if individual fixed effects are not included in the specification, part of this is captured by the department-time fixed effects. Appendix Tables C.1, C.2 and C.3 reproduce the variance analysis respectively for the probability to publish, publication quantity and top quality. For all the variables, the conclusions are similar.

To sum up and keeping in mind that, if there is insufficient mobility between departments – which we discuss further below, individual fixed effects cannot always be properly identified separately from department effects, the lower bound of the explanatory power of department effects is around a quarter of the explanatory power of individual effects. However, at the upper bound, without individual fixed effects, but still including a fairly large set of individual observable characteristics, department effects can explain as much as individual effects. This contrasts with the findings in the literature. Waldinger (2012) finds no peer effects among physicists, chemists and mathematicians in Nazi Germany. Similarly, Dubois et al. (2014) find no effects for modern-day mathematicians. Kim et al. (2009) find that the effect of being in the top 25 economics and finance departments gradually disappears between the 1970s and the 1990s in the USA. Note that these authors comment on the fact that department effects are or are not significant, but do not, as we do in this paper, discuss their overall explanatory power. Thus, the perspective is different and, in our view, our approach is relevant for assessing the share of individual productivity explained by local effects. Another possible explanation of the difference between these results is that individual or department fixed effects are not always properly identified. Unfortunately, whether mobility is sufficiently high to identify both individual and department fixed effects is difficult to test formally and is not attempted in any of the literature. Finally, it might be that research habits differ between a European country, such as France, and the USA, in terms of both research technology (e.g. intensity of internet use for collaborations) and institutional design. For instance, the possibility of academics to capture their publication performance is considered lower for most European countries where wages and positions are less closely tied to publication records (see Combes et al., 2008b, for France). All of these factors might be affecting the relative roles of individual and department effects.

Another important result emphasised by Combes et al. (2008a) is related to the sorting of workers across space. More able workers locate in more favourable locations, that is, those where the location effects reflecting local externalities are strongest. The "Sorting" columns in Table 2 for publication quality, and in Tables C.1, C.2 and C.3 for the probability to publish and the quantity and top quality of publications, report the correlation between the effect of a variable or group of variables and the overall department effects. Typically, it is positive for observed individual characteristics. Workers with individual observed characteristics that promote higher numbers of publications and higher quality, are located in departments that provide larger external effects.[14] When individual fixed effects are not controlled for, the correlation of all individual observed effects together with overall department effects, is large at 0.19. However, it falls to 0.03 when individual fixed effects are controlled for and we observe a negative correlation at −0.10 between individual fixed effects and overall department effects. Therefore, we find that the effect of spatial sorting of academics is quite large and positive on the observed characteristics, but that it decreases hugely if individual fixed effects are considered; the overall sorting turns negative (at −0.08) due to a more than compensating negative sorting on unobservable characteristics.

The sorting results for quality lie between those obtained for quantity and top quality, which are reported in Appendix C. In relation to quantity, there is almost no sorting on observed characteristics (correlation at +0.01 and −0.04 without and with fixed effects respectively) and a fairly large negative sorting for individual fixed effects (correlation at −0.16). For top quality, there is a positive sorting effect on observed characteristics (+0.24 and +0.04 without and with fixed effects respectively), thus, stronger than for quality, and a fairly small negative sorting for individual fixed effects (−0.04). Unfortunately, none of the papers assessing the magnitude of peer effects in science computes these correlations between individual and department effects, which would have allowed us to compare these important results with those from other fields or for other periods.

The presence of a negative sorting on unobserved academic characteristics would be a striking result. However, we cannot exclude that a negative sorting on unobserved individual characteristics is the result of weak separate identification of individual and department fixed effects. Further investigation renders this hypothesis more credible. We re-estimate the model after removing the departments and academics contributing, a priori, the most to identification, that is, first, the three largest departments, which account for around 15% of French academics, and second, the 15% most productive academics. In both cases, the negative correlation between individual fixed effects and department effects increases to −0.19 in the first case and to −0.41 in the second case, for instance, for the quantity of publications (see full results in Bosquet and Combes, 2017). As Andrews et al. (2012) emphasise, this is a sign of possible lack of identification power.

4. Role of individual characteristics

Before investigating in Section 5 the channels for local effects, we discuss the role of individual characteristics. The results in Table 1 suggest that women and older academics publish less. This is consistent with the findings in the literature and this is even more intuitive in our case since we control for type of position held. Once a given position is achieved, full professor for instance, the number and quality of publications decrease with age. Part of the effect might result also from a cohort effect (previous generations had less incentives to publish than younger academics).

As described in Bosquet and Combes (2017), the results for the impact of different positions are as expected. The higher the rank (professor, research professor and, especially, *INSEE* or *Ponts-*

the value 1.000 for the line 'Department effects'). This allows us to assess which variable (or group of variables) is the most strongly correlated to the department component of individual productivity.

[14] Age is negatively correlated to department fixed effects and has a negative effect on output, so, again, we find that academics with "good" characteristics (younger age), tend to locate in better quality departments. The women dummy is the only exception here, although the correlation to the dependent variable is small.

et-Chaussées engineers, compared to assistant professor or junior research fellow) and the more time allocated to research (research versus teaching positions), the higher the quantity and quality of publications. This applies also to academics in pure economics departments compared to those who work also in business or mathematics, for instance. Therefore, even if part of the promotion process in France does not depend on publications, as shown in Combes et al. (2008b), those who are appointed to more senior positions tend to publish more on average. Note that our aim here is not to give a causal interpretation for these variables, but to control for individual ability when estimating the role of departments.

Note, also, that we control not only for some of the standard "ability" variables considered in wage and productivity equations but also for variables characterising the academic's research. The variables showing the largest correlation to the dependent variable are the connection variables. Academic economists with more co-authors working abroad (both in the USA and elsewhere), also publish more and are generally over-represented in better departments. Again, the direction of the causalities cannot be determined here.

The average number of co-authors per publication also has a large explanatory power. Its impact on published quantity is mostly negative. Having more co-authors decreases the number of published papers, which means that attributing only part of the publication to each co-author corresponds to a stronger effect than the effect of producing more papers with more co-authors. In other words, the quantity published is subject to decreasing returns to scale in terms of the number of authors; academics would publish more papers were they to work alone. However, the average number of co-authors has a large positive effect on average publication quality, and this is almost three times larger for top quality. Therefore, a larger number of co-authors decreases the equivalent number of publications written alone, but increases their quality. Thus, there is a trade-off, which only an analysis such as that presented here can identify. For instance, according to the estimates controlling for individual fixed effects (although magnitudes are very similar without), an academic with an average of two co-authors rather than one (per publication) has 31.3% fewer publications, but their average quality is 8.1% higher and average top quality is 24.4% higher.[15] Combining these two effects, having two co-authors rather than one, decreases the quality-adjusted number of publications (a measure that is frequent in the literature and here is obtained by multiplying the quantity by the average quality) for both standard and top quality, although less in the case of top quality. Therefore, we should not expect a higher number of publications, even in good quality journals, based on co-authoring, but we can expect a higher possibility of publication in top journals. The article by Sauer (1988), which is one of the earliest contributions on the impact of co-authorship on publication, finds almost no effect, and two other studies, also on economists, Hollis (2001) and Medoff (2003), conclude that there is a negative effect of co-authorship on publication quality. Dubois et al. (2014) identify an overall negative effect of co-authorship among mathematicians on their citations-adjusted publication index, but the effect of collaboration with co-authors with different specialisation is positive. Ductor (2015) finds for economists a negative effect of co-authorship between 1970 and 2011, which turns positive if unobserved heterogeneity and endogenous co-authorship formation are accounted for. Overall, our results are more in line with studies that find a negative impact of co-authorship; however, ours is the only work that estimates the effect of co-authorship on the average quality of publication. We find that this is positive, the more the more selective the journal.

