Do large departments make academics more productive? Sorting and agglomeration economies in research

Clément Bosquet[ab]
ThEMA (University of Cergy-Pontoise) and SERC (LSE)

Pierre-Philippe Combes[ac]
Aix-Marseille School of Economics and *Sciences Po*

November 13, 2015

Abstract

We study how departments' characteristics impact academics' quantity and quality of publications in economics. Individual time-varying characteristics and individual fixed-effects are controlled for. Departments' characteristics have an explanatory power at least equal to a fourth of that of individual characteristics and possibly as high as theirs. An academic's quantity and quality of publications in a field increase with the presence of other academics specialised in that field and with the share of the field's output in the department. By contrast, department's 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 at least for some publication measures but only when individual fixed effects are not controlled for. This suggests a role for individual positive sorting where these characteristics only attract more able academics. A residual negative sorting between individuals' and departments' unobserved characteristics is simultaneously exhibited.

Key words: Research productivity, Local externalities, Skills sorting, Peer effects, Co-author networks, Economics of science.
JEL classification: R23, J24, I23.

[a]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 and participants to various seminars and conferences for useful suggestions. This study was made possible and financed by the French Ministry of Research and Higher Education. Pierre-Philippe Combes is a researcher at the Cnrs, whose financial support is gratefully acknowledged.

[b]ThEMA, UMR 8184, Université Cergy-Pontoise, Cnrs, F-95000 Cergy-Pontoise. 33 boulevard du Port, 95011 Cergy-Pontoise cedex, France. c.bosquet@u-cergy.fr; http://sites.google.com/site/clementbosquet/.

[c]Aix-Marseille University (Aix-Marseille School of Economics), CNRS, & EHESS, GREQAM, 2 Rue de la Charité, 13002 Marseille, France and Sciences Po, Economics Department, 28, Rue des Saints-Pères, 75007 Paris, France. ppcombes@gmail.com; http://www.vcharite.univ-mrs.fr/pp/combes/. Also affiliated with the Centre for Economic Policy Research (CEPR).

1 Introduction

Every academic has an opinion about what makes a good department. Surprisingly enough, there are few hard studies quantifying this precisely, although possible implications for an optimal design of education and research policy are numerous. We focus here on the effect on individual publication records of both individual characteristics and a large set of department characteristics. We develop a careful strategy that controls for possible spatial selection of academics and missing variables. Both our identification strategy and the results we get for academic research are relevant more generally, for instance for knowledge-intensive industries where individual abilities play a crucial role (as in R&D departments of many manufacturing industries, or in finance, law and health services).

We propose to answer three sets of questions. First, what makes individuals productive, their own abilities or the firm or city where they work? Applied to science, this translates into whether academics publish more because they have better abilities (gender, age or any other individual characteristics possibly unobserved) and a more rewarding publication strategy (research field, number and location of co-authors) or because they are located in departments that provide a better local environment with stronger externalities? Using an exhaustive panel of French academics over 19 years (1990-2008) and their quality-adjusted publication records in EconLit, we find that both types of explanations are relevant. In particular, location matters and partly determine the individual quantity and quality of publications. When individual fixed effects are not controlled for, location explains almost as much as observed individual characteristics. When individual unobserved characteristics are controlled for, location still represents at least a fourth of the explanatory power of all individual characteristics. We observe a positive sorting of more able individuals according to observed characteristics in departments generating larger local effects. Maybe surprisingly, a negative sorting on unobserved individual characteristics is obtained, which slightly more than compensates the former positive sorting.

We also exhibit the role of a number of time-varying individual characteristics. The individual strategy of publishing with foreign co-authors and with a high number of co-authors per paper is the most rewarding in terms of publication quality. The latter conclusion suggests the presence of increasing returns to scale at the co-author team level. Larger teams reach better journals on average. By contrast, the number of publications per individual (a publication with n co-authors counts $1/n$) reduces with the number of co-authors. For example, having two co-authors instead of one reduces the number of publications per individual by around 30% but increases quality by between 8 and 25% depending on the quality measure. The average quality of academics' publications also increases with their number of publications, suggesting the presence of increasing returns to scale also at the individual level. Everything else equal (including the field of specialisation, all the aforementioned variables, and a department-time fixed effect), women and older academics publish less (for a given position type).

We then move to the dual question of the extent to which more productive firms or locations simply attract more productive employees or on the contrary generate more productive environments. That is, for academia, whether good departments (defined as those where the average quantity and quality of publications per academic is high) are those where highly-productive academics locate or those that generate more externalities. We find that local effects and the spatial sorting of academics matter equally when explaining the publication-based ranking of departments. Best departments are not only places where most able academics gather but also those generating stronger local effects.

Finally, and most importantly for the optimal design of firms, cities, or academic institutions, what are the channels of local external effects? We try to enter the black box of department effects and to assess the relative magnitude of the channels through which they operate. However, standard local variables usually correlated with the strength of local externalities have a very small explanatory power here when individual fixed effects are controlled for. This holds both for variables capturing within-field externalities ('localisation externalities') and overall department observed characteristics ('urbanisation externalities'). Many more departments' characteristics are significant when individual fixed effects are not controlled for but the explanatory power remains lower than standard estimates for market-activities.

Therefore, we exhibit a double puzzle. First, there is a far from zero explanatory power of local effects but it is neither driven by standard variables considered in urban economics nor by a number of departments' research characteristics. Some departments' unobserved characteristics seem to matter a lot. Second, academics sort positively but also negatively in better departments depending on the individual characteristics considered. For both puzzles, academics' mobility in France might be too low to have simultaneous individual and department fixed effects properly identified. In any case, this suggests to replicate our strategy on other countries to confirm, or not, these findings and to propose some further variables that could explain the role of both departments' and individuals' unobserved characteristics. Possibly, the teaching load and the quality of students, or the local administration efficiency, which cannot be controlled for here, are important. More subtle effects of the atmosphere in the department, sometimes mentioned by academics, could be at work too.

Still, some department variables have a significant impact, especially those relating to localisation effects. Being specialised in a field significantly and largely increases the quantity and quality of publications in this field even when individual fixed effects are controlled for. The mere presence of other academics in the department publishing in a field already generates positive externalities, increasing the average quality of another academic's publications in this field by 40%. This is reinforced when the field share of the department increases. For instance, the number of an academic's publications in a field increases by 6% when other academics' share of publications in this field doubles. At least when individual fixed effects are not controlled for, some department overall characteristics also have a decent explanatory power and exert a significant impact. Department size significantly increases quality but not quantity, with an elasticity close to usual estimates, at around 0.04. Proximity to nearby departments has a larger marginal elasticity at around 0.3 for quality and 0.1 for quantity. The presence of academics with co-authors abroad also present rather large marginal effects. Again, all these urbanisation effects are not significant when individual fixed effects are controlled for, localisation effects being the only ones to be robust to this inclusion.

The next section comments our contributions with respect to the literature. Section 3 presents the theoretical framework, the studied variables and the econometric strategy. Data is detailed in Section 4. Results are presented in Section 5 for the relative roles of individual characteristics and local effects for individual output, in Section 6 as regard the variance analysis of departments' performances, and in Section 7 for the determinants of local effects. Section 8 concludes.

2 Contributions to the literature

The paper first relates to the new surge of interest over the last fifteen years in the estimation of agglomeration economies (for reviews see Rosenthal and Strange, 2004; Combes and Gobillon, 2015). Assessing how much is gained by further spatially concentrating economic activities, by increasing regional specialisation or concentrating certain types of occupations, is indeed a crucial preliminary step to evaluate the efficiency of the spatial equilibrium or of some regional policies. A parallel can be made with universities or governments providing incentives to make academic departments larger or more specialised, for instance. Aghion et al. (2010) is a recent example of the more general concern for the optimal governance of universities as regards research. Another strand of literature examines the effect of peer and network effects on various school or labour market outcomes (for reviews see Jackson, 2011; Sacerdote, 2011). The focus is on the effect of the composition (by gender, ethnic or social origin) of a group or network on the individual outcomes of its members. Clearly, similar questions for departments arise as regards their optimal ratio of assistant professors to full professors, of women, of elder academics, or of certain academics who are particularly talented or connected to co-authors in other institutions. For both agglomeration and peer effects, the access to new data sets, typically encompassing information at the individual level (firms, workers, students/pupils or academics as here) and the search for relevant econometric strategies has greatly widened identification possibilities and clarified the direction of causalities, a trend we also follow here.

A couple of recent papers to which we compare our results consider a sub-set of the effects identified here. Waldinger (2012) concludes there were no localised peer effects in Germany among physicists, mathematicians and chemists under the Nazi regime. Kim et al. (2009) reach a similar conclusion, that being affiliated to one of the top 25 US universities no longer had an effect on the individual outcomes of academics in economics and finance in the 1990s, unlike in the 1970s and 1980s. This is confirmed for mathematics by Dubois et al. (2014), who show that the best departments do not necessarily stimulate positive externalities even if they are the most successful in hiring the most promising academics. Oyer (2006) shows that a top placement for new PhD economists has long-term benefits in terms of career but no benefit in terms of enhanced productivity, also supporting the view that top departments had no productivity spillovers (in the 1990s). Therefore, our conclusion that departments explain a rather large share of academics' individual productivity is somewhat discordant. It may be explained either by 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 analysis.

We decompose individual productivity into three components: the probability to publish in a given period, the number of publications and the average quality of these publications. We study the determinants of these three dimensions separately, at a detailed sub-field level within economics. The literature usually considers only the number of publications adjusted for quality as dependent variable, and for broader fields. We show that the effect of some variables does differ from one productivity dimension to another, which means that the optimal strategy for an individual or a department depends on the targeted dimension. Moreover, we also perform our estimations on two different indexes, more or less selective, of publication quality. Typically, the (individual and department) determinants of publications in top journals might differ from those in field journals. Our use of all EconLit publications and a corresponding impact factor index for all the 1200 EconLit journals enables us to study such differences, whereas the literature usually restricts itself to a small number of journals: 23 in Waldinger (2012), 41 in Kim et al. (2009), 98 in Dubois et al. (2014).

We take seriously the concern of a possible spatial sorting of talents that might influence the measurement of department effects. Such possible selection effects are mentioned in most papers on peer effects (Sacerdote, 2011) and are also central in recent assessments of agglomeration economies in market activities (Combes and Gobillon, 2015). We use individual data to properly tackle them. We run estimations on individual productivity considering both individual and department variables, and both department and individual fixed effects. 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 uses the premature death of superstars academics) do to identify peer effects.[1] However, we do provide estimates of local effects net of the possible academic spatial sorting, whether based on time-varying observed or time-constant

[1]This is also the strategy of Borjas and Doran (2012) who show that the inflow of Soviet mathematicians to the US after 1992 mainly substituted for local mathematicians, whose publications fell sharply while overall publications slightly increased. However, the effects of location within the US and of department characteristics are not simultaneously assessed.

unobserved individual effects, which corresponds to a pretty general model. The French context we use helps too. While initial affiliation, which is captured by the individual fixed effect, certainly relates to individual characteristics in France, most subsequent moves are clearly not driven by publication performance but rather either friendship connections or personal/family reasons. This is due to the features of the French academic system. For instance, moving does not affect wages since academics are civil servants paid the same everywhere. Moreover, the most frequent way for an assistant professor to become full professor, the largest source of movements in France, consists in passing the 'Agregation' contest after which spatial allocation is random for most candidates. Therefore, even if we do not benefit from a fully random experiment, we do not think that individual time-varying publication shocks conditional on individual and department-time fixed effects could largely affect location choices and then bias our evaluation of local effects. Not using a natural experiment also presents the advantage of providing more general results, and thus greater external validity. For instance, the co-authors of superstars, or the scientists dismissed by the Nazis, may not share the same characteristics as average current academics.

