ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Goldschlag, Nathan; Bianchini, Stefano; Lane, Julia; SanMartin Sola, Joseba; Weinberg, Bruce A.

Working Paper

Research Funding and Regional Economies

IZA Discussion Papers, No. 10081

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Goldschlag, Nathan; Bianchini, Stefano; Lane, Julia; SanMartin Sola, Joseba; Weinberg, Bruce A. (2016) : Research Funding and Regional Economies, IZA Discussion Papers, No. 10081, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/145215

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 10081

Research Funding and Regional Economies

Nathan Goldschlag
Stefano Bianchini
Julia Lane
Joseba SanMartin Sola
Bruce Weinberg

July 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Research Funding and Regional Economies

Nathan Goldschlag
U.S. Census Bureau

Joseba SanMartin Sola
FECYT, Madrid

Stefano Bianchini
University of Strasbourg

Bruce Weinberg
Ohio State University and IZA

Julia Lane
New York University and IZA

Discussion Paper No. 10081
July 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 10081
July 2016

ABSTRACT

Research Funding and Regional Economies[1]

Public support of research typically relies on the notion that universities are engines of economic development, and that university research is a primary driver of high wage localized economic activity. Yet the evidence supporting that notion is based on aggregate descriptive data, rather than detailed links at the level of individual transactions. Here we use new micro-data from three countries - France, Spain and the United States - to examine one mechanism whereby such economic activity is generated, namely purchases from regional businesses. We show that grant funds are more likely to be expended at businesses physically closer to universities than at those farther away. In addition, if a vendor has been a supplier to a grant once, that vendor is subsequently more likely to be a vendor on the same or related grants. Firms behave in a way that is consistent with the notion that propinquity is good for business; if a firm supplies a research grant at a university in a given year it is more likely to open an establishment near that university in subsequent years than other firms.

JEL Classification: O30, R10

Keywords: science policy, innovation, regional economic development, UMETRICS

Corresponding author:

Julia Lane
Wagner School of Public Policy
New York University
The Puck Building
295 Lafayette Street
New York, NY 10012-9604
USA
E-mail: julia.lane@nyu.edu

[1] Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research was supported by the National Center for Science and Engineering Statistics. NSF SciSIP Awards 1064220 and 1262447; NSF Education and Human Resources DGE Awards 1348691, 1547507, 1348701, 1535399, 1535370; NSF NCSES award 1423706; NIHP01AG039347; and the Ewing Marion Kaufman and Alfred P. Sloan Foundations. Data were generously provided by the Committee on Institutional Cooperation and its member institutions. We thank Cameron Conrad, Ahmad Emad, Evgeny Klochikhin, for research support, Greg Carr, Marietta Harrison, David Mayo, Mark Sweet, Jeff Van Horn and Stephanie Willis for help with data issues, and Clara Eugenia Garcia Garcia, Cecilia Cabella Valdes, Barbara McFadden Allen, Jay Walsh, Roy Weiss, and Carol Whitacre for their continuing support. Chris Goetz provided extremely helpful comments and suggestions. The research agenda draws on work with many coauthors, but particularly Nikolas Zolas, Patrick Llerena, Ron Jarmin, and Jason Owen Smith.

1. Introduction

The link between university research and economic growth has been identified as a key driver of regional economies (Saxenian 1996; Nelson 2005; Zucker & Darby 1996). The geography innovation literature has identified important possible reasons for the link– the role of geographical closeness in the transmission of tacit knowledge(Gertler 2003) and the important of spatial proximity in regional knowledge production functions(Charlot et al. 2015; Caragliu & Nijkamp 2016). Concomitantly Owen Smith and Powell describe the importance of geographic location and network connections and the strategic benefits of location in dense alliance networks (Owen-Smith & Powell 2004).

However, much of the literature has been based on aggregate data with links inferred rather than directly measured (Hausman 2012; Kantor & Whalley 2014; Saha et al. 2015). This paper uses extremely granular data about purchases made in the process of conducting research to directly examine the links between universities and businesses in three countries (France, Spain and the United States). A particular strength of the data is that, in the US case, we can compare the characteristics of vendor firms and establishments with the characteristics of *all* U.S. firms and establishments. We are able, as a result and for the first time, to examine the persistence of project specific purchases from individual vendors as well as the behavior of firms which sell to research projects; both sets of analyses lend support to the notion that network connections are a source of regional economic impact.