We find also that a higher diversity of research fields decreases the number of an academic's publications, but has no impact on their quality and, possibly, has a positive impact on top-quality publications, although this disappears if we control for individual fixed effects. Dubois et al. (2014) find a positive effect of field diversity for mathematicians.

Finally, in Bosquet and Combes (2017) we report estimations for the quality determinants that control for individual quantity, although endogeneity concerns might be more severe than for other variables. This allows us to test for the presence of increasing returns to scale on quality at the individual level (as opposed to the co-author team level, assessed before based on the number of co-authors). This has been overlooked in most previous work. We find increasing returns to number of publications for average quality and, even more so, for top quality. The more academics publish, the higher the average quality of their publications. An academic with twice as many publications has an average publication quality 6.6% higher, and a top publication quality 40.8% higher when not controlling for individual fixed effects.[16] It follows, also, that the impact of the number of co-authors per publication on quality is even stronger if controlling for quantity.

Notice that all these results hold within field since we control for Jel code fixed effects. Jel codes appear to have quite large explanatory power, especially in relation to publication quality and top quality. This reflects the fact that not all fields are equal in terms of publication opportunities. To the best of our knowledge, no one has attempted to assess whether this is due to a purely "fashion" effect (some topics are more "fashionable", which makes publication easier) or to some selection effects (more able academics self-select in certain fields or particular fields attract more able academics). We do not aim in this paper to tackle this difficult question. However, in terms of interpretation, here, individual and department effects are estimated net of the direct role of publication fields.

5. Channels of department externalities

Recall, that we want not only to emphasise that local effects are present in academia, which is in line with the literature and discussed in Section 3, but also to assess whether the standard local characteristics considered in urban economics, completed by a number of others potentially relevant for academic research activity, matter. This requires us to study both the impact of the field presence and specialisation variables in the first-step estimation, and the determinants of the department fixed effects in the second step. The results are presented in Table 3. The first two rows report the results of the first-step estimations (Table 1); subsequent rows report the estimations of specification (2) on the panel of department-time fixed effects.

As shown by the low within-time R^2 values reported in the penultimate row of Table 3, we can conclude that, broadly, department-time fixed effects are difficult to explain when controlling for individual fixed effects. Even when not controlling for individual fixed effects, the explanatory power of overall department characteristics is lower than usually obtained for the case of market activities. As reported in Table 2 for quality and in Appendix C for the other publication measures, the two field-specific department characteristics, introduced in the first-step specification, also have much lower explanatory power than the department-time fixed effects. Generally, the explanatory power of department characteristics is slightly higher at the two extremes of the publication measures as shown by the full variance analysis provided in Table 4, which does not account for individual fixed ef-

[15] $1.5^{-0.925} - 1$, $1.5^{0.192} - 1$ and $1.5^{0.539} - 1$, respectively.

[16] $2^{0.092} - 1$ and $2^{0.494} - 1$ respectively.

Table 3
The effects of department characteristics.

	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field presence (from 1st step)	0.063^{a}	0.122^{a}	0.345^{a}	0.334^{a}	0.114^{a}	0.087^{a}	0.359^{a}	0.318^{a}
	(0.001)	(0.002)	(0.022)	(0.022)	(0.021)	(0.019)	(0.055)	(0.048)
Specialisation (from 1st step)	0.014^{a}	0.024^{a}	0.098^{a}	0.088^{a}	0.036^{a}	0.020^{a}	0.132^{a}	0.084^{a}
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.003)	(0.009)	(0.008)
Size	0.003^{a}	−0.002	0.009	0.000	0.034^{a}	−0.013	0.055	−0.033
	(0.001)	(0.002)	(0.011)	(0.015)	(0.013)	(0.014)	(0.035)	(0.035)
% women	0.009	0.008	−0.041	0.157	−0.002	0.068	0.011	−0.110
	(0.007)	(0.014)	(0.083)	(0.120)	(0.097)	(0.109)	(0.264)	(0.276)
Average age	0.001^{a}	0.002^{a}	−0.002	0.002	−0.006^{c}	−0.005	−0.026^{a}	−0.025^{a}
	(0.000)	(0.001)	(0.003)	(0.004)	(0.003)	(0.004)	(0.009)	(0.009)
% rank A	−0.013^{b}	−0.037^{a}	0.101	−0.149	0.286^{a}	−0.022	1.061^{a}	−0.150
	(0.005)	(0.011)	(0.067)	(0.097)	(0.079)	(0.089)	(0.214)	(0.225)
Diversity	0.001	−0.011^{a}	−0.073^{a}	−0.059^{b}	0.043^{c}	−0.027	0.054	−0.043
	(0.002)	(0.003)	(0.020)	(0.027)	(0.023)	(0.024)	(0.064)	(0.062)
Research access	0.001^{b}	−0.003^{b}	0.025^{a}	−0.012	0.036^{a}	0.001	0.122^{a}	0.033
	(0.001)	(0.001)	(0.007)	(0.010)	(0.008)	(0.009)	(0.022)	(0.024)
Heterogeneity	−0.022^{a}	−0.021^{a}	0.000	−0.014	0.098^{a}	0.025	0.382^{a}	0.141^{c}
	(0.002)	(0.004)	(0.026)	(0.034)	(0.031)	(0.032)	(0.084)	(0.079)
USA connections	0.139^{a}	0.094^{a}	−0.270	0.048	1.052^{a}	0.251	3.080^{a}	0.700
	(0.017)	(0.025)	(0.181)	(0.219)	(0.220)	(0.205)	(0.597)	(0.515)
Non-USA connections	0.162^{a}	0.031	0.267^{c}	0.192	0.302^{c}	−0.289^{c}	1.324^{a}	−0.549
	(0.013)	(0.020)	(0.144)	(0.177)	(0.175)	(0.166)	(0.474)	(0.416)
Positions' shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE in 1st step	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.62	0.72	0.69	0.58	0.56	0.49	0.56	0.62
OLS within-time R^2	0.47	0.15	0.08	0.09	0.22	0.04	0.28	0.06
Observations	1208	1208	1208	1208	1208	1208	1208	1208

Notes: Feasible General Least Squares. Standard error between brackets. ^{a}, ^{b}, ^{c}: significant at the 1%, 5% and 10% levels respectively.

Table 4
Variance analysis of the determinants of department fixed effects.