Our data set is not only exhaustive on all academics present in France but it also presents the second important advantage of reporting non-publishing academics. Studies that only use bibliometric sources necessarily ignore this group. This means that the department characteristics of these studies, computed on publishers only, are affected by measurement error, possibly large as non-publishers often represent 30% of a department (see Combes and Linnemer, 2003, for both European and US departments). For instance, department size is not the actual number of academics that we use here but the number of academics in the department who publish over the studied period (which is usually a short one). Last, we also have more individual characteristics, such as age, gender and position held, which can affect publication output and are usually absent from the data sets used in other studies. All of this can influence the results obtained and explain our new findings. We acknowledge that they remain estimated on French economists only.

3 Theory and estimation strategy

Publication output, individual and local effects

Let y_{ift} denote the publication yearly output (presented in Section 4) 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 share their time between many departments, let α_{idt} denote the share of academic i's output attributed to department d at date t.[2] 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 $Y_{dt} = \sum_f Y_{dft} = \sum_{i \in dt} \alpha_{idt}\ y_{it}$.

[2]90% of our academics have only one affiliation, in which case α_{idt} is equal to 1 for one departement only, and to 0 for all other ones. How α_{idt} is computed for others is detailed below.

We assume that y_{ift} is given by:

$$y_{ift} = e_{it} \, A_{d(i,t)ft} \ ,$$

where e_{it} is the academics' contribution to their own output ('individual efficiency') at date t and $A_{d(i,t)ft}$ is the field-specific contribution of department $d(i,t)$ to which academic i belongs at date t ('departement's externality').[3] The department's externality in field f at date t, A_{dft}, is then assumed to depend on two components, as follows:

$$A_{dft} = \text{Department's Overall Characteristics}_{dt} \times \text{Department's Field-Specific Characteristics}_{dft} \ . \tag{1}$$

'Department's Overall Characteristics' includes a first set of variables regarding the demographic structure of the department: the logarithm of the department size ($\text{Size}_{dt} = \sum_{i \in dt} \alpha_{idt}$, the department's full-time equivalent number of academics), the average age of academics, the share of women, and the shares of the different types of positions in the department.

Department size plays the role of the total employment density variable considered in standard estimations of agglomeration economies. It reflects possible local externalities from the overall size of the local economy. The list of possible positive effects from department size is long. To give but a few examples, academics in larger departments may benefit from larger and more intensive scientific interactions with their peers, from more numerous administrative or research assistance staff or the sharing of computing facilities, from greater bargaining power within the university or at the national level, allowing them to get more research funds, or from a better overall visibility that makes external network effects stronger. We cannot exclude the possibility of congestion effects causing size to also generate some negative effects. As most often in this literature, only the total net effect of size is identified.

For a given 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 suggested by Hellerstein et al. (1999), these composition effects must be introduced into the specification in the form of their proportions in the total number of academics in the department. This allows us to assess whether local externalities are stronger from different types of academics.[4] For instance, older academics may provide their experience, women may generate more externalities than men, and similarly for the various types of positions. A literature in industrial organisation (see for instance Besley and Ghatak, 2008; Auriol et al., 2012) studies the role of status incentives with implications for the optimal share of the different positions within the firms, which is another interpretation

[3] For academics belonging to more than one department, one such specification is assumed for each department, A_{dft} being specific to each of them in this case, and each observation is weighted by α_{idt} in the regressions. This assumes no specific externalities between the academic's departments. By contrast A_{dft} will be made a function of all nearby departments' output.

[4] Here again, as emphasised by Ciccone and Peri (2006), only a combination of the externality and of some possibly negative substitution effects is identified.

of such variables. The French academic system is rather complex in terms of possible academic status. We first distinguish between lower and upper positions (assistant professor versus full professor, which we name 'Rank A'). Moreover, some academics have research 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 of them are attached to domains other than economics (business, mathematics, etc.). Each type of position can generate more or less externalities since time devoted to research and incentives to cooperate locally both differ. We distinguish 10 different positions.

In order to identify different sources of department externalities, a second type of variables that relates more to the research characteristics of the department is considered. First, we evaluate the role of departments' field-specific research characteristics through two variables capturing the role of what urban economics names 'localisation economies'. These externalities emerge within industries, within research fields here, and are typically stronger the larger the industry/field locally. This relates to 'Department's Field-Specific Characteristics' that are considered in specification (1). As many fields are absent from many departments, we consider a non-linear specification for such effects. We consider first a dummy variable that takes value 1 when at least one academic in the department other than the one we consider has once published in the field:

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

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. We use here a time discount factor since the first publication of each academic, while we use a moving average over three years to compute the dependent variable as described in section 4. This better captures the academics' fields of specialisation than their publications in the field over three years only.[5]

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

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

Marshall (1890) first developed the idea that the relative size of an industry within the local economy can generate stronger local externalities for this industry, for instance when it uses specific local public goods, specific inputs or labour types. The same intuition can be developed for a field

[5]Strictly speaking, the variable, and some others below, depends on each individual and not the department only. Still, they capture the notion of external effects that the literature has in mind, and we prefer to keep these notations for the sake for clarity. For market activities, it is less crucial to exclude the own individual from the computation of local variables as, in any case, any of them is often negligible and the measure is almost not affected. Here, an individual can represent alone a large share of the department's output in a field, and this makes interpretation cleaner.

in academic research, 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 measure of publication at the field level allows us to test whether academics in departments that are specialised in a particular field publish more in that field.

Conversely, it has been argued since Jacobs (1969) that the overall diversity of the local activity can be beneficial to local productivity, especially in research-intensive sectors. According to this viewpoint, diversity encourages the cross-fertilisation of ideas between industries, thus strengthening innovation and growth. A large literature has attempted to test this idea by introducing a diversity index into estimated specifications, typically a Herfindahl index on the shares of each industry in the local economy. We proceed similarly here with the shares of each Jel codes in the department publications to obtain the departement's overall field diversity. A problem arises because, by construction, such a crude diversity index is highly correlated with department size. This is because departments with few academics have many Jel codes without any publications. To remove this size effect, which is absent from 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 would choose their Jel codes randomly. The diversity index net of the size effect is then:

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

where $\sum_f \left(\frac{\breve{Y}_{dft}}{\breve{Y}_{dt}}\right)^2$ is the randomly-generated Herfindahl index built by simulations.[6]

The second department's research overall characteristic concerns the physical proximity to other departments with which academics could interact or on the contrary compete. We capture this effect by an external research access variable, 'Research Access'. This tells us whether externalities emerge between different but nearby departments, as urban economics has highlighted over the last decade for market activities. 'Research Access' 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'.[7]

Departments also differ in terms of the co-authorship patterns of their academics. Having

[6]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 national 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 1,000 such procedures.

[7]Alternative specifications of the research access variable, with squared distance or square root of distance in the denominator have been tested and lead to qualitatively similar results. We keep the most standard one.

academics connected to foreign academic institutions can generate positive externalities through network effects for instance, which has been emphasised recently both in market activities and for 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 who is located outside 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').

There are debates about whether or not hiring top academics is a good strategy for the other academics in the department. We test more generally the possible effect that the department's heterogeneity in terms of academics' publication records has on individual publication records. This 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 Standard Deviation$(\tilde{y}_{it})_{i\in(d,t)}$ and 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.

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

We also include variables that reflect more the academic's individual research characteristics. To test for the presence of economies of scale within co-author teams, we introduce the average number of authors per article written by academic i at date t, ('Authors per publication'). This variable is central in many studies on the determinants of publication records that ignore the role of location but evaluate the returns to co-authorship following Sauer (1988). We also consider academic i's field diversity, ('Individual Diversity'), to assess whether academics benefit from knowledge acquired in other fields to publish in field f. This tests the presence of complementarities between fields at the individual level. The variable consists in the logarithm of the number of fields in which the academic has published:

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

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. Finally, some estimations also consider individual fixed effects that capture the role of any individual characteristic that is constant over time.

Econometric specifications

To separate agglomeration effects from the role of individual characteristics, we follow the econometric strategy proposed by Combes et al. (2008a). This is a two-step procedure in which, in the first step, the logarithm of individual productivity in a given field is regressed on individual effects (and possibly an individual fixed effect), a department-time fixed effect (β_{dt}) and the department characteristics that depend on the field:

$$\begin{aligned} \log\ y_{ift} \ &= \ \theta_i \ + \ \text{Individual Characteristics}_{it}\, \varphi \ + \ \beta_{d(i,t)t} \\ &\quad + \ \text{Department's Field-Specific Characteristics}_{d(i,t)ft}\, \eta \ + \ \mu_{ft} \ + \ \varepsilon_{ift} \ , \end{aligned} \tag{2}$$

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 (i.i.d.) across individuals, fields and periods.

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

$$\hat{\beta}_{dt} \ = \ \text{Department's Overall Characteristics}_{dt}\, \gamma \ + \ \delta_t \ + \ \upsilon_{dt} \ , \tag{3}$$

where δ_t is a time fixed effect and υ_{dt} is a random component at the department level assumed to be i.i.d. across departments and periods.

The main advantage of the two-step procedure is that it allows for a more general specification than could be made by directly considering department variables next to individual effects in a single step. The one-step estimation ignores unobserved aggregate shocks whereas the two-step procedure allows us to consider both individual and aggregated random components, ε_{ift} and υ_{dt}. Beyond extending the model, this also deals with the heteroscedasticity issues raised by Moulton (1990). Another advantage is that the first-step estimates for the department's field-specific characteristics and for individual characteristics are independent of the specification chosen for the department's overall characteristics effects. Changing the specification of the second step does not affect estimates from the first step.[8] The estimation of the second-step dependent variable in the first step creates measurement error issues. We deal with it by using Feasible Generalised Least Squares (FGLS) in the second step.