There is strong evidence of regional vendor effects in all three countries. In the United States, an establishment within a university's state is 11 times more likely to be a vendor than an establishment outside its state. We also find vendors have characteristics associated with higher productivity (Syverson 2011): in the United States, establishments supplying publicly funded research tend to be older, larger, pay higher wages, and much more likely to be operated by an R&D performing firm than are non-vendor establishments.

We find support for the notion that tacit knowledge plays a role in determining research purchases - because we have data on all transactions and all individuals on all grants, we are able to determine that there is substantial persistence in purchases from a given vendor. In the United States data, if researchers on a grant purchase from a given vendor, they are 35-42% more likely to purchase from that same vendor in the following year. In the French data, although only 8% of vendors frequently (more than 20 times a year) do business with the research projects at the University of Strasbourg, they account for more than 50% of the value of research expenditures. In the Spanish data, the proportion of frequent vendors is the same – 8% - but they account for more than 70% of the value of the research expenditures.

We are also able to find support for the importance of propinquity by observing firm behavior by investigating where vendors to university research open new establishments. Strikingly, we show that firms that are vendors to universities are more likely to locate new establishments near the universities with which they do business and that the strength of the relationship increases the

larger the value of the transaction. Similar results hold among the subset of firms that open new establishments in a given year – those new establishments are more likely to be near the universities with which they do business, especially if the transactions are large.

2. Background

There is qualitative evidence that universities are engines of regional economic development – the existence of Silicon Valley has been traced to its propinquity to Stanford, Boston's growth has been attributed to its great research universities, and the Research Triangle to the research performed at Duke, University of North Carolina and North Carolina State University(Liu 2015). Part of the effect could be due to the effects of educational investments. Although it has been difficult to identify these relationships, work in Sweden has shown that universities have important effects on the productivity of workers (Andersson et al. 2004) and work in the United States suggests that a 10% increase in higher education spending increases local non-education sector labor income by about 0.5% (Kantor & Whalley 2014).

One mechanism by which this is thought to occur is through human contact: when people are physically close to each other, skills are easier to acquire and knowledge is easier to exchange (Duranton & Puga 2004; Gertler 2003). Another possibility is that the flows of students to jobs in regional businesses directly increase the human capital in the workforce, and result in more productive businesses (Stephan 2007; Zolas et al. 2015). As Hausman points out, many of the mechanisms, particularly hiring, patenting and spinoffs, have a local bias (Hausman 2012).

We provide detailed micro-evidence on one link which has not been fully examined – the role of universities in driving innovation through the purchase of scientific equipment. Such a link is plausible – customers and consumers are key drivers to firm innovation in other sectors (Von Hippel 2005; Freeman et al. 1968). And university researchers are important customers of R&D performing firms; in life sciences, for example, about 8% of R&D funds are directly spent on equipment while core research expenditures are likely to be much more (Stephan 2014). Certainly, universities are more likely to buy than to make their inputs (Cassiman & Veugelers 2006). We also know that innovation depends not just on technical knowledge, but also on the knowledge of customers, market segments and potential applications (Simard & West 2006). The literature suggests that physical closeness is an important source of this knowledge (Whittington et al. 2009), particularly for small and young firms which are beginning to establish customer ties.

3. Data

The source of US data is the enhanced STAR METRICS data, or UMETRICS data (Lane et al. 2015; Zolas, Goldschlag, R. S. Jarmin, et al. 2015). We use UMETRICS data from 8 major public research universities located across the Midwest (Ohio State University, Pennsylvania State University, Purdue University, and the Universities of Indiana, Iowa, Michigan, Minnesota and Wisconsin), provided as a result of a collaboration with the Big Ten Academic Alliance; our sample universities are large research intensive public institutions in the MidWest and as such

are not representative of the universe of all research universities although they do account for more than 10% of federal academic R&D expenditures.

The vendor data are unusually rich. For every purchase on every grant, the data give the date of each transaction, the amount and identifying information on the vendor (a unique vendor ID, the name, address and, when available, the DUNS Number). Taken as a whole our data cover over 749,000 transactions totaling over $1 billion on 15,923 grants from over 60,000 vendor establishments (an establishment is the physical place where business is conducted, and the unit of observation at which industry and geographic location are defined; firms can own one or more establishments). For cross country consistency, we use data from 2010 – 2012, although the results are consistent in previous years.