	Probability		Quantity		Quality		Top Quality	
	Std. dev.	Corr.	Std. dev.	Corr.	Std. dev.	Corr.	Std. dev.	Corr.
Explained: Dep. fixed eff.	0.027	1.000	0.285	1.000	0.350	1.000	0.979	1.000
Observed characteristics	0.019	0.686	0.082	0.288	0.166	0.474	0.514	0.525
Composition effects	0.011	0.065	0.059	0.205	0.104	0.366	0.278	0.396
Gender	0.001	−0.032	0.001	0.011	0.012	0.084	0.039	0.089
Age	0.003	−0.090	0.003	−0.030	0.011	−0.019	0.059	−0.011
Rank A	0.002	−0.222	0.023	0.167	0.042	0.342	0.180	0.393
Positions	0.009	0.145	0.050	0.164	0.082	0.280	0.164	0.222
Research characteristics	0.021	0.587	0.057	0.200	0.108	0.375	0.354	0.450
Size	0.003	0.043	0.002	0.011	0.020	−0.008	0.025	−0.003
Research access	0.002	0.210	0.038	0.168	0.055	0.288	0.185	0.333
Diversity	0.001	0.138	0.039	0.109	0.011	0.016	0.015	−0.004
USA connections	0.008	0.463	0.019	−0.099	0.058	0.333	0.181	0.392
Non-USA connections	0.010	0.486	0.021	0.132	0.027	0.313	0.105	0.366
Heterogeneity	0.009	0.341	0.001	−0.001	0.032	−0.108	0.127	−0.089
Residuals	0.020	0.728	0.273	0.958	0.308	0.881	0.833	0.851

fects. Department characteristics matter more for either just publishing or at the other extreme of the distribution, publishing in a high quality journal, and are less important for number of publications.

The only department characteristics that exert a significant positive impact on individual publications, both with and without individual fixed effects, and for all publication dimensions, are the two field-specific characteristics. The marginal effects of field-specific characteristics are quite large. For instance, the presence of an academic's field in the department increases the number of other academics' publications in that field, and their average top quality by almost 40% (39.7% and 37.4% respectively). The effect is lower for the probability to publish and average standard quality (13.0% and 9.1%).[17] Similarly, the elasticity of specialisation, which for market activities is in the range 0.01-0.05 for productivity, is significantly larger here for quantity and top quality. Doubling the department's share of publications in a field (corresponding to an increase of one standard deviation at the median), increases individual quantity and average top quality by 6.3% and 6.0% respectively.[18] From the regressions reported in Bosquet and Combes (2017), we see that controlling for individual quantity in the first step estimation, positive field presence and specialisation impact are due not just to the fact that these characteristics increase indi-

[17] $e^{0.334}-1$, $e^{0.318}-1$, $e^{0.122}-1$, and $e^{0.087}-1$, respectively.

[18] $2^{0.088}-1$ and $2^{0.084}-1$.

vidual quantity, which, in turn, increases average quality as shown in Section 4. The impact of the department-field variables remains significant when controlling for individual quantity, and lower by only around one third. Overall, more academics in the department in a given field helps other academics to publish more and to produce publications of higher quality, in that field. The so-called localisation effects that reflect local economies of scale within the field, similar to their effect within industries in market activities, impact significantly publications in economics in France.

The impact of size of the local economy on local productivity has been the focus of many urban economics papers. We could have evaluated the role of the size of the city where the university is located. However, we believe that local externalities may be even more localised in relation to academic activities, for instance, because these activities require more face-to-face contact. Therefore, we use department size, defined as the number of full-time equivalent academics. This is an interesting variable for policy since it is, at least partly, down to the department head, the university or central government (in many European countries). We test the relevance of our choice of spatial scale in two ways. First, we include in the specification a research access variable for proximity to other departments, which allows us to separate very local size externalities from more diffuse ones. Second, we provide estimates at the employment area level, a more aggregated spatial classification.

Department size, similar to overall department characteristics, has no significant impact on department fixed effects when controlling for individual fixed effects. Without individual fixed effects, the positive elasticity obtained for quality is of the same magnitude as the estimates in the literature for market activities. Doubling department size would increase average quality by 3%-4%. The fact that the effect disappears when individual fixed effects are introduced might mean that academics with better characteristics join larger departments.

The findings for the “Research access” variable seem to indicate a further role of local size at a larger spatial scale, through proximity to nearby departments. When not controlling for individual fixed effects, the elasticity of research access is positively significant for all dimensions of publication productivity. It is fairly high for top quality, since multiplying research access by 5, which is around one standard deviation at the median, increases average top quality by 21.7%. The impact is still 6.0% for quality and 4.1% for quantity.[19] Again, this may be the result of the sorting of more efficient academics according to unobserved characteristics, in departments with better research access since the effects are no longer significant when individual fixed effects are controlled for. The explanatory power is also much higher for market access than for size, with a still quite low standard deviation of the effects, but a rather large correlation to the dependent variable (0.33 for top quality from Table 4 for instance).

More in line with the analyses in urban economics, we can assess the role of size on a larger spatial scale by aggregating the data at city level. We use employment areas, which refer to 341 spatial units in France built by INSEE, the French national institute of statistics, specifically to study the role of local labour markets. For many employment areas, there are either no universities or only one (for 38 employment areas): thus, considering department or employment area is the same. Six employment areas host two departments, four employment areas host three departments, and three employment areas host four or more departments. The results in Bosquet and Combes (2017) are very similar at this scale compared to university departments. In particular, the elasticity of local size, which should be the most affected by this change of spatial scale, is generally only slightly smaller and less significant. Research access still matters, and its impact on top quality remains significant when individual fixed effects are controlled for, with a high value at 0.099. Overall, it is difficult to assess whether local effects matter more at the department or at the city level, possibly because these two levels are too similar. However, moving to a larger spatial scale, that is, the region, would reduce the number of observations too much and would make less sense in terms of connections between departments.

In addition to size and market access, some other department variables have significant explanatory power when not controlling for individual fixed effects. This contrasts with the findings in much of the urban economics literature, which show that these two variables mostly explain productivity differences across locations. We study the role of co-authors' locations based on the connection variables. The literature on academic networks (see for instance Laband and Tollison, 2000; Rosenblat and Mobius, 2004) shows that distance to co-authors has increased significantly over time. If the links to co-authors were not controlled for, access to departments outside France could have been computed and we might have found a positive effect. Here, having co-authors abroad, either in the USA or elsewhere, increases both the individual quantity and quality of publications, as shown in Section 3. Also, these two variables are among those that have the largest explanatory power of department fixed effects. The elasticities are large even though, again, they reduce and lose significance when individual fixed effects are controlled for. The role of USA connections is the strongest for the extreme dimensions of publication, that is, the probability to publish, and the average top quality, while non-USA connections affect all dimensions of publication activity (probability to publish, number of publications and publication quality). In this case, the effect increases moving across these dimensions (when individual fixed effects are not controlled for). In a world where distance matters much less than previously, being connected to other academics elsewhere and, in particular, to colleagues in the USA, remains important. This is in line with the major role of networks in academia that is underlined in the literature.