The literature agrees on the fact that considering individual fixed effects in the first step allows capturing the role of unobserved individual effects that could otherwise bias the estimation of local effects. For instance, Combes et al. (2008a) show that the impact of employment density on productivity is twice as low when individual fixed effects are introduced into the specification. However, to identify these effects separately from local effects, one needs large data sets and enough mobility of individuals between locations over time. Gobillon (2004) shows that the exact identification conditions are difficult to check empirically, and it is never done in practice in the literature. Given the pretty low mobility of French academics across departments and the much lower sample size by comparison with standard employer-employee data set, it is difficult to be sure that individual and local effects are properly identified here. Notice also that information about age, gender, position, fields and connections are not often available in other studies from the literature, which makes it more important for them to control for individual fixed effects than here. We present the two sets of estimations, with and without individual fixed effects, and provide comments when conclusions differ between the two.

Introducing field-time fixed effects corresponds to an interpretation issue. If one assumes that differences in publication records between fields at the world level are only a matter of fashion, and not of talents and true differences of productivity between academics and departments specialised in different fields, then one should remove world wide differences by introducing field fixed effects to focus on spatial variability independently of specialisation choices. Conversely, if one believes that a higher number of publications in a field at the world level truly corresponds to higher productivity in that field, then field fixed effects should not be introduced into the specification. We adopt the former position here, and introduce field fixed effects. This is the most conservative strategy as it removes the role of possible correlations between local effects and field specialisation choices. It is also the viewpoint adopted in urban economics, which systematically considers 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.

Finally, we need to comment on possible endogeneity concerns. Beyond structural approaches, the only way to deal with them in the first-step estimation consists in using natural experiments,

[8] As shown in Bosquet and Combes (2015), most of the second-step results (determinants of the departments' externalities) are nevertheless robust to a specification in one-step.

as proposed by Azoulay et al. (2010) with the premature death of stars or by Waldinger (2012) with the dismissal of scientists by the Nazi government.[9] Then one has to believe that the natural experiment is not correlated with any co-variate and, most importantly, that estimates obtained from the natural experiment would also hold in other circumstances. Alternatively, most of the literature does not deal with this possible endogeneity. We argue that considering both individual and department-time fixed effects should remove most if not all sources of endogeneity, as is the case when location choices are based on location characteristics only and not on individual temporary shocks. Since this seems to be a reasonable assumption for academics, and as a choice between two evils, we follow this strategy here.[10,11]

Endogeneity biases can also be present in the second-step estimation. For instance, Combes et al. (2008a) show that estimates of agglomeration economies in market industries decrease by around 20% when local variables are instrumented. The department's overall characteristics are endogenous when academics are mobile and have their department choices driven by the departments' publication output. Given the number of department characteristics we consider here, instrumenting them all would be difficult and not make much sense, particularly with respect to possible weak instrument issues. We are not aware of any other paper on agglomeration and peer effects in academia that proposes an instrumentation of department characteristics. Therefore we leave this issue for later contributions. Most estimation issues in this literature, which are also present here, are further discussed in Combes and Gobillon (2015).

Decomposing output

Academic i's productivity can be decomposed as follows:

$$y_{ift} \equiv \mathbb{1}(\text{Quantity}_{ift} > 0) \times \text{Quantity}_{ift} \times \frac{y_{ift}}{\text{Quantity}_{ift}} \tag{4}$$

where Quantity_{ift} is the number of publications of academic i in field f at date t. 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 academic i in field f at date t. One contribution we make consists in separately studying the determinants of each of these components of academics' publication output. For instance, we can state whether a department's characteristic affects the probability to publish, the quantity or

[9] Or by Borjas and Doran (2012) with the inflow of Soviet mathematicians to the US after 1992, except that those effects are not local].

[10] A third strategy would involve first specifying a model for the academics' choice of department and then estimating our two-equation model conditionally on that choice. However, exclusion restrictions must then be satisfied, namely finding variables that explain the department choice but not the publication record. We cannot see any candidate for this, since even family characteristics for instance can explain the latter.

[11] We also run regressions without connection (USA and non-USA) variables, which we consider to be the most endogenous variables of our analysis beyond the department choice. Results are qualitatively similar.

the quality of publications in the same direction, or if they are substitutes to each other. This is important from a policy perspective. More precisely, we assume that specifications (2) and (3) hold for each component of the individual publication record in equation (4).

Finally, first-step estimations need to weight 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}. Second, the academics output is split between all their publications' Jel codes and we have, for each academic, date and department, one observation for each field with weight $\frac{\text{Quantity}_{ift}}{\text{Quantity}_{it}}$. To take both effects into account, we weight by α_{idt} each observation in the probability to publish first-step estimation and, because we take logs, by $\alpha_{idt}\frac{\text{Quantity}_{ift}}{\text{Quantity}_{it}}$ each observation in the quantity and quality first-step estimations, in order to give the same weight to each academic in all regressions. OLS are used even for the probability to publish due to the presence of many fixed effects. We checked that a probit estimation leads to similar results.

4 Data

Measure of output

We measure the publication output of academic i in field f at date t as a weighted sum of his publications in field f listed in EconLit[12] over period τ. In most tables, τ corresponds to years $t+1$, $t+2$, $t+3$ and the output at date t is a moving average over these three years. This choice is standard in the literature, adopted for instance recently by Ductor et al. (2014). It seems to correspond to the average reality of the profession in terms of the time needed to write papers and publication delays.[13] As a robustness check and because such a choice is both somewhat arbitrary and could result in the autocorrelation of residuals, we also present in Bosquet and Combes (2015) regressions where τ is reduced to year $t+2$. The results are very similar to those based on the three-year moving average.

The previous discussion regards the dependent variable. Because the academics' field choice in a given year may not perfectly reflect their average field of specialisation, we prefer to compute the individual and local externality explanatory variables that depend on field on a longer time span. In that case, we discount over time the past publications in the field: All publications are taken into account but most recent ones count more than older ones.[14]

[12]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.

[13]Note also that the list of the department's academics at date t is established in September of that year.

[14]The discount factor t' years before t correspond 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 for 1, year $t-1$ ones count for 0.94, year $t-2$ ones count for 0.75, and so on, and, for instance, after 10 years they count for 0.20, and after twenty years for 0.10 only.

Beyond the number of publications, some further weights allow us to assess the quality of the publications. Each publication p is first weighted by the quality of the journal, $W(p)$, 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[15] and from Google Scholar citations.[16] 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 publications of a journal get the same weight independently 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 journal) and CLh which is more selective (ranging from 100 for the Quarterly Journal of Economics to 0.0007 for the last 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.

As is common practice in the literature, article a is also weighted by the inverse of its number of authors, $n(a)$. Since the publication output of a department is the sum of the 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 by using the average number of authors per publication as one of the independent variables.

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

Finally, 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 world level (see above discussion). We use Jel codes at the first digit level (letter) and we ignore the fields "Y - Miscellaneous Cate-

[15]http://www.webofknowledge.com/
[16]http://scholar.google.com/

gories" and "Z - Other Special Topics". We also slightly modify the codes C and D by merging code 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 to be 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:

$$y_{ift} = \frac{1}{\mathrm{Card}(\tau)} \sum_{a(i)\in\tau} \frac{W(a)}{n(a)} \frac{p(a)}{\overline{p}} \frac{j_f(a)}{j(a)}$$

where $\mathrm{Card}(\tau)$ is the number of years in period τ.

Academics and universities

The French Ministry of Higher Education and Research, Cnrs and Inra[17] provided us with the list of academics in economics in France for the period 1990 to 2008. Each academic is affiliated to at least one university department or to a Cnrs or Inra research centre. We merge together these affiliations at the university level to obtain what we call a "department". This is either an economics department, if that is the only body to which economists are affiliated in a university (which is the majority of the cases), or the aggregation of all departments or research centres containing economists in the university. We believe that this notion of slightly aggregated economics departments better matches the French reality of academic research than a more detailed approach. Robustness checks using detailed affiliations lead to fully consistent results, which are presented in Bosquet and Combes (2015).

The French system allows for multiple affiliations and around 10% of academics belong to 2 or 3 departments. In this case, each department is equally weighted (parameter α_{idt} in above definitions). For a few cases of academics who have positions both in France and abroad, we use their CV to evaluate the share that should be attributed to the French department. Last, we want the analysis to focus only on academics that can really be considered as forming a local group of academics working together. Therefore we only keep departments larger than 4 full-time equivalent academics, excluding economists that are isolated in universities without a real economics department. We have performed a variant keeping only departments larger than 9 full-time equivalent academics and obtained very similar results, which are given in Bosquet and Combes (2015).

The data set includes a number of individual characteristics such as gender, age and position. We merge this with EconLit by surname and the initial of the first name. First names are too

[17] 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, respectively.

badly recorded in EconLit to be used in full. Keeping only the initial very slightly increases the number of academics with identical names; we deal with their publication records manually. For all academics and for each year between 1990 and 2008, we obtain then a data set with their individual characteristics, departments of affiliation, and publication record with weighted outputs.

Descriptive statistics

Using a three-year moving average for the publication output prevents us from considering the years 2006, 2007, and 2008. Therefore, time observations in our data set span from 1990 to 2005. Both the number of academics and the number of departments are monotonically increasing over time, from 1,753 academics and 69 departments in 1990 to 2,914 academics and 81 departments in 2005. Over the 16 years of our panel, this leads to 38,742 academic-year observations and 1,267 department-year observations.

Table 6 panel (a) in Appendix A 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 corresponding to research academics. The line 'Rank A' reports that 35% of academic-year observations correspond to a rank A position, *i.e.*, equivalent to full professor as opposed to assistant professor.

Not all academics publish over a given three-year period. The line 'Publisher' in Table 6 panel (a) reports that one third of them 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 6 provides descriptive statistics on the sub-group of academics who have 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 line 'Quantity' in Table 6 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 little, but partly due to the fact that many academics do not publish any papers at all. Conditional on having at least one publication over the three-year period, we read in Table 6 panel (b) that the average number of publications is three times higher, corresponding to, for instance, one publication alone and one publication with a co-author every three years. As regards quality, we also confirm the large disparities existing among academics, a well-documented fact since Lotka (1926). The mean publication is worth the equivalent of one publication per year in the 150th journal but the median publication is lower, around the 350th journal. By contrast, the top decile average quality publication corresponds to one publication per year in the 50th journal or one publication in a top 5 journal 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 the publishers 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.

Regressions are performed at the field level. Since there are 18 possible fields, the 38,577 academic-year observations translate into 694,386 field-academic-year observations, which is then increased by the fact that some observations are duplicated for academics with multiple affiliations (as explained in Section 3). As a result, the number of observations we have in the first-step estimation for the probability to publish is 771,426. As appear in Table 1 column 1, this reduced to 760,014 when department-year fixed effects are considered because of department-year without any publisher. This further reduce to 425,394 when individual fixed effect are considered (Table 1 column 2) because of some academics without publication. Because some academics do not publish at all, and, more importantly, because they do not publish in all fields, many of these observations correspond to zero publications (in a given field). There are 'only' 39,004 non-zero observations, which are the observations used for the first-step quantity and quality estimations. The line 'Individual Diversity' in Table 6 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 has 5 fields. At the national level, 'Microeconomics' is largely the most represented field in France with 16.8% of the number of publications. This is larger than its share for EconLit as a whole, which is 10.2%. Then, there are 10 fields each representing more than 4%.[18]

Panel (c) in Table 6 reports descriptive statistics at the department level. The average department has 31.2 academics who are 45 years old on average, 24% are women, 34% have rank A positions, and 33% are publishers. The figures are comparable to the averages over all academics. Importantly, all variables present quite a lot of variations between departments, which is in particular observed for the publication output. The average department has 5.4 publications per year, 0.18 per academic, and the average quality indexes are in the same ranges as for individual academics. The line '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's cumulative output represents 20% of this one. Departments are therefore 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

[18]'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%).

values even at the top decile, meaning that departments are less diversified than they would be with random Jel code choices.