The transaction level data are extremely granular: every purchase made on every grant is included in the data, including very small transactions of just a few dollars. Most transactions are quite small, and reflect small scale purchases such as office supplies or trips of short duration. There are also a number of very large transactions that are from important scientific vendors. Some vendors are responsible for tens of thousands of transactions, but the simple mean of transactions per vendor is 9. Thus the distribution of the number of transactions and of transactions by value is right skewed.

The US data can be used for even more granular analysis because they provide information about each purchase from each grant at each of the universities in the sample. Those purchases can then be matched to data on all US non-agricultural employer establishments with their geographic location (longitude and latitude) from the US Census Bureau's Longitudinal Business Database (LBD) (Jarmin & Miranda 2002). This permits us to construct comparison groups of businesses and also directly estimate the likelihood that a given establishment is selected to be a vendor as a function of distance. It is worth noting that distance effects are likely to be nonlinear for two reasons. First, the universities are in the MidWest, so a natural boundary is the two coasts. Second, California, a particularly R&D intensive state, is between 1,500 and 2,000 miles away from the sample universities.

In order to construct the US comparison groups, we use a subsample of the LBD that includes all establishments associated with R&D performing firms. R&D performing firms include all firms that report non-zero expenditures in R&D in any given year between 1976 and 2012. The firm identifiers and R&D expenditures are collected from two separate surveys collected over two separate time periods. The R&D data from 1976 to 2007 are collected from the Standard Industrial Research and Development Survey (SIRD) and the R&D data from 2008 until 2012 are collected from the updated version of this survey called the Business Research and Development and Innovation Survey (BRDIS)[2]. Both surveys are jointly administered by the US

[2] https://bhs.econ.census.gov/bhs/brdis/about.html

Census Bureau and the National Science Foundation and represent a national sample of firms beginning in 1992. All firms that report conducting R&D in one year are retained to the next year, with additional firms sampled (based on survey weights). R&D performing firms make up a small share of all firms in the United States. Of the 5M+ firms in existence in the United States in 2012, fewer than 12,500 report conducting R&D. These firms are also known to significantly differ from the typical U.S. firms on a number of dimensions including being much larger and much more likely to engage in international trade and multinational activity (Davis et al. 2007).

The French data are derived from an analysis of the financial records data from 2011-2014 from the University of Strasbourg which is a major research university located in the Bas Rhein region about 300 miles from Paris. We examine data from about 500 grants, expending almost 15 million euros on over 19,000 transactions from 1291 vendors. An advantage of the French data is that each transaction is classified into one of 58 object codes (a *rubrique*) that reflect the nature of the purchase, and which we aggregate into 4 broad categories – research materials, meetings, operating costs and other expenses.

The Spanish data are derived from the same source: human resource and financial records in 2012 from the Complutense University of Madrid, the Instituto de Ciencias Matemáticas in Madrid and the Centre for Genomic Regulation in Barcelona. Data are available for 292 grants, expending almost 8 million euros on almost 11,000 transactions from 1,006 vendors. The Spanish data can be matched to industry information and compared to US data.

4. Basic Facts

Our analysis of the data generates two stylized facts: (i) university vendors have characteristics that are associated with higher levels of productivity than do non-vendors and (ii) vendors are more likely to be geographically closer to the university than non-vendors.

Vendors are disproportionately in high technology industries. For example, in the United States 11% of the establishments are in Scientific Research and Development Services (NAICS code 5417), while only 1.2% of US establishments are from that industry. Other major industries include Professional and Commercial Equipment and Supplies Merchant Wholesalers (9% vs. .9% nationally), Book, Periodical and Music Stores (5.4% vs. .04% nationally) and Pharmaceutical and Medicine manufacturing (4.5% vs. .37% nationally)

Table 1: Comparison of Vendor Industries with All US Establishments

Industry Code	NAICS Description	Proportion of Establishments	
		All Vendors	All U.S.
5417	Scientific Research and Development Services	11.05%	1.21%
4234	Professional and Commercial Equipment and Supplies Merchant Wholesalers	9.48%	0.93%
4512	Book, Periodical, and Music Stores	5.38%	0.03%
3254	Pharmaceutical and Medicine Manufacturing	4.54%	0.37%
6113	Colleges, Universities, and Professional Schools	5.54%	1.72%
8139	Business, Professional, Labor, Political, and Similar Organizations	4.07%	0.42%
3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	3.96%	0.53%
7223	Special Food Services	3.05%	0.22%
3273	Cement and Concrete Product Manufacturing	1.85%	0.12%

Source: UMETRICS and LBD, author's calculations.