Having heterogeneous academics in a department enhances average publications quality and, especially top quality, with an effect that remains slightly significantly positive at 10% even when controlling for individual fixed effects. We are not aware of a similar finding in the literature. The presence of top people in the department may help others publishing in the best journals. The explanatory power of this variable is also quite large. By contrast, heterogeneity has no effect on publications quantity and has a small negative effect on the probability to publish. Field diversity in the department has a rather small impact in general, most often negative, but, in any case, quite sensitive to the specification chosen. Industry diversity for market activities is also often found to have a not very robust impact.

Finally, we consider as department characteristics the shares in the department of the various individual characteristics considered in the first step. Taken together, they have similar explanatory power to department overall research characteristics (respectively rows ‘Composition effect’ and ‘Research characteristics’ in Table 4). However, it is the department composition in terms of positions that matters most. Recall that when individual position is controlled for in the first-step, a larger share of higher/the most productive positions tend, in general, to increase the publication output of all other academics in the department. Bosquet and Combes (2017) find that this is especially true when not controlling for individual fixed effects, and a few effects remain significant even when individual fixed effects are controlled for. The share of women, and average age explain much less. Older academics, in a given position in the department, exert a significant effect even

[19] $5^{0.122} - 1$, $5^{0.036} - 1$ and $5^{0.025} - 1$, respectively.

when controlling for individual fixed effects. The effect is positive, but negligible for the probability to publish, and slightly larger and negative for top quality. Increasing department's average age by three years (which is close to one standard deviation at the median), for a given composition of positions, decreases top quality by 7.2%.[20] The share of women in the department has no impact on publication output, neither positive nor negative.

6. Conclusions

Location matters for the publication performance of French economists and economics departments. A careful variance analysis of individual publication determinants shows that the explanatory power of department effects accounts for at least a quarter of the explanatory power of individual effects. This corresponds to what many academics would expect. However, it contrasts sharply with previous findings of the literature, which show the presence of small or no local effects. We attribute this difference to our exhaustive data covering all academic economists in France with many individual variables, which allow us (i) to follow them over time and across locations even if they do not publish, (ii) to consider more local effects and (iii) to develop a more complete econometric strategy. Moreover, we separately studied the determinants of the probability to publish, the number of publications and their average quality, whereas most studies in the literature consider only the quality-adjusted number of publications.

A new and puzzling result emerges, which future research should investigate: neither standard variables considered in urban economics nor a number of departments' research characteristics explain overall local effects. Nevertheless, department field specialisation, which measures localisation economies at the field level has a robust and rather large positive impact on the department academics' publications in that field. We find some evidence that academic mobility in France might be too low to properly identify individual and department-time fixed effects simultaneously. This suggests that our strategy should be replicated with other countries to confirm or reject these findings and to propose some further variables that might explain the roles of both department and individual unobserved characteristics. It is possible that teaching load and student quality or the efficiency of the local administration, which cannot be controlled for here, are important. Also, more subtle effects, such as the atmosphere in the department, which sometimes is mentioned by academics, might also have an effect.

Due to possible missing variables and reverse causality when estimating the agglomeration effects, we do not claim to provide a conclusive assessment of the role of department effects on individual performance. We would argue that endogenous individual location choices are of little concern, at least in the French context, while historical natural experiments lack external validity. Issues arising from the endogeneity of department characteristics might be more serious and should be treated. The possibility of combining bibliometric and administrative sources, as in the current paper, should be extended to cover longer periods, other fields and other countries. This would allow researchers to identify even more sources of exogenous variation in order to properly assess the role of endogenous and exogenous individual and department characteristics. Adding more structure to the underlying network formation and agglomeration and peer effects models, which here are only implicit and, therefore, treated as a black box, would help to improve the estimated specifications. Ultimately, this and future work should provide important results to contribute to the better design of higher education and research policies.

[20] $e^{-3\times 0.025}-1$.

Appendix A. Definition of variables

A1. Publication measures

Academics' publication output at date t is given by the sum of the value of all their publications over period τ (τ corresponds to years $t+1$, $t+2$, and $t+3$ in our implementation). To assess the value of a publication, each publication a is first weighted by the quality of the journal, $W(a)$, in which it is published. We use the Combes and Linnemer (2010) journal weighting scheme. Each journal weight is a weighted average of various recursive impact factors built from Thomson Reuters Web of Knowledge impact factors[21] and from Google Scholar citations.[22] For journals not listed in the Web of Knowledge, Combes and Linnemer (2010) use an econometric model to infer their weight. This leads to a ranking of all EconLit journals. Unfortunately, the ranking is constant over time and all of a journal's publications receive the same weight independent of their publication year. Then a function is applied to the ranking to obtain more or less selective weighting schemes. Here, we compare the determinants of publications using two of them, *CLm* in which selectivity is moderate (ranging from a weight of 100 for the *Quarterly Journal of Economics* through 55.1 for the *Journal of Labor Economics*, for instance, to a weight of 4 for the lowest ranked journal) and *CLh* which is more selective (ranging from 100 for the *Quarterly Journal of Economics* to 0.0007 for the lowest ranked journal, via 16.7 for the *Journal of Labor Economics*). We refer to these two schemes as the 'Quality' and 'Top quality' publication measures, respectively.

Publication a is also weighted by the inverse of its number of authors, $n(a)$. Since a department's publications output is the sum of the publication outputs of its academics, we do not want a publication written by two members of the department to account for more (or less) than the same publication written by a single author. As mentioned above, we evaluate the presence of increasing or decreasing returns to scale within co-author teams, using the average number of authors per publication as one of the independent individual variables.

The output measure also takes account of the number of pages in the article, $p(a)$, relative to the average length of all articles in EconLit in the same journal in the same year, $\overline{p}$. This captures the idea that longer articles should contain more ideas and innovations. A natural example is provided by the difference between short and regular papers in the *American Economic Review*. Importantly, these means are computed within each journal-year. This assumes that the editorial policy of the journal is consistent within a year, an article 20% shorter than the journal average representing 20% less output, for instance. Conversely, differences in article length between journals, which can stem from different page and font sizes or from real contribution differences, are assumed to be directly and fully reflected in the journals' quality weight, $W(a)$. In some sense, our choice is intermediate between fully ignoring publication length, and using the absolute number of pages as sometimes done in the literature.

Finally, publication output is measured at the field level to enable us to study the effect of field-specific characteristics and to control for between-field differences at the French level. We use Jel codes at the first digit level (letter) and we ignore the fields "Y - Miscellaneous Categories" and "Z - Other Special Topics". We also slightly modify the codes C and D by merging codes C7 (Game Theory and Bargaining Theory) and C9 (Design of Experiments) with Microeconomics (code D) and removing them from Mathematical and Quantitative Methods (code C), which we believe is

[21] http://www.webofknowledge.com/.

[22] http://scholar.google.com/.

more coherent. This leaves us with 18 fields. The weight of publication a attributed to academic i is first divided by the publication's number of Jel codes, $J(a)$, and then multiplied by the publication's number of Jel codes corresponding to field f, $J_f(a)$.