Finally, Table 7 in Appendix A presents simple correlations between the variables at the department level. First, quantity and quality are largely positively correlated even for the top quality index. Those departments that publish more also produce higher quality publications and no trade-off seems to take place between the two. This is in keeping with what Combes and Linnemer (2003) find 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 with the share of academics having co-authors abroad and in the USA (connection variables), and again large for heterogeneity, which is positively and negatively correlated with quantity and quality respectively. The correlation of size with quantity is not very strong but it increases for quality, and even more for top quality. We must now investigate whether these correlations are driven by the fact that rank A position researchers, or researchers with high 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. This is the purpose of the econometric analysis developed in the next sections.

5 Productive academics: Individual abilities versus local effects

This section studies the determinants of individual productivity and assesses the relative weight of individual and department effects. We regress individual productivity in a specific field on individual characteristics that relate 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. Columns (1) and (2) ('Publishing') in Table 1 concern 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') concern the log of the average publication quality using the standard and top journal quality indexes respectively. For each output measure, the first column does not include individual fixed effects, which are included in the next column.[19]

Before turning to the effect of each variable, let us start with some variance analysis. A first remark regards the large increase of the R^2, by 17% for the probability to publish, by 22% for the quantity published, 32% for the publication quality and 28% for top quality when department

[19]The effects of women and age are not displayed in columns (2), (4), (6), and (8), since it is not possible to separately identify them from individual and year fixed effects. There are some variations for the rank A and position variables within individual but, as they are very rare, we prefer to remove these variables too when individual fixed effects are considered. Results are virtually not affected when they are included.

Table 1: Determinants of individual publications

	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Women	-0.016^a		-0.120^a		-0.069^a		-0.274^a	
	(0.001)		(0.009)		(0.009)		(0.024)	
Age	-0.005^a		-0.034^a		-0.023^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		0.135^a		0.541^a	
	(0.001)		(0.009)		(0.009)		(0.024)	
Authors per publication			-0.948^a	-0.923^a	0.187^a	0.193^a	0.511^a	0.543^a
			(0.011)	(0.015)	(0.010)	(0.013)	(0.027)	(0.033)
Individual diversity			-0.096^a	-0.131^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	0.192^a	0.307^a	0.084^a	1.129^a	0.341^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	0.610^a
			(0.015)	(0.019)	(0.014)	(0.016)	(0.038)	(0.042)
Dep. field presence	0.064^a	0.123^a	0.346^a	0.335^a	0.116^a	0.088^a	0.363^a	0.319^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	0.099^a	0.088^a	0.037^a	0.021^a	0.134^a	0.085^a
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.003)	(0.009)	(0.008)
Position FE	Yes	No	Yes	No	Yes	No	Yes	No
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes
Field-time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department-time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.07	0.13	0.33	0.54	0.37	0.65	0.46	0.72
Observations	760,014	425,394	39,004	39,004	39,004	39,004	39,004	39,004

Standard error between brackets. ^{a}, ^{b}, ^{c}: significant at the 1%, 5% and 10% levels respectively. 'Publishing', 'Quantity', 'Quality', and 'Top quality': OLS estimates. Variables are defined in Section 3.

effects are considered, compared to estimations (not reported here) where department characteristics are not introduced in the model (and individual fixed effects are introduced in neither cases). Moreover, the explanatory power of the model also largely increases when individual fixed effects are further introduced. It is now 50 to 80% higher than when individual variables and department effects only are included and reaches levels comparable, even if a bit lower, to what is obtained in standard individual wage or productivity equations, with R^2 between 0.54 for quantity and 0.72 for top quality.[20] A last conclusion is that the model better explains the average quality of publications than the number of publications, even more so when a top quality index is considered. The publication quality relates more to individual and department characteristics than does the number of publications, for which the random component is larger. This would probably make sense to all academics, since publishing in good journals needs more specific skills, captured by the model, than just publishing.

[20]The R^2 is lower for the probability to publish but we have no benchmark value for that and this obviously arises, at least partly, from the large number of zero observations.

To obtain more precise insights as regards the sources of output variations, a more detailed variance analysis is provided in Table 2. First, the 'Std. dev.' columns report the standard deviation of the effect of a variable or of a group of variables for the quality estimations presented in Table 1 columns (5) and (6). The higher it is 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, and most importantly, a variable or group of variables has a large explanatory power when its effect is largely correlated with the dependent variable. This is reported in the 'Correlation' columns.[21]

Table 2: Variance analysis of the individual publication quality

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Quality	0.452	1.000		0.452	1.000	
Individual effects	0.159	0.452	0.193	0.340	0.753	-0.079
Individual fixed effect	-	-	-	0.331	0.636	-0.097
Observable individual effects	0.159	0.452	0.193	0.178	0.253	0.029
Women	0.016	0.050	0.016	-	-	-
Age	0.082	0.132	-0.025	0.163	0.134	0.015
Position	0.045	0.151	0.060	-	-	-
Rank A	0.038	0.101	0.136	-	-	-
Authors per publication	0.040	0.215	0.092	0.041	0.215	0.035
Individual diversity	0.004	0.141	0.120	0.001	0.141	0.031
Non-USA connection	0.061	0.276	0.122	0.017	0.276	0.007
USA connection	0.075	0.317	0.174	0.031	0.317	0.038
Department effects	0.154	0.434	1.000	0.110	0.188	1.000
Department-time fixed effect	0.152	0.418	0.987	0.109	0.171	0.991
Department-field-time effects	0.025	0.135	0.171	0.015	0.142	0.135
Field presence	0.016	0.065	0.115	0.012	0.065	0.032
Specialisation	0.028	0.082	0.086	0.016	0.082	0.101
Field-time fixed effect	0.091	0.302	0.121	0.063	0.257	0.024
Residuals	0.359	0.795	0	0.268	0.593	0

The table presents the variance analysis of the estimations reported in Table 1 columns (5) and (6). All variables are first centered with respect to their annual mean. Therefore all the variables are detrended and the variance analysis is performed in the within-time dimension. The 'Std. dev.' columns report the standard deviation of the effect of a variable or a group of variables. For the first line, it reports the standard deviation of the dependent variable. The 'Correlation' columns report the correlation between the effect of a variable or a group of variables and the dependent variable. The 'Sorting' columns report the correlation between the effect of a variable or of a group of variables and the department overall effect. Line 'Individual effects' corresponds to the simultaneous role of both individual observed and fixed effects if any. Line 'Observable individual effects' correspond to the role of all observed individual effects together. Line 'Department effects' corresponds to the role of both the department-time fixed effect and all department-field-time effects. Line 'Department-field-time effects' correspond to the simultaneous role of 'Field presence' and 'Specialisation'.

A first crucial remark is that the variance analyses considering or not individual fixed effects, the left- and right-hand sides of the table respectively, largely differ. According to the model

[21] The variance analysis is computed on variables centered with respect to their annual means and therefore performed in the within-time dimension, to focus on spatial variations. See Abowd et al. (1999) for details on this type of variance analysis.

without individual fixed effects, the explanatory power of individual and department effects are very similar, both in terms of the standard deviation of the effects (each at around a third of the standard deviation of the dependent variable) and of the correlation with the dependent variable (slightly lower than 0.5). The two groups of variables would contribute to explaining the individual publication output to the same extent.

By contrast, individual effects have a much larger explanatory power when individual fixed effects are considered. From the right-hand side of Table 2, we can see that the standard deviation of individual effects is now two-third of the one of the dependent variable, three times as large as that of department effects. The correlation with the dependent variable is four times larger for individual effects. This means that some unobserved individual effects do significantly influence the quality of publications and part of them are captured by the department-time fixed effects when individual fixed effects are not introduced in the specification. Tables 8, 9 and 10 in Appendix B reproduce the variance analysis for the probability to publish, the publication quantity and top quality respectively. Similar conclusions are obtained for all variables.

To sum up, and keeping in mind the observation in Section 3 that individual fixed effects cannot always be properly identified separately from department effects if there is not enough mobility between departments – which could be the case here –, the lower bound for the explanatory power of department effects is around a fourth of the explanatory power of individual effects. However, at the upper bound, without individual fixed effects but still with a pretty large set of individual observable characteristics, department effects could explain as much as individual effects. This contrasts with the findings of the literature. Waldinger (2012) does not find any peer effects among physicists, chemists and mathematicians in Nazi Germany, and this is also the conclusion of Dubois et al. (2014) for modern-day mathematicians. Kim et al. (2009) find that the effect of being in a top 25 economics and finance department gradually disappears between the 1970s and the 1990s in the USA. Notice that these authors comment the fact that department effects are or are not significant, but they do not discuss their global explanatory power, which we do here. This is not exactly the same point of view, and we believe our approach to be relevant to 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, as discussed above, the fact that mobility is high enough to identify both individual and department fixed effects is difficult to test formally, and none study does it. A last explanation could be that research habits differ between a European country like France and the USA, both in terms of research technology (e.g., the intensity of internet use for collaborations) and in terms of institutional design. For instance, the possibility of individuals capturing their publication performance is considered to be lower in most European countries where wages and positions are much less closely tied to publication records.[22] All of these factors

[22]Combes et al. (2008b) document this for France.

could affect the relative role of individual and department effects.

Another crucial result emphasised by Combes et al. (2008a) regards the sorting of workers across space. More able workers locate in more favourable locations, where location effects reflecting local externalities are the strongest. From the econometric point of view, it is important to assess whether department effects would be biased if individual effects were ignored. From the policy point of view, it is interesting to know whether more productive academics gather in some departments, which makes these departments appear to be more productive than others, or if locations choices are rather arbitrary but some departments generate stronger externalities, and thus are more productive. The 'Sorting' columns in Table 2 for publication quality, and in Tables 8, 9 and 10 in Appendix B 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 department overall effects. It is typically found to be positive for observed individual effects. Workers with individual observed characteristics that make them publish more and with a higher quality are located in the departments that provide larger external effects.[23] When individuel fixed effects are not controlled for, the correlation with the department overall effects of all individual observed effects together, at 0.19, is large. Here the correlation between observed individual effects and department overall effects is smaller when individual fixed effects are controlled for, at 0.03 and a negative correlation at -0.10 is observed between individual fixed effects and department overall effects. Therefore we find that the spatial sorting of academics is rather large and positive on observed characteristics but that it much decreases when individual fixed effects are considered, the overall sorting turning to negative (at -0.08) due to a more than compensating negative sorting on unobservable characteristics. This is a striking conclusion, even if we cannot exclude that such a negative sorting on unobserved individual characteristics results from a weak separate identification of individual and department fixed effects.