To sum up, the publication output of academic i at date t in field f is given by:

$$y_{ift} = \frac{1}{\text{Card}(\tau)} \sum_{a \in (i,\tau)} \frac{W(a)}{n(a)} \frac{p(a)}{\overline{p}} \frac{J_f(a)}{J(a)} ,$$

where $\text{Card}(\tau)$ is the number of years in period τ, typically $\text{Card}(\tau) = 3$ for $\tau = t+1, t+2, t+3$.

Then, y_{ift} is decomposed as follows:

$$y_{ift} \equiv \mathbb{1}(\text{Quantity}_{ift} > 0) \times \text{Quantity}_{ift} \times \frac{y_{ift}}{\text{Quantity}_{ift}} ,$$

where Quantity_{ift} is the number of publications of academic i in field f at date t, $\text{Quantity}_{ift} = \frac{1}{\text{Card}(\tau)} \sum_{a \in (i,\tau)} \frac{J_f(a)}{n(a)J(a)}$. The first component is a dummy variable equal to 1 when at least one of academic i's publications at date t refers to Jel code f. The second component measures the publication quantity of active academic i in field f at date t. The last component corresponds to the average quality of publications of active academics i in field f at date t. The observation weighting in estimations uses the share of field f in academic i's output at date t, $S_{ift} = \frac{\text{Quantity}_{ift}}{\text{Quantity}_{it}}$, multiplied by α_{idt}, the share of academic i's output attributed to department d at date t defined below.

A2. Department characteristics

Let y_{ift} denote the publication yearly output of academic i in field f at date t. The total yearly output of academic i at date t is the sum of y_{ift} over all fields, $y_{it} = \sum_f y_{ift}$. Since some academics split their time between multiple departments, let α_{idt} denote the share of academic i's output attributed to department d at date t. 90% of our academics have only one affiliation, in which case α_{idt} is equal to 1 for one department only and to 0 for all others. Then an equal splitting between all affiliations is assumed, and α_{idt} is 1 over the total number of affiliations. Department d's yearly output in field f at date t is given by $Y_{dft} = \sum_{i \in dt} \alpha_{idt}\, y_{ift}$ and its total output by $Y_{dt} = \sum_f Y_{dft} = \sum_{i \in dt} \alpha_{idt}\, y_{it}$.

The time span used here to measure activity at date t consists of the moving average of the publications over the three years that follow t. Importantly, because academics' field choice in a given year may not perfectly reflect their average field of specialisation, we compute individual variables that depend on field over a longer time span. In this case, as suggested by Combes and Linnemer (2010), we discount, over time, past publications in the field: all publications until τ are taken into account, but more recent ones count more than older ones. The discount factor t' years before t corresponds to a logistic function given by $\frac{1-\exp\left(-10/(t'+1)^{1.8}\right)}{1+\exp\left(-20/(t'+1)^{1.2}\right)}$. Year t publications count 1, year $t-1$ publications count 0.94, year $t-2$ publications count 0.75, and so on. For instance, after 10 years they count 0.20, and after 20 years 0.10 only.

The first 'Department's Field-Specific Characteristic' consists of a dummy variable for the presence of the field in the department:

$$\begin{aligned} \text{Field Presence}_{dft} &= 1 \quad \text{if } \tilde{Y}_{dft} - \tilde{y}_{ift} > 0, \\ &= 0 \quad \text{else}, \end{aligned}$$

where $\tilde{Y}_{dft}$ and $\tilde{y}_{ift}$ are the same publication measures as those used to measure academics' output, but calculated over a different time span, which considers all publications, but discounts them over time.

Note that, strictly speaking, this variable and some described subsequently, depend on each individual and not just the department. However, they capture the notion of external effects and we prefer to keep these notations for the sake of clarity. For market activities, it is less crucial to exclude the own individual values from the computation of local variables since, in any case, many are negligible and the measure is almost not affected. Here, an individual can, on his or her own, represent a large share of the department's output in a field, and this makes interpretation cleaner.

Then a specialisation variable - the share of department d's output in field f at date t (other than the academic's output) - is given by:

$$\text{Specialisation}_{dft} = \log \frac{\tilde{Y}_{dft} - \tilde{y}_{ift}}{\tilde{Y}_{dt} - \tilde{y}_{it}} .$$

The diversity index net of the size effect is given by:

$$\text{Diversity}_{dt} = \log \left[\sum_f \left(\frac{\tilde{Y}_{dft}}{\tilde{Y}_{dt}} \right)^2 \right]^{-1} - \log \left[\sum_f \left(\frac{\check{Y}_{dft}}{\check{Y}_{dt}} \right)^2 \right]^{-1} ,$$

where $\sum_f \left(\frac{\check{Y}_{dft}}{\check{Y}_{dt}} \right)^2$ is a randomly-generated Herfindahl index built by simulations. Indeed, a problem arises because, by construction, the crude Herfindhal diversity index is highly correlated to department size. This is because departments with small numbers of academics have many Jel codes without any publications. To remove this size effect, which is absent in standard urban economics studies because there are few locations without any activity in an industry, we subtract from the gross diversity index the value it would take if all academics in the department were to randomly choose their Jel codes. We first attribute random Jel codes to each publication, assuming that the probability to publish in each Jel code follows a binomial law with a probability of success given by the share of output in each Jel code at the French level. Then, the department diversity index is recomputed using these new Jel codes. The randomly-generated Herfindahl index for the department is the average of 1000 such procedures.

The research access variable is the spatially-discounted sum of the sizes of all other departments:

$$\text{Research Access}_{dt} = \log \sum_{d' \neq d} \frac{\text{Size}_{d't}}{Dist_{dd'}} ,$$

where $Dist_{dd'}$ is the geographical distance between departments d and d'. Alternative specifications of the research access variable, with squared distance or the square root of distance in the denominator, were tested and led to qualitatively similar results. We keep the most standard one.

The department's heterogeneity in terms of academics' publication records is measured by the within-department coefficient of variation of individual output:

$$\text{Heterogeneity}_{dt} = \log \frac{\text{Standard Deviation}(\tilde{y}_{it})_{i \in (d,t)}}{\text{Average}(\tilde{y}_{it})_{i \in (d,t)}} ,$$

where $\text{Standard Deviation}(\tilde{y}_{it})_{i \in (d,t)}$ and $\text{Average}(\tilde{y}_{it})_{i \in (d,t)}$ are the standard deviation and the average of individual publication outputs within department d at date t.

Finally, field diversity at the individual level is given by:

$$\text{Individual Diversity}_{it} = \log \left[\sum_f \mathbb{1}(\tilde{y}_{ift} > 0) \right] ,$$

where $\mathbb{1}(\tilde{y}_{ift} > 0)$ is a dummy variable equal to 1 when academic i has at least once published in field f until date t.

Appendix B. Descriptive statistics and simple correlations

Panel (a) in Table B.1 presents descriptive statistics for all academics. The average academic is 45.6 years old and 25% are women. We do not present the share of each of the 10 positions distinguished, but we create two aggregate variables that characterise them. The line 'Teaching' reports that 83% of academic-year observations have statutory teaching loads, others correspond to researcher academics. The line 'Rank A' reports that 35% of academic-year observations correspond to a rank A position, that is, equivalent to full professor as opposed to assistant professor.

Not all academics publish over a given three-year period. The 'Publisher' row in Table B.1 panel (a) reports that one-third have published at least one article over the three-year period, possibly co-authored, and in any field. This is one of the figures that changed quite substantially over time, rising from 0.17 in 1990 to 0.42 in 2005. Panel (b) in Table B.1 provides descriptive statistics for the sub-group of academics who published at least one article over the three years. They are almost three years younger, slightly less likely to be women, and more likely to hold non-teaching and rank A positions.

The row 'Quantity' in Table B.1 panel (a) reveals that the average academic publishes 0.17 papers equivalent alone per year, which is one paper with one co-author every three years. This is small, but due partly to the fact that many academics do not publish any papers. Conditional on having at least one publication over the three-year period, Table B.1 panel (b) shows that the average number of publications is three times higher, corresponding to, for instance, one single authored publication and one publication with a co-author, every three years. In relation to quality, we can confirm the well documented large disparities existing among academics (see Lotka, 1926). The mean publication is worth the equivalent of one publication per year in the 150th ranked journal, but the median publication is lower, at around the 350th ranked journal. By contrast, the top decile average quality publication corresponds to one publication per year in the 50th ranked journal or

Table B1
Descriptive statistics.

	Mean	Standard deviation	1st decile	Median	Last decile
Panel (a): All academics					
Age	45.6	9.1	32	46	58
Women	0.25	0.41	0	0	1
Rank A	0.35	0.45	0	0	1
Teaching	0.83	0.36	0	1	1
Publisher	0.33	0.45	0	0	1
Quantity	0.17	0.36	0	0	0.57
Quality	4.3	10.2	0	0	12.1
Top quality	0.80	5.31	0	0	0.22
Panel (b): Publishers					
Age	42.7	9.1	31	41	56
Women	0.22	0.38	0	0	1
Rank A	0.49	0.47	0	0	1
Teaching	0.75	0.40	0	1	1
Quantity	0.52	0.46	0.17	0.33	1.06
Quality	13.3	14.3	4.0	7.9	29.4
Top quality	2.44	8.92	0.01	0.04	4.94
Authors per publication	1.9	0.7	1	2	3
Non-USA connection	0.1	0.3	0	0	1.0
USA connection	0.07	0.24	0	0	0
Individual diversity	2.6	1.6	1	2	5
Panel (c): Departments					
Size	31.6	34.6	7.5	18.0	82.0
Age	45.0	3.5	40.6	45.0	49.3
Women	0.24	0.12	0.10	0.24	0.39
Rank A	0.34	0.19	0.13	0.31	0.64
Teaching	0.79	0.34	0	0.97	1
Publishers	0.34	0.20	0.11	0.30	0.62
Quantity	5.47	8.21	0.44	2.76	12.64
Quantity per academic	0.18	0.18	0.04	0.13	0.37
Quality	11.82	8.16	5.67	9.19	20.85
Top quality	1.94	4.48	0.02	0.27	5.04
Field presence	0.60	0.27	0.21	0.61	0.94
Specialisation	0.27	0.18	0.12	0.21	0.50
Diversity	−0.54	0.46	−1.18	−0.46	−0.01
Research access	49.7	63.5	6.0	17.0	154.4
Non-USA connection	0.04	0.07	0	0.01	0.13
USA connection	0.02	0.06	0	0	0.07
Heterogeneity	2.3	0.9	1.3	2.1	3.5

Notes: Variables are defined in Section 2.2. To match what is done in the econometric section, publication variables are first computed as three-year moving averages before descriptive statistics are computed. The number of observations for panels (a), (b) and (c) are 39,266, 12,924, and 1208 respectively. Descriptive statistics at the department level (panel (c)) are calculated on the sub-sample of departments in which there is at least one published author and, hence, for which all variables are defined. Specialisation defined at the Jel code level is first averaged by department (weighted by the share of the Jel code in the department output), then the statistics are computed.

Table B2
Simple correlations at the department level.

	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Quantity (1)	0.92	0.73	0.06	−0.15	−0.19	0.52	−0.40	0.51	−0.26	0.23	0.26	0.52	0.48	−0.47
Quality (2)	1	0.91	0.10	−0.16	−0.19	0.56	−0.46	0.54	−0.20	0.14	0.31	0.60	0.59	−0.32
Top quality (3)		1	0.20	−0.12	−0.17	0.51	−0.40	0.56	−0.18	0.04	0.31	0.54	0.54	−0.09
Size (4)			1	0.18	−0.05	0.06	0.15	0.71	−0.44	0	−0.11	0.01	−0.05	0.38
Age (5)				1	−0.18	0.28	0.02	0.01	0	−0.08	0.19	−0.23	−0.15	0.27
Women (6)					1	−0.27	0.14	−0.14	0.09	−0.08	0.13	−0.18	−0.12	0.07
Rank A (7)						1	−0.58	0.29	−0.07	−0.01	0.43	0.41	0.44	−0.27
Teaching (8)							1	−0.07	−0.09	0.05	−0.42	−0.40	−0.47	0.36
Field presence (9)								1	−0.58	0.30	0.07	0.30	0.26	0.06
Specialisation (10)									1	−0.78	0.06	−0.05	−0.01	0.06
Diversity (11)										1	−0.04	0.04	0.01	−0.31
Research access (12)											1	0.26	0.31	−0.14
Non-USA connection (13)												1	0.61	−0.26
USA connection (14)													1	−0.22
Heterogeneity (15)														1

Notes: Variables are defined in Section 2.2. Specialisation defined at the Jel code level is first averaged by department (weighted by the share of the Jel code in the department output) and all variables are detrended before statistics are computed.

one publication in one of the top 5 journals every three years. The average quality of publications of academics in France appears to be better in terms of the top quality index, since the mean is now around the 50th journal, the median around the 100th, and the top decile around the 30th journal. 10% of publishing academics have at least one co-author abroad, but not in the USA, and 7% have at least one co-author in the USA. The average number of authors per paper is 1.9 and, more precisely, 44.7%, 38.0% and 14.8% of the publications have one, two, and three authors respectively. Only 2.5% of the publications have strictly more than three authors.

The row ‘Individual Diversity’ in Table B.1 panel (b) reveals that the average number of fields per publishing academic over a three-year period is 2.6 and the very diversified academic at the top decile publishes in 5 fields. At the national level, ‘Microeconomics’ is the most represented field in France with 16.8% of publications. This is larger than its share in EconLit as a whole, which is 10.2%. Then, there are 10 fields each representing more than 4%.[23]

Panel (c) in Table B.1 reports the descriptive statistics at department level. The average department has 31.6 academics who are 45 years old on average, 24% are women, 34% have rank A positions, and 34% publish. The figures are comparable to the averages for all academics. Importantly, all the variables present quite a lot of variation between departments, in particular, for publication output. The average department has 5.5 publications per year, 0.18 per academic, and the average quality indexes are in the same ranges as for individual academics. The row ‘Field presence’ reveals that cumulative publications in the average department cover 60% of Jel codes at the first digit level (letter). Specialisation of the median department means that a Jel code present in the department represents 21% of the department’s cumulative output. Thus, departments are fairly specialised, given that there are 18 different possible Jel codes. In the very specialised department at the top decile of specialisation, each Jel code represents half of the publications. This is confirmed by the diversity index, which almost always takes negative values even at the top decile, meaning that departments are less diversified than they would be with random Jel code choices.