The sorting results for quality are intermediate between those obtained for quantity and top quality, which are reported in Appendix B. As regards quantity, there is almost no sorting on observed characteristics (correlation at +0.01 and -0.04 without and with fixed effects respectively) and a pretty large negative sorting as regards individual fixed effects (correlation at -0.17). For top quality, there is a positive sorting on observed characteristics (+0.24 and +0.04 without and with fixed effects respectively), thus stronger than for quality, and a pretty small negative sorting for individual fixed effects (-0.04). Unfortunately, none of the papers assessing the magnitude of peer effects in science computes such correlations between individual and department effects, which would have allowed us to compare these important conclusions with those in other fields or for other periods.

We turn to the role of each individual variable. From Table 1, women and older academics

[23] Age presents a negative correlation with department fixed effects but it plays negatively on output, so again we have that academics with the good characteristic (being younger) locate more in better departments. The women dummy is the only exception but its correlation with the dependent variable is small.

appear to publish less. This is consistent with previous findings in the literature, even more so since we control here for the type of position held. Once a given position is achieved, for instance becoming a full professor, the number and quality of publications decreases with age. Part of the effect might also result from a cohort effect (previous generations had weaker incentives to publish than younger academics).

As detailed in Bosquet and Combes (2015), expected results are obtained for the impact of the various positions. The higher the rank (professor, research professor, and even more so Insee or *Ponts-et-Chaussées* Engineers as opposed to assistant professors or research junior fellows) and the more time allocated to research (research versus teaching positions), the larger the published quantity and quality, which is also the case for the academics purely in economics compared to those involved in business or mathematics for instance. Therefore, even if part of promotions in France does not relate to publications, as established by Combes et al. (2008b), those who get better positions do publish more on average. Note that our purpose here is not to give a causal interpretation to such variables but to control at best for individual abilities when estimating the role of departments.

Interestingly, we control not only for some of the standard 'ability' variables considered in wage or productivity equations, like gender, age or position (which plays the role of occupation), but also for variables characterising the research of academics. The variables that have the largest correlation with the dependent variable are the connection variables. Academics who have co-authors abroad (both in the USA and elsewhere) also publish more and they are largely over-represented in better departments. Again, the direction of the causalities cannot be stated here.

Then, the average number of co-authors per publication also has a large explanatory power. Its impact on published quantity is largely 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 one 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 if they would work alone. However, the average number of co-authors has a large positive effect on the average publication quality, which is almost three times larger for top quality. Therefore, a larger number of co-authors decreases the number of publications equivalent written alone but increases their quality. There is a trade-off between the two, and only an analysis such as our can identify the two effects separately. For instance, according to the estimates controlling for individual fixed effects (but magnitudes are very similar without), an academic who has on average two co-authors instead of only one (per publication) has 31.2% less publications but their average quality is 8.1% higher and their average top quality is 24.6% higher.[24] Combining the two effects, having two co-authors instead of one decreases the quality-weighted quantity of publications (a measure frequently found in the literature, consisting

[24] $1.5^{-0.923} - 1$, $1.5^{0.193} - 1$ and $1.5^{0.543} - 1$, respectively.

here in multiplying the quantity by the average quality) for both the standard and top quality index but less in the latter case. Therefore, one should not expect to publish more, even in decent journals, thanks to co-authoring, but possibly to reach 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 to a negative effect of co-authorship on publication quality. Dubois et al. (2014) identify an overall negative effect of co-authorship for mathematicians on their citation-weighted publication index but the effect of collaboration with co-authors having different specialisation is positive. Ductor (2015) finds a negative effect of co-authorship for economists between 1970 and 2011 that turns positive when unobserved heterogeneity and endogenous co-authorship formation are taken into account.[25] Overall our results are more consistent with the studies enlightening a negative impact of co-authorship but it is the only one that estimates the effect of co-authorship on the average quality of publication, and it is found to be positive, the more the more selective the journals are.

We also find that having a higher diversity of research fields decreases the number of publications of academics but it has no impact on the publication quality and possibly a positive impact on top-quality publication, which disappears however when individual fixed effects are controlled for. Dubois et al. (2014) find a positive effect of field diversity for mathematicians.

Finally, we report in Bosquet and Combes (2015) estimations of quality determinants that control for quantity. This allows us to test for the presence of increasing returns to scale for quality at the individual level now (as opposed to the co-author team level assessed through the number of co-authors). This is usually omitted in the literature. There are indeed increasing returns for average quality with respect to the number of publications, 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 higher by 6.6% and a top publication quality higher by 40.8% when individual fixed effects are not controlled for.[26] It also follows that the impact of the number of co-authors per publication on quality is even stronger when quantity is controlled for.

Notice that all these results hold within field (within Jel codes), since we control for Jel code fixed effects. Jel codes appear to have a pretty large explanatory power, especially as regards publication quality and top quality. This reflects the fact that all fields are not equal in terms of publication opportunities. To the best of our knowledge, nobody has yet been able to assess whether this is due to a pure 'fashion' effect (some topics are more fashionable, which makes them easier to publish) or to some selection effects (more able academics self-select in certain fields or those fields attract more able academics). It is not the purpose of the present article to tackle this difficult question. Still, in terms of interpretation, department effects are estimated here net of the

[25] See Section 3 for a discussion about endogeneity concerns in the first-step estimation.
[26] $2^{0.092} - 1$ and $2^{0.494} - 1$ respectively.

direct role of academic composition in terms of research fields.

6 Productive departments: Sorting versus local effects

The previous section studies how much location matters for the individual productivity of academics. We now move to the dual question of the extent to which the department's composition in terms of individual characteristics shapes the department's performance, compared with the presence of local externalities. To this end, we repeat the previous variance analysis using the same estimates as before (Table 1) but, before, we aggregate the variables by department-year. In other words, we now want to study the determinants of the departments' relative performance. Table 3 presents the results for quality while Tables 11, 12 and 13 in Appendix C present the results for the probability to publish, quantity, and top quality respectively.

Table 3: Variance analysis of average publication quality at the department level

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Quality	0.426	1.000		0.426	1.000	
Individual effects	0.160	0.564	0.153	0.375	0.743	-0.239
Individual fixed effect	-	-	-	0.380	0.622	-0.275
Observable individual effects	0.160	0.564	0.153	0.166	0.254	0.090
Women	0.014	0.054	0.012	-	-	-
Age	0.077	0.079	-0.111	0.146	0.064	0.043
Position	0.062	0.208	-0.015	-	-	-
Rank A	0.037	0.206	0.140	-	-	-
Authors per publication	0.045	0.358	0.152	0.047	0.358	0.144
Individual diversity	0.004	0.238	0.203	0.001	0.238	0.050
Non-USA connection	0.050	0.414	0.175	0.014	0.414	0.030
USA connection	0.052	0.484	0.234	0.021	0.484	0.069
Department effects	0.351	0.880	1.000	0.289	0.457	1.000
Department-time fixed effect	0.351	0.877	0.998	0.289	0.450	0.998
Department-field-time effects	0.021	0.093	0.068	0.016	0.133	0.036
Field presence	0.036	0.149	0.106	0.027	0.149	-0.006
Specialisation	0.033	-0.104	-0.073	0.018	-0.104	0.040
Field-time fixed effect	0.085	0.311	-0.008	0.063	0.249	-0.070
Residuals	0.001	-0.013	0	0.001	-0.033	0

Same notes as for Table 2. The effect of variables or group of variables is averaged by department-year before calculating the standard deviations and correlations.

Both Kim et al. (2009) for economics and business and Dubois et al. (2014) for mathematics argue that a good department nowadays is primarily made by bringing together academics with good individual characteristics without much local effects taking place. This is not what we find here. Even according to the specification that includes individual fixed effects, in the right-hand side of Table 3 for quality, the standard deviation of all individual effects together is only slightly higher than the standard deviation of the department effects all together. Still, the correlation with the average department quality of publications is twice as high for the former. When individual fixed effects are not controlled for, both the standard deviation of department effects and their correlation with the dependent variable are now higher than those for individual effects. From Tables 11, 12, and 13 in Appendix C, similar conclusions are obtained for quantity and top quality, to slightly larger and smaller extents respectively. Overall, the characteristics of academics explain at best slightly more publication disparities between departments than the external effects at play in these departments. Department effects are strong and therefore important in explaining the ranking of academic institutions.

When the correlation between individual and department effects is positive, as for market activities in France according to Combes et al. (2008a), disparities in terms of individual characteristics and of strength of externalities cumulate and generate pretty large productivity disparities between locations. As reported in columns 'Sorting' of Tables 3, 11, 12, and 13, this is partly reversed here. As regards quality for instance, the effects of observed individual characteristics are almost all positively correlated with the departments' effects, and the correlation with them all together is 0.09. But individual fixed effects are largely negatively correlated with them, at -0.27. All individual effects together present a correlation with department effects at -0.24. Departments' externalities and the positive sorting of academics according to observed characteristics are compensated by a pretty strong negative sorting of academics according to unobserved characteristics when explaining the ranking of departments. This result will be further investigated and discussed in section 7.

The hierarchy of individual characteristics in terms of explanatory power of department disparities is similar to the one for individual disparities. Connections to co-authors abroad play the larger role, followed by the number of authors per publication, and a somewhat large positive sorting with department effects occurs for them as we have just mentioned. Importantly for policy implications, age, position type and rank A share, which results from a mix of the department's, Ministry of Higher Education and individual decisions, also play some role but sorting is lower in that case. Individual diversity, and the gender composition to a lower extent, do not vary much between departments and their impact on average department quality is lower. As presented in Appendix C, such conclusions are broadly confirmed for the number and top quality of publications with some variables sometimes having a slightly stronger or weaker explanatory power for one of these publication measures.

Kim et al. (2009) argue that the declining role of economics and finance department externalities is a recent trend that gradually emerged in the eighties and nineties, by comparison with the seventies. To assess whether such a trend is also present for French economics departments, and because our panel spans a fairly long period of time, we repeat all our analyses for two sub-periods separately, 1990–1997 and 1998–2005. These periods are interesting because it was only at the end of the nineties that, first, the internet started to be systematically used to search for literature and circulate papers and second, that publications in peer-reviewed non-French journals became the norm when evaluating academics. Both may have contributed to a change in the effect of departments on publications. However, as the results reported in Bosquet and Combes (2015) show, all the conclusions we draw in this paper are broadly stable between the two periods and correspond to what is found over the period 1990-2005. The minor observed changes relate to the lower productivity of women, which is more pronounced over the 1998-2005 period, a slightly larger positive effect of the number of co-authors per publication for this period, and to the positive sorting on individual observed characteristics, which is slightly larger over the 1998-2005 period. In particular, we do not observe any strong decline in the strength of department effects over time.