Finally, Table B.2 presents simple correlations between the variables at the department level. First, quantity and quality are mostly positively correlated even for the top quality index. Those departments that publish more also produce higher quality publications and there seems to be no trade-off between the two. This is in line with Combes and Linnemer (2003) findings at the European level. Academics are also, on average, more productive in departments where the share of rank A is higher and the share of teaching positions lower, and where field diversity and research access are high. Correlations are also positive, but lower for share of academics with co-authors abroad, and in the USA (connection variables), and, again, large for heterogeneity, which is positively and negatively correlated to quantity and quality respectively. The correlation of size to quantity is not very strong, but increases for quality and even more so for top quality. The challenge is to investigate whether these correlations are driven by the fact that rank A position researchers or researchers with higher abilities more generally, are over-represented in some departments through selection effects and/or by the fact that some academics or department characteristics generate more externalities.

Appendix C. Variance analyses of individual probability to publish, publication quantity and top quality

[23] ‘Industrial organization’ (9.5% vs 8.8% for EconLit as a whole), ‘Development/Growth’ (8.8 vs 10.0%), ‘Finance’ (8.8 vs 10.9%), ‘Macro/Monetary economics’ (8.2 vs 7.2%), ‘Labour/Demography’ (8.2 vs 8.3%), ‘International economics’ (7.6 vs 7.8%), ‘Agricultural/Environmental economics’ (5.6 vs 7.0%), ‘Economics history’, ‘Thoughts and methodology’ (5.4 vs 2.2%), ‘Public economics’ (4.2 vs 4.3%), ‘Urban and regional economics’ (4.2 vs 5.0%).

Table C1
Variance analysis of the individual probability to publish.

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Corr.	Sorting	Std. dev.	Corr.	Sorting
Explained: Publishing	0.203	1.000		0.266	1.000	
Individual effects	0.033	0.170	0.056	0.073	0.259	−0.096
Indiv. fixed effect	-	-	-	0.089	0.191	−0.071
Obs. indiv. effects	0.033	0.170	0.056	0.061	0.031	−0.012
Women	0.007	0.030	−0.012	-	-	-
Age	0.024	0.079	0.008	0.061	0.031	−0.012
Position	0.018	0.105	0.045	-	-	-
Rank A	0.020	0.085	0.048	-	-	-
Department effects	0.028	0.181	1.000	0.041	0.185	1.000
Dep.-time fixed eff.	0.022	0.119	0.754	0.026	0.053	0.590
Dep.-field-time eff.	0.018	0.133	0.625	0.034	0.188	0.779
Field presence	0.024	0.072	0.372	0.044	0.094	0.325
Specialisation	0.025	0.028	0.100	0.042	0.050	0.274
Field-time fixed effect	0.020	0.145	0.338	0.034	0.196	0.437
Residuals	0.197	0.966	0	0.249	0.935	0

Notes: Same notes as for Table 2.

Table C2
Variance analysis of the individual publication quantity.

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Corr.	Sorting	Std. dev.	Corr.	Sorting
Explained: Quantity	0.457	1.000		0.457	1.000	
Individual effects	0.214	0.471	0.007	0.324	0.654	-0.156
Indiv. fixed effect	-	-	-	0.261	0.502	−0.163
Obs. indiv. effects	0.214	0.471	0.007	0.197	0.410	−0.039
Women	0.027	0.093	0.019	-	-	-
Age	0.066	0.028	0.035	0.034	0.012	0.062
Position	0.029	0.090	0.022	-	-	-
Rank A	0.061	0.085	0.075	-	-	-
Authors per pub.	0.204	0.353	−0.077	0.199	0.353	−0.066
Individual diversity	0.031	0.046	0.008	0.042	0.046	0.018
Non-USA connection	0.074	0.110	0.064	0.038	0.110	0.032
USA connection	0.060	0.117	0.068	0.033	0.117	0.042
Department effects	0.123	0.274	1.000	0.134	0.169	1.000
Dep.-time fixed eff.	0.104	0.234	0.841	0.120	0.117	0.891
Dep.-field-time eff.	0.067	0.141	0.529	0.061	0.141	0.434
Field presence	0.048	0.033	0.186	0.047	0.033	0.151
Specialisation	0.074	0.105	0.357	0.066	0.105	0.291
Field-time fixed effect	0.071	0.158	0.012	0.078	0.102	−0.075
Residuals	0.377	0.825	0	0.313	0.685	0

Notes: Same notes as for Table 2.

Table C3
Variance analysis of the individual publication top quality.

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Corr.	Sorting	Std. dev.	Corr.	Sorting
Explained: Top quality	1.305	1.000		1.305	1.000	
Individual effects	0.528	0.527	0.238	1.028	0.807	−0.018
Indiv. fixed effect	-	-	-	1.050	0.622	−0.043
Obs. indiv. effects	0.528	0.527	0.238	0.686	0.258	0.039
Women	0.062	0.069	0.017	-	-	-
Age	0.270	0.141	−0.015	0.647	0.144	0.029
Position	0.126	0.180	0.108	-	-	-
Rank A	0.152	0.141	0.150	-	-	-
Authors per pub.	0.109	0.236	0.103	0.116	0.236	0.025
Individual diversity	0.035	0.179	0.125	0.008	0.179	0.017
Non-USA connection	0.223	0.334	0.138	0.068	0.334	0.019
USA connection	0.236	0.360	0.193	0.090	0.360	0.044
Department effects	0.468	0.489	1.000	0.285	0.209	1.000
Dep.-time fixed eff.	0.459	0.468	0.982	0.280	0.178	0.979
Dep.-field-time eff.	0.087	0.155	0.187	0.058	0.164	0.185
Field presence	0.050	0.068	0.102	0.044	0.068	0.011
Specialisation	0.099	0.103	0.114	0.063	0.103	0.161
Field-time fixed effect	0.265	0.337	0.168	0.159	0.281	0.038
Residuals	0.962	0.737	0	0.695	0.533	0

Notes: Same notes as for Table 2.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jue.2017.05.003.

References

Abowd, J.M., Kramarz, F., Lengerman, P., Perez-Duarte, S., 2004. Are good workers employed by good firms? a test of a simple assortative matching model for France and the United States. Mimeograph. Crest.
Abowd, J.M., Kramarz, F., Margolis, D.N., 1999. High wage workers and high wage firms. Econometrica 67 (2), 251–333.
Aghion, P., Dewatripont, M., Hoxby, C.M., Mas-Colell, A., Sapir, A., 2010. The governance and performance of research universities: evidence from Europe and the U.S.. Econ. Policy 25 (61), 7–59.
Agrawal, A., Kapur, D., McHale, J., 2008. How do spatial and social proximity influence knowledge flows? evidence from patent data. J. Urban Econ. 64 (2), 258–269.
Andrews, M.J., Gill, L., Schank, T., Upward, R., 2012. High wage workers match with high wage firms: clear evidence of the effects of limited mobility bias. Econ. Lett. 117 (3), 824–827.
Auriol, E., Friebel, G., Lammers, F., 2012. The firm as the locus of social comparisons: internal labor markets versus up-or-out. Discussion Paper 8831. Centre for Economic Policy Research.
Azoulay, P., Graff Zivin, J.S., Wang, J., 2010. Superstar extinction. Q. J. Econ. 125 (2), 549–589.
Baum-Snow, N., Pavan, R., 2012. Understanding the city size wage gap. Rev. Econ. Stud. 79 (1), 88–127.
Bayer, P., Hjalmarsson, R., Pozen, D., 2009. Building criminal capital behind bars: peer effects in juvenile corrections. Q. J. Econ. 124 (1), 105–147.
Beaudry, P., Green, D.A., Sand, B., 2014. Spatial equilibrium with unemployment and wage bargaining: theory and estimation. J. Urban Econ. 79, 2–19.
Besley, T., Ghatak, M., 2008. Social incentives. Am. Econ. Rev. 98 (2), 206–211.
Borjas, G.J., Doran, K.B., 2012. The collapse of the Soviet Union and the productivity of American mathematicians. Q. J. Econ. 127 (3), 1143–1203.
Bosquet, C., Combes, P.-P., 2017. Supplement to 'Sorting and agglomeration economies in French economics departments': additional appendices. Available on. https://sites.google.com/site/clementbosquet/.
Bosquet, C., Combes, P.-P., Garcia-Peñalosa, C., 2014. Gender and promotions: evidence from academic economists in France. LIEPP Working Paper 29.
Carlino, G., Kerr, W., 2015. Agglomeration and innovation. In: Duranton, G., Henderson, V., Strange, W. (Eds.), Handbook of Urban and Regional Economics, 5A. North-Holland, Amsterdam.
Ciccone, A., Peri, G., 2006. Identifying human capital externalities: theory with an application to US cities. Rev. Econ. Stud. 73 (2), 381–412.
Combes, P.-P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters!. J. Urban Econ. 63 (2), 723–742.
Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., Roux, S., 2012. The productivity advantages of large markets: distinguishing agglomeration from firm selection. Econometrica 80 (6), 2543–2594.
Combes, P.-P., Duranton, G., Gobillon, L., Roux, S., 2010. Estimating agglomeration effects with history, geology, and worker fixed-effects. In: Glaeser, E.L. (Ed.), Agglomeration Economics. Chicago University Press, Chicago, IL, pp. 15–65.
Combes, P.-P., Gobillon, L., 2015. The empirics of agglomeration economies. In: Duranton, G., Henderson, V., Strange, W. (Eds.), Handbook of Urban and Regional Economics, 5A. North-Holland, Amsterdam.
Combes, P.-P., Linnemer, L., 2003. Where are the economists who publish? publication concentration and stock rankings in europe. J. Eur. Econ. Assoc. 1 (6), 1250–1308.
Combes, P.-P., Linnemer, L., 2010. Inferring missing citations. A quantitative multi-criteria ranking of all journals in economics. GREQAM Working Paper 2010-25.
Combes, P.-P., Linnemer, L., Visser, M., 2008. Publish or peer-rich. the role of skills and networks in hiring economics professors. Labour Econ. 15 (3), 456–474.
Damm, A.P., 2014. Neighborhood quality and labor market outcomes: evidence from quasi-random neighborhood assignment of immigrants. J. Urban Econ. 79, 139–166.
Dubois, P., Rochet, J.-C., Schlenker, J.-M., 2014. Productivity and mobility in academic research: evidence from mathematicians. Scientometrics 98 (3), 1669–1701.
Ductor, L., 2015. Does co-authorship lead to higher academic productivity? Oxf. Bull. Econ. Stat. 77 (3), 385–407.
Ductor, L., Fafchamps, M., Goyal, S., van der Leij, M.J., 2014. Social networks and research output. Rev. Econ. Stat. 96 (5), 936–948.
Gould, E., 2007. Cities, workers, and wages: a structural analysis of the urban wage premium. Rev. Econ. Stud. 74, 477–506.
Hellerstein, J.K., Kutzbach, M.J., Neumark, D., 2014. Do labor market networks have an important spatial dimension? J. Urban Econ. 79, 39–58.
Hellerstein, J.K., Neumark, D., Troske, K.R., 1999. Wages, productivity, and worker characteristics: evidence from plant-level production functions and wage equations. J. Labour Econ. 17 (3), 409–446.
Hollis, A., 2001. Co-authorship and the output of academic economists. Labour Econ. 8 (4), 503–530.
Jackson, M.O., 2011. An overview of social networks and economic applications. In: Jess Benhabib, A.B., Jackson, M.O. (Eds.), Handbook of Social Economics, 1. North-Holland, Amsterdam, pp. 511–585.
Jacobs, J., 1969. The Economy of Cities. Random House, New York.
Kang, C., 2007. Classroom peer effects and academic achievement: quasi-randomization evidence from South Korea. J. Urban Econ. 61 (3), 458–495.
Kim, H.E., Morse, A., Zingales, L., 2009. Are elite universities losing their competitive edge? J. Financ. Econ. 93 (3), 353–381.
Laband, D.N., Tollison, R.D., 2000. Intellectual collaboration. J. Political Econ. 108 (3), 632–662.
Lavy, V., Paserman, M.D., Schlosser, A., 2012. Inside the black box of ability peer effects: evidence from variation in the proportion of low achievers in the classroom. Econ. J. 122 (559), 208–237.
Lefgren, L., 2004. Educational peer effects and the chicago public schools. J. Urban Econ. 56 (2), 169–191.
Lotka, A., 1926. The frequency distribution of scientific productivity. J. Whashington Acad. Sci. 16 (12), 317–323.
Marshall, A., 1890. Principles of Economics. Macmillan, London.
Medoff, M.H., 2003. Collaboration and the quality of economics research. Labour Econ. 10 (5), 597–608.
Oyer, P., 2006. Initial labor market conditions and long-term outcomes for economists. J. Econ. Perspect. 20 (3), 143–160.
Rosenblat, T.S., Mobius, M.M., 2004. Getting closer or drifting apart? Q. J. Econ. 119 (3), 971–1009.
Rosenthal, S.S., Strange, W.C., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, 4. North-Holland, Amsterdam, pp. 2119–2171.
Sacerdote, B., 2011. Peer effects in education: how might they work, how big are they and how much do we know thus far?. In: Eric A. Hanushek, S.M., Woessmann, L. (Eds.) Handbook of the Economics of Education, 3. Elsevier, Amsterdam, pp. 249–277.
Sauer, R.D., 1988. Estimates of the returns to quality and coauthorship in economic academia. J. Political Econ. 96 (4), 855–866.
Waldinger, F., 2012. Peer effects in science: evidence from the dismissal of scientists in Nazi Germany. Rev. Econ. Stud. 79 (2), 838–861.
Zenou, Y., 2003. The spatial aspects of crime. J. Eur. Econ. Assoc. 1 (2–3), 459–467.