7 The channels of department externalities

The last step in our analysis is to identify the channels through which department externalities operate. This consists in studying both the impact of the field presence and specialisation variables in the first-step estimation and the determinants of the department fixed effect in the second step. Results are provided in Table 4, where the first two lines are simply reported from the first-step estimations (Table 1) and following lines correspond to the estimation of specification (3) on the panel of department-time fixed effects. Since the dependent variable is estimated in a first step, we must correct for measurement errors on it. We do so using Feasible Generalised Least Squares but obtain similar results using alternative strategies, as often met in the literature (see Combes and Gobillon, 2015).

As testified by the low within-time R^2 values reported at the bottom of Table 4, the broad conclusion is that department effects are very difficult to explain when individual fixed effects are controlled for. Even when individual fixed effects are not controlled for, the explanatory power of all department overall characteristics together is lower than what is usually obtained for market activities. The full variance analysis when individual fixed effects are not considered is provided in Table 5. It appears that departments' observed characteristics have an explanatory power lower than half of the one of department fixed effects. The largest explanatory power is obtained at the two extremes of the publication measures. It is at around two-third for the probability to publish and slightly above one half for top quality. Observed department characteristics matter for either just being a publisher, at the lower end of the academic skills distribution probably, or for reaching top journals, at the other extreme of the distribution. There are less important for standard quality and matter little for the number of publications.

Therefore, a general conclusion is that a large share, if not almost all, of department effects are not explained by the department's characteristics we consider, even if the list is pretty long. Remember that we do not want only to emphasise that local effects are present in academia, as the literature usually does and which we emphasize here even when individual fixed effects are controlled for, but we would like to exhibit whether the standard local characteristics considered

Table 4: The effects of department characteristics

	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field presence (from 1st step)	0.064^a	0.123^a	0.346^a	0.335^a	0.116^a	0.088^a	0.363^a	0.319^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	0.099^a	0.088^a	0.037^a	0.021^a	0.134^a	0.085^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	-0.001	0.035^a	-0.013	0.057	-0.031
	(0.001)	(0.002)	(0.011)	(0.015)	(0.013)	(0.014)	(0.035)	(0.034)
Women	0.009	0.007	-0.036	0.181	-0.006	0.059	0.002	-0.125
	(0.007)	(0.014)	(0.083)	(0.119)	(0.097)	(0.109)	(0.264)	(0.275)
Age	0.001^a	0.002^a	-0.002	0.002	-0.006^c	-0.005	-0.025^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	0.102	-0.151	0.284^a	-0.022	1.054^a	-0.152
	(0.005)	(0.011)	(0.067)	(0.097)	(0.079)	(0.088)	(0.214)	(0.224)
Diversity	0.001	-0.011^a	-0.072^a	-0.058^b	0.043^c	-0.029	0.052	-0.047
	(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	-0.013	0.037^a	0.002	0.123^a	0.034
	(0.001)	(0.001)	(0.007)	(0.010)	(0.008)	(0.009)	(0.022)	(0.024)
Heterogeneity	-0.022^a	-0.021^a	0.001	-0.009	0.096^a	0.021	0.377^a	0.136^c
	(0.002)	(0.004)	(0.026)	(0.034)	(0.031)	(0.032)	(0.085)	(0.080)
USA connections	0.140^a	0.092^a	-0.268	0.062	1.054^a	0.248	3.089^a	0.693
	(0.017)	(0.025)	(0.181)	(0.220)	(0.220)	(0.205)	(0.598)	(0.514)
Non-USA connections	0.163^a	0.031	0.275^c	0.210	0.292^c	-0.294^c	1.314^a	-0.553
	(0.013)	(0.020)	(0.143)	(0.178)	(0.175)	(0.165)	(0.474)	(0.415)
Position 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	0.59	0.56	0.49	0.56	0.62
Within-time OLS 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

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

in urban economics, completed by a number of other ones potentially relevant for the academic research activity, matter. The answer is mostly negative as regards the overall department characteristics when individual fixed effects are controlled for and the field-specific ones play a significant role but have a small explanatory power. Alternatively, there is maybe not enough mobility among French academics to separately identify individual and department fixed effects. Without individual fixed effects, many more overall department characteristics have a significant impact and a higher explanatory power.

We document in Table 2 a slightly negative sorting (on unobserved characteristics) of individuals in departments generating larger local effects. We now refine this statement by computing the correlation between individual fixed effects and the departments' observed characteristics. Individual fixed effects appear to be pretty largely positively correlated with the departments' characteristics that increase output when fixed effects are not considered. For instance, the correlation of individual fixed effects for quality is 0.20 with research access, 0.32 and 0.33 for non-USA and USA connections, and 0.36 for the share of rank A positions. Therefore, the departments' observed characteristics that increase output do seem to attract most able academics (in terms of unobserved characteristics) without generating much local externalities. But simultaneously some departments' unobserved (good) characteristics would be negatively correlated with the individual unobserved (good) characteristics, or bad unobserved departments' characteristics would be positively correlated with good individual characteristics. Possible candidates can be the administrative or the teaching loads that would be higher, which would penalise research, in departments where most able academics (in terms of research) are, possibly because students are more demanding or because less resources are allocated to the administrative support for instance. Remember that these conclusions only hold if individual and department fixed effects are correctly identified, which is not fully granted here. Future contributions will have to further investigate these puzzles, of low explanatory power of departments' observed characteristics and possibly negative sorting on unobservables.

Table 5: 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	0.284	1.000	0.350	1.000	0.979	1.000
Observed characteristics	0.019	0.689	0.082	0.288	0.166	0.474	0.514	0.525
Composition effects	0.011	0.071	0.059	0.208	0.104	0.367	0.278	0.397
Gender	0.001	-0.032	0.001	0.009	0.012	0.084	0.039	0.089
Age	0.003	-0.090	0.002	-0.030	0.010	-0.019	0.058	-0.011
Rank A	0.002	-0.221	0.023	0.168	0.042	0.343	0.180	0.393
Positions	0.010	0.152	0.050	0.166	0.082	0.280	0.165	0.222
Research characteristics	0.021	0.588	0.057	0.201	0.108	0.375	0.354	0.450
Size	0.003	0.043	0.002	0.012	0.021	-0.008	0.027	-0.003
Research access	0.002	0.208	0.038	0.169	0.055	0.289	0.186	0.333
Diversity	0.001	0.138	0.038	0.107	0.011	0.016	0.014	-0.005
USA connections	0.008	0.465	0.019	-0.101	0.058	0.332	0.181	0.392
Non-USA connections	0.010	0.486	0.022	0.134	0.027	0.312	0.104	0.366
Heterogeneity	0.009	0.343	0.001	-0.002	0.031	-0.110	0.125	-0.091
Residuals	0.020	0.725	0.272	0.958	0.308	0.880	0.833	0.851

Let us move to the role of departments' characteristics. The only ones 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. Remember that these characteristics assess whether having many academics in the department involved in a given field helps them to publish more in that field. Such so-called localisation effects can reflect local economies of scale taking place within the field, as they do within industries for market activities. As reported in Tables 2 and 3, these field-specific variables, although they exert a significant effect, have an explanatory power much lower than the department-time fixed effects. By contrast, their explanatory power is stable when individual fixed effects are introduced contrary to the one of department fixed effects that much decreases. The variance of field-specific effects is low and the correlation with the dependent variable is around three times lower than that of the department fixed effects when individual fixed effects are controlled for, and ten times when they are not.

The marginal effects of field-specific characteristics is pretty large. For instance, the mere presence of an academic's field in the department increases the number of publications of other academics in that field or their average top quality by almost 40% (39.8% and 37.6% respectively). The effect is lower for the probability to publish and the average standard quality (13.1% and 9.2%).[27] Similarly, the elasticity of specialisation, which is usually found in the range 0.01-0.05 for productivity in market activities, 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 6.1% respectively.[28]

The impact of the size of the local economy on local productivity is one of the most studied questions in urban economics. We could have evaluated the role of the total size of the city where the university is located. However, we believe that local externalities can be even more localised as regards academic activities, for instance because they require more face-to-face contacts. Therefore, we use the size of the department, defined as its number of academics. It is in itself an interesting variable for policy since it is at least partly in the hands of the department head, the university, or the central government (in many European countries, for instance). We then test the relevance of our choice of spatial scale in two ways. First, we also include in the specification a research access variable that corresponds to the proximity to other departments, which allows us to separate very local size externalities from more extended ones. Second, we provide estimates in Bosquet and Combes (2015) at the employment area levels, a more aggregated spatial classification.

Department size, as almost all of overall department characteristics, has no significant impact on department fixed effects when individual fixed effects are controlled for. Without individual fixed effects, the positive elasticity obtained for quality, and for top quality even if not significant in that case due to a too large standard error, is of the same magnitude as the usual estimates obtained for market activities. Doubling the size of a department would increase average quality by 3-4%. The fact that the effect disappears when individual fixed effects are introduced may mean that larger departments are only places were academics with better unobserved characteristics gather, as discussed above.

Conclusions for the 'research access' variable seem to indicate a further role of local size at a larger spatial scale, through the proximity to nearby departments. When individual fixed effects

[27] $e^{0.335} - 1$, $e^{0.319} - 1$, $e^{0.123} - 1$, and $e^{0.088} - 1$, respectively.
[28] $2^{0.088} - 1$ and $2^{0.085} - 1$.

are not controlled for, the research access elasticity is positively significant for quantity, quality and top quality all together. It is pretty high for this last variable as multiplying research access by 5, which is around one standard deviation at the median, increases average top quality by 21.9%, and the impact is still 6.1% for quality and 4.1% for quantity. Again, this may only result from the sorting of more efficient academics according to unobserved characteristics in departments with better research access as the effects are not significant anymore when individual fixed effects are controlled for. The explanatory power is also much higher for market access than for size, with a still pretty low standard deviation of the effects but a rather large correlation with the dependent variable (0.33 for top quality from Table 5 for instance).

Closer to what is usually done in urban economics, we can assess the role of size at a larger spatial scale by aggregating the data at the city level. We use employment areas, which are 341 spatial units that fully cover France and were specifically built by Insee, the French national institute of statistics, to study the role of local labour markets. For many employment areas, there are either no universities or only one (for 38 universities): Considering department or employment area is therefore the same. Six employment areas host two departments, four employment areas host three departments, and three employment areas host four or more departments. Results provided in Bosquet and Combes (2015) show conclusions that are very similar at this scale by comparison with university departments. In particular, the elasticity for local size that should be the most affected by this change of spatial scale is only slightly lower and less significant generally speaking. Research access still matters, and its impact on top quality remain significant when individual fixed effects are controlled for, with a high value at 0.100. Overall it is difficult to assess whether local effects matter more at the department or the city level, possibly because these two levels are too close to each others. Unfortunately, moving to a larger spatial scale, the region, would too much reduce the number of observations.

Beyond size and market access, some other department variables have a significant explanatory power when individual fixed effects are not controlled for. This contrasts with the usual findings of the urban economics literature, where these two variables are found to be the main explanation of productivity differences across locations. We study the role of the co-authors' location through 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 significantly increased over time. If the links to co-authors were not controlled for, access to departments outside France could have been computed and could have had also a positive effect, at least for the first years in our sample when internet use was less widespread. Here, not only having co-authors abroad, either in the USA or elsewhere, increases both the individual quantity and quality of publications, as shown in Section 5, but a larger proportion of academics in the department with co-authors abroad has an explanatory power of department fixed effects among the largest. Elasticities too are large even if, 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, i.e. the probability to publish and the average top quality, while non-USA connections affect all dimensions of the publication activity (probability to publish, number of publications and their quality). In that case the effect is larger and larger as one moves across these dimensions (when individual fixed effects are not controlled for). In a world where distance matters much less than before, being connected with other academics elsewhere, and in particular in the USA, is important. This is fully consistent with the large role of networks in academia underlined by the literature.

Having heterogeneous academics in a department enhances the average quality of publications, and even more top quality, this effect remaining slightly significantly positive (at 10%) even when individual fixed effects are controlled for. We are not aware of any 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 pretty large too. By contrast, heterogeneity plays no role on the quantity published, and a small negative role on the probability to publish. Field diversity in the department has a rather small impact in general, most often negative but in any case pretty sensitive to the specification chosen. Industrial diversity for market activities is often found to have a non very robust impact too.

Finally, we consider as department characteristics the shares in the department of the various individual characteristics considered in the first step. All together they have an explanatory power similar to the explanatory power of the department overall research characteristics (lines 'Composition effect' and 'Research characteristics' respectively in Table 5). Still, this is the department's composition in terms of positions that matters the most. A larger share of higher/most productive positions tend in general to increase publication output, as it is detailed in Bosquet and Combes (2015), especially when individual fixed effects are not controlled for but a few effects remain significant even when they are. The share of women and the average age explain much less. Older academics, at a given position, exert a significant effect even when individual fixed effects are controlled for. It is positive but negligible for the probability to publish and slightly larger and negative for top quality. Increasing the deparment's average age by 3 years (which is close to one standard deviation at the median), for a given composition in terms of positions, decreases top quality by 7.2%.[29] The share of women in the department has no impact on publication output, neither positive nor negative.

8 Conclusion

Location matters for the publication performance of French economists and economic departments. A careful variance analysis of individual publication determinants shows that the explanatory power of department effects represents at least a fourth of the explanatory power of individual

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

effects. When explaining departments' performance, selection and local effects have similar explanatory power. As argued by Waldinger (2012), this corresponds to what many academics have in mind. However, it is in sharp contrast with previous findings from the literature, which concludes to the presence of small, if not totally absent, local effects. We attribute this difference in conclusions to the fact that we have access to an exhaustive data set of all academic economists in France, whom we can follow over time and across locations even when they do not publish, which presents the further advantage of not biasing the computation of departments' characteristics. Moreover, we also have access to more individual variables, some of them time-varying, which are usually unavailable. We also separately study the determinants of the probability to publish, the number of publications and their average quality, whereas only a quality-adjusted number of publications is generally considered in the literature. By contrast to the literature on agglomeration effects in market activities, we also exhibit two new puzzles that future contributions will have to further investigate. First, the explanatory power of local variables is low, close to zero when individual fixed effects are considered. This implies that unobserved local effects would be large. Second, to the extent that fixed effects are well identified, individuals sort negatively across good departments according to their unobserved characteristics.

Due to possible missing variables and reverse causality involved when estimating agglomeration effects, we do not claim to present a conclusive assessment of the role of department characteristics on individual performance. We argue in Section 2 that endogenous individual location choices is probably not much of a concern, at least in the French context, while historical natural experiments lack of external validity. Issues due to the endogeneity of department's characteristics could be more severe, and should be treated. The possibility of combining bibliometric and administrative sources as we do here will certainly be extended in the future (to longer periods, other fields and countries). This should allow researchers to find even more sources of exogenous variations to properly assess the role of endogenous and exogenous individual and department characteristics. Putting more structure to the underlying models of network formation and agglomeration and peer effects, which are only implicit here and therefore treated as a black box, should also certainly help to improve estimated specifications. Ultimately, important results should be obtained for the better design of higher education and research policies.

References

Abowd, John M., Kramarz, Francis, and Margolis, David N. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

Aghion, Philippe, Dewatripont, Mathias, Hoxby, Caroline M., Mas-Colell, Andreu, and Sapir, André. The governance and performance of research universities: Evidence from Europe and the U.S. *Economic Policy*, 25(61):7–59, 2010.

Auriol, Emmanuel, Friebel, Guido, and Lammers, Frauke. The firm as the locus of social comparisons: Internal labor markets versus up-or-out. Discussion Paper 8831, Centre for Economic Policy Research, 2012.

Azoulay, Pierre, Graff Zivin, Joshua S., and Wang, Jialan. Superstar extinction. *Quarterly Journal of Economics*, 125(2):549–589, 2010.

Besley, Tim and Ghatak, Maitreesh. Social incentives. *American Economic Review*, 98(2):206–211, 2008.

Borjas, George J. and Doran, Kirk B. The Collapse of the Soviet Union and the Productivity of American Mathematicians. *Quarterly Journal of Economics*, 127(3):1143–1203, 2012.

Bosquet, Clément and Combes, Pierre-Philippe. Supplement to 'Do large departments make academics more productive? Agglomeration and peer effects in research': Additional appendices. Available on https://sites.google.com/site/clementbosquet/, 2015.

Ciccone, Antonio and Peri, Giovanni. Identifying human capital externalities: Theory with an application to US cities. *Review of Economic Studies*, 73(2):381–412, 2006.

Combes, Pierre-Philippe and Gobillon, Laurent. The empirics of agglomeration economies. In Duranton, Gilles, Henderson, Vernon, and Strange, Will, editors, *Handbook of Urban and Regional Economics*, volume 5A. North-Holland, Amsterdam, 2015.

Combes, Pierre-Philippe and Linnemer, Laurent. Where are the economists who publish? Publication concentration and stock rankings in Europe. *Journal of the European Economic Association*, 1(6):1250–1308, 2003.

Combes, Pierre-Philippe and Linnemer, Laurent. Inferring missing citations. a quantitative multi-criteria ranking of all journals in economics. *GREQAM Working Paper*, (2010-25), 2010.

Combes, Pierre-Philippe, Duranton, Gilles, and Gobillon, Laurent. Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742, 2008a.

Combes, Pierre-Philippe, Linnemer, Laurent, and Visser, Michael. Publish or peer-rich (skills and networks in hiring economics professors). *Labour Economics*, 15(3):456–474, June 2008b.

Dubois, Pierre, Rochet, Jean-Charles, and Schlenker, Jean-Marc. Productivity and mobility in academic research: Evidence from mathematicians. *Scientometrics*, 98(3):1669–1701, 2014.

Ductor, Lorenzo. Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics*, 77(3):385–407, June 2015.

Ductor, Lorenzo, Fafchamps, Marcel, Goyal, Sanjeev, and van der Leij, Marco J. Social networks and research output. *Review of Economics and Statistics*, 96(5):936–948, 2014.

Gobillon, Laurent. The estimation of cluster effects in linear panel models. Processed, INED, 2004.

Hellerstein, Judith K., Neumark, David, and Troske, Kenneth R. Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labour Economics*, 17(3):409–446, 1999.

Hollis, Aidan. Co-authorship and the output of academic economists. *Labour Economics*, 8(4): 503–530, 2001.

Jackson, Matthew O. An overview of social networks and economic applications. In Jess Benhabib, Alberto Bisin and Jackson, Matthew O., editors, *Handbook of Social Economics*, volume 1, pages 511–585. North-Holland, Amsterdam, 2011.

Jacobs, Jane. *The Economy of Cities*. Random House, New York, 1969.

Kim, Han E., Morse, Adair, and Zingales, Luigi. Are elite universities losing their competitive edge? *Journal of Financial Economics*, 93(3):353–381, 2009.

Laband, David N. and Tollison, Robert D. Intellectual collaboration. *Journal of Political Economy*, 108(3):632–662, 2000.

Lotka, Alfred James. The frequency distribution of scientific productivity. *Journal of the Whashington Academy of Sciences*, 16(12):317–23, 1926.

Marshall, Alfred. *Principles of Economics*. Macmillan, London, 1890.

Medoff, Marshall H. Collaboration and the quality of economics research. *Labour Economics*, 10 (5):597–608, 2003.

Moulton, Brent R. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics*, 72(2):334–338, 1990.

Oyer, Paul. Initial labor market conditions and long-term outcomes for economists. *Journal of Economic Perspectives*, 20(3):143–160, 2006.

Rosenblat, Tanya S. and Mobius, Markus M. Getting closer or drifting apart? *Quarterly Journal of Economics*, 119(3):971–1009, 2004.

Rosenthal, Stuart S. and Strange, William C. Evidence on the nature and sources of agglomeration economies. In Henderson, Vernon and Thisse, Jacques-François, editors, *Handbook of Regional and Urban Economics*, volume 4, pages 2119–2171. North-Holland, Amsterdam, 2004.

Sacerdote, Bruce. Peer effects in education: How might they work, how big are they and how much do we know thus far? In Eric A. Hanushek, Stephen Machin and Woessmann, Ludger, editors, *Handbook of the Economics of Education*, volume 3, pages 249 – 277. Elsevier, Amsterdam, 2011.

Sauer, Raymond D. Estimates of the returns to quality and coauthorship in economic academia. *Journal of Political Economy*, 96(4):855–866, 1988.

Waldinger, Fabian. Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *Review of Economic Studies*, 79(2):838–861, 2012.

A Descriptive statistics and simple correlations

Table 6: 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
Upper position	0.35	0.45	0	0	1
Teaching	0.83	0.35	0	1	1
Publisher	0.33	0.44	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	9.0	31	41	56
Women	0.22	0.38	0	0	1
Upper position	0.49	0.46	0	0	1
Teaching	0.75	0.40	0	1	1
Quantity	0.52	0.46	0.17	0.33	1.06
Quality	13.2	14.3	4.0	7.9	29.4
Top quality	2.44	8.91	0.01	0.04	4.94
Authors number	1.9	0.7	1	2	3
Non-USA openness	0.1	0.3	0	0	1.0
USA openness	0.07	0.24	0	0	0
Individual diversity	2.6	1.6	1	2	5
Panel (c): Departments					
Publishers	0.34	0.20	0.11	0.30	0.62
Quantity	5.46	8.20	0.44	2.75	12.62
Quantity per academic	0.18	0.18	0.04	0.13	0.37
Quality	11.80	8.10	5.67	9.18	20.85
Top quality	1.93	4.44	0.02	0.27	5.06
Specialisation	0.28	0.20	0.12	0.20	0.50
Size	31.6	34.6	7.5	18.0	82.0
Women	0.24	0.12	0.10	0.24	0.39
Age	45.0	3.5	40.6	45.0	49.3
Upper position	0.34	0.19	0.13	0.31	0.64
Teaching	0.79	0.34	0	0.97	1
Department diversity	-0.52	0.46	-1.18	-0.42	-0.01
Research Access	11.4	17.6	0.8	2.8	37.9
Non-USA openness	0.04	0.07	0	0.01	0.14
USA openness	0.02	0.06	0	0	0.07
Heterogeneity	2.1	0.8	1.2	1.9	3.1
Stars	0.01	0.05	0	0	0.02

Variables are defined in Section 3. 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 38,577, 12,591, and 1,208 respectively. 165 individuals that have missing values for some variables are excluded from the sample. 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) before statistics are computed.

Table 7: Simple correlations at the department level

	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Quantity (1)	0.92	0.75	0.08	-0.07	-0.08	0.48	-0.36	0.59	0.32	0.55	0.48	-0.78	0.42	
Quality (2)	1	0.92	0.11	-0.08	-0.09	0.52	-0.43	0.52	0.37	0.63	0.59	-0.69	0.54	
Top quality (3)		1	0.21	-0.07	-0.06	0.48	-0.38	0.43	0.36	0.57	0.55	-0.51	0.49	
Size (4)			1	-0.02	0.19	0.06	0.15	0.16	-0.09	0.03	-0.04	0.16	-0.09	
Women (5)				1	-0.08	-0.28	0.16	0	0.23	-0.07	-0.07	-0.02	-0.15	
Age (6)					1	0.26	0.03	0.03	0.23	-0.15	-0.11	0.19	-0.14	
Upper position (7)						1	-0.58	0.15	0.40	0.38	0.43	-0.37	0.40	
Teaching (8)							1	-0.05	-0.36	-0.37	-0.46	0.41	-0.36	
Diversity (9)								1	0.19	0.25	0.18	-0.50	0.15	
Research Access (10)									1	0.30	0.34	-0.31	0.30	
Non-USA Openness (11)										1	0.61	-0.45	0.51	
USA Openness (12)											1	-0.39	0.72	
Heterogeneity (13)												1	-0.29	
Stars (14)													1	

Variables are defined in Section 3. Specialisation defined at the Jel code level is first averaged by department (weighted by the share of the Jel code in the department output) before statistics are computed.

B Variance analysis of individual probability to publish, publication quantity and top quality

Table 8: Variance analysis of the individual probability to publish

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Publishing	0.203	1.000		0.266	1.000	
Individual effects	0.033	0.170	0.055	0.073	0.259	-0.097
Individual fixed effect	-	-	-	0.089	0.192	-0.072
Observable individual effects	0.033	0.170	0.055	0.060	0.031	-0.012
Women	0.007	0.030	-0.012	-	-	-
Age	0.024	0.079	0.008	0.060	0.031	-0.012
Position	0.018	0.104	0.043	-	-	-
Rank A	0.020	0.085	0.048	-	-	-
Department effects	0.028	0.181	1.000	0.041	0.185	1.000
Department-time fixed effect	0.022	0.119	0.754	0.026	0.052	0.588
Department-field-time effects	0.019	0.133	0.626	0.034	0.188	0.780
Field presence	0.024	0.072	0.373	0.045	0.094	0.326
Specialisation	0.025	0.028	0.099	0.043	0.049	0.274
Field-time fixed effect	0.020	0.145	0.338	0.034	0.196	0.437
Residuals	0.196	0.966	0	0.249	0.935	0

Same notes as for Table 2.

Table 9: Variance analysis of the individual publication quantity

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Quantity	0.456	1.000		0.456	1.000	
Individual effects	0.214	0.471	0.006	0.324	0.653	-0.160
Individual fixed effect	-	-	-	0.261	0.503	-0.166
Observable individual effects	0.214	0.471	0.006	0.196	0.410	-0.043
Women	0.028	0.094	0.020	-	-	-
Age	0.066	0.027	0.035	0.036	0.012	0.063
Position	0.028	0.090	0.022	-	-	-
Rank A	0.061	0.085	0.075	-	-	-
Authors per publication	0.204	0.353	-0.078	0.198	0.353	-0.070
Individual diversity	0.031	0.046	0.008	0.042	0.046	0.014
Non-USA connection	0.074	0.110	0.064	0.038	0.110	0.032
USA connection	0.060	0.118	0.069	0.033	0.118	0.045
Department effects	0.123	0.274	1.000	0.134	0.168	1.000
Department-time fixed effect	0.104	0.234	0.840	0.120	0.115	0.891
Department-field-time effects	0.067	0.141	0.531	0.061	0.141	0.436
Field presence	0.048	0.033	0.185	0.047	0.033	0.151
Specialisation	0.074	0.106	0.359	0.066	0.106	0.293
Field-time fixed effect	0.071	0.158	0.011	0.078	0.102	-0.073
Residuals	0.376	0.825	0	0.312	0.685	0

Same notes as for Table 2.

Table 10: Variance analysis of the individual publication top quality

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Top quality	1.302	1.000		1.302	1.000	
Individual effects	0.527	0.527	0.237	1.027	0.808	-0.019
Individual fixed effect	-	-	-	1.050	0.621	-0.043
Observable individual effects	0.527	0.527	0.237	0.688	0.257	0.038
Women	0.063	0.069	0.016	-	-	-
Age	0.269	0.142	-0.014	0.649	0.144	0.028
Position	0.126	0.180	0.108	-	-	-
Rank A	0.151	0.141	0.150	-	-	-
Authors per publication	0.110	0.236	0.102	0.117	0.236	0.023
Individual diversity	0.035	0.179	0.125	0.008	0.179	0.015
Non-USA connection	0.223	0.334	0.138	0.067	0.334	0.018
USA connection	0.236	0.360	0.193	0.090	0.360	0.043
Department effects	0.467	0.489	1.000	0.283	0.207	1.000
Department-time fixed effect	0.459	0.468	0.982	0.278	0.176	0.979
Department-field-time effects	0.089	0.155	0.190	0.058	0.165	0.188
Field presence	0.050	0.067	0.101	0.044	0.067	0.011
Specialisation	0.101	0.104	0.118	0.064	0.104	0.164
Field-time fixed effect	0.264	0.337	0.168	0.158	0.281	0.038
Residuals	0.959	0.737	0	0.694	0.533	0

Same notes as for Table 2.

C Variance analysis of average probability to publish, publication quantity and top quality at the department level

Table 11: Variance analysis of average probability to publish at the department level

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Publishing	0.038	1.000		0.041	1.000	
Individual effects	0.023	0.729	0.200	0.040	0.627	-0.410
Individual fixed effect	-	-	-	0.046	0.505	-0.338
Observable individual effects	0.023	0.729	0.200	0.027	0.080	-0.040
Women	0.002	0.144	0.041	-	-	-
Age	0.009	0.266	0.096	0.027	0.080	-0.040
Position	0.015	0.629	0.132	-	-	-
Rank A	0.008	0.533	0.194	-	-	-
Department effects	0.027	0.817	1.000	0.035	0.453	1.000
Department-time fixed effect	0.027	0.831	0.992	0.036	0.423	0.977
Department-field-time effects	0.003	-0.202	-0.046	0.008	0.100	0.022
Field presence	0.014	0.376	0.433	0.028	0.344	0.048
Specialisation	0.014	-0.441	-0.460	0.024	-0.377	-0.050
Field-time fixed effect	0.001	0.028	0.014	0.001	0.001	0.009
Residuals	0.001	-0.021	0	0.001	0.060	0

Same notes as for Table 2. The effect of variables or group of variables is averaged by department-year before calculating the standard deviations and correlations.

Table 12: Variance analysis of average publication quantity at the department level

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Quantity	0.374	1.000		0.374	1.000	
Individual effects	0.222	0.646	0.093	0.397	0.635	-0.457
Individual fixed effect	-	-	-	0.339	0.404	-0.531
Observable individual effects	0.222	0.646	0.093	0.216	0.535	-0.008
Women	0.024	0.112	0.054	-	-	-
Age	0.063	-0.055	-0.020	0.032	-0.088	0.099
Position	0.033	0.232	0.106	-	-	-
Rank A	0.060	0.199	0.135	-	-	-
Authors per publication	0.230	0.486	0.011	0.224	0.486	-0.031
Individual diversity	0.028	0.112	0.055	0.038	0.112	0.030
Non-USA connection	0.061	0.079	0.035	0.031	0.079	0.017
USA connection	0.041	0.112	0.068	0.023	0.112	0.018
Department effects	0.286	0.805	1.000	0.327	0.380	1.000
Department-time fixed effect	0.285	0.805	0.977	0.326	0.379	0.983
Department-field-time effects	0.061	0.001	0.108	0.060	0.012	0.105
Field presence	0.107	0.095	0.145	0.103	0.095	0.115
Specialisation	0.087	-0.117	-0.101	0.078	-0.117	-0.072
Field-time fixed effect	0.058	0.012	-0.094	0.065	-0.041	-0.049
Residuals	0.001	-0.003	0	0.001	-0.013	0

Same notes as for Table 2. The effect of variables or group of variables is averaged by department-year before calculating the standard deviations and correlations.

Table 13: Variance analysis of average publication top quality at the department level

	Without individual fixed effects			With individual fixed effects		
	Std. dev.	Correlation	Sorting	Std. dev.	Correlation	Sorting
Explained: Top quality	1.283	1.000		1.283	1.000	
Individual effects	0.524	0.641	0.221	1.133	0.822	-0.163
Individual fixed effect	-	-	-	1.185	0.662	-0.189
Observable individual effects	0.524	0.641	0.221	0.635	0.232	0.062
Women	0.054	0.065	0.015	-	-	-
Age	0.254	0.094	-0.112	0.583	0.072	0.037
Position	0.176	0.280	0.067	-	-	-
Rank A	0.149	0.266	0.188	-	-	-
Authors per publication	0.124	0.346	0.126	0.132	0.346	0.068
Individual diversity	0.032	0.247	0.192	0.007	0.247	0.033
Non-USA connection	0.184	0.453	0.199	0.056	0.453	0.067
USA connection	0.163	0.525	0.280	0.062	0.525	0.074
Department effects	0.978	0.871	1.000	0.734	0.416	1.000
Department-time fixed effect	0.980	0.867	0.997	0.736	0.407	0.997
Department-field-time effects	0.070	0.032	0.008	0.057	0.096	0.001
Field presence	0.112	0.149	0.101	0.099	0.149	-0.027
Specialisation	0.119	-0.122	-0.090	0.076	-0.122	0.036
Field-time fixed effect	0.239	0.399	0.096	0.158	0.296	-0.099
Residuals	0.001	0.016	0	0.001	0.078	0

Same notes as for Table 2. The effect of variables or group of variables is averaged by department-year before calculating the standard deviations and correlations.