Interestingly, the vendors also have characteristics that are associated with greater firm level productivity: the establishments are larger (39 times larger) than the typical U.S. establishment, higher wage, older, much more likely to be owned by an R&D performing firm and much more likely to be owned by a firm that patents. They also have substantially higher employment growth rates (9.3% versus 6.5%).

Table 2: Comparison of Vendor Characteristics with All US Establishments

		Vendors	US	Ratio
Employment	Mean	651.70	16.80	38.79
	Median	34.50	3.40	10.15
Average Payroll per worker	Mean	67,000	38,000	1.76
	Median	53,000	25,000	2.12
Firm Age	Mean	24.80	17.00	1.46
	Median	33.00	14.30	2.31
Employment Growth		9.33%	6.49%	1.44
Percent within 50 Miles of University		4.37%	0.61%	7.10
Percent within 50-100 Miles of University		4.74%	1.09%	4.34
Percent within 100-250 Miles of University		2.47%	9.39%	0.26
Percent within State of University		25.51%	2.42%	10.54
Percent Performing R&D		29.15%	4.55%	6.41
Percent Patenting (of R&D Performing)		28.19%	10.61%	2.66
Percent Firm Age Under 3 Yrs		5.83%	16.92%	0.34
Percent Firm Age Under 5 Yrs		9.79%	24.03%	0.41

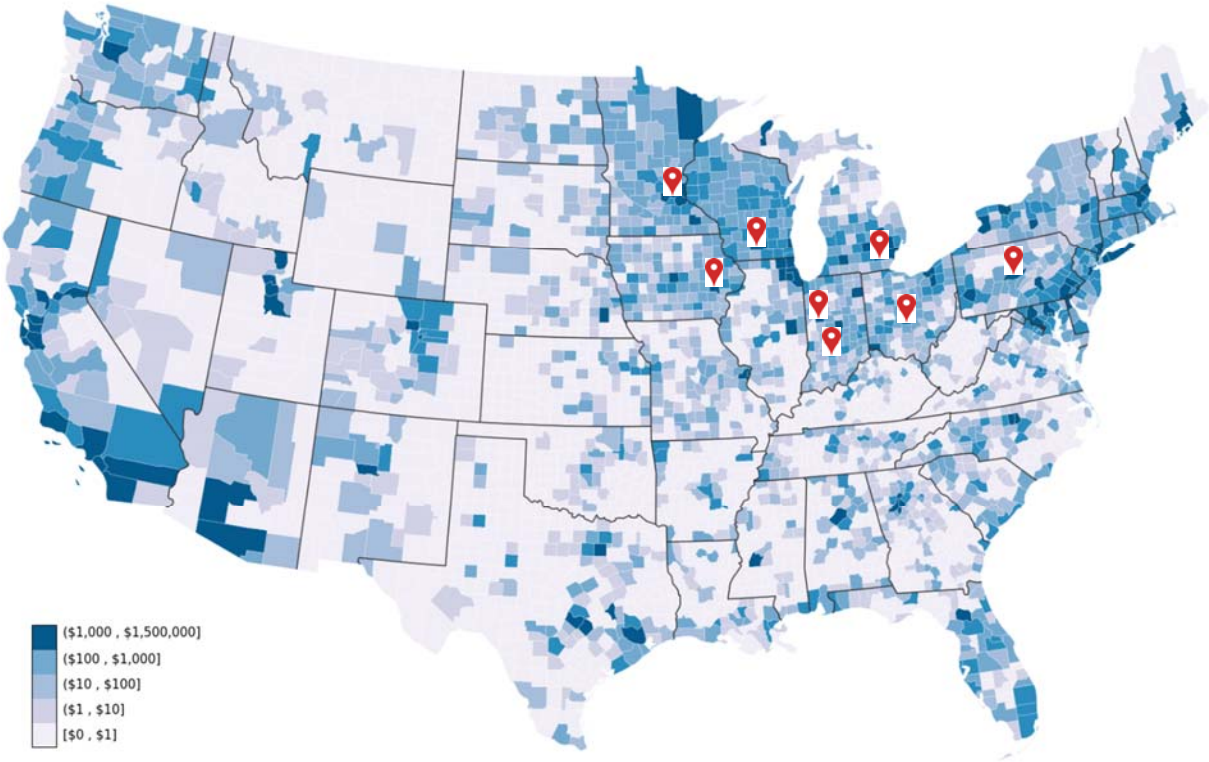
Source: UMETRICS and LBD, author's calculations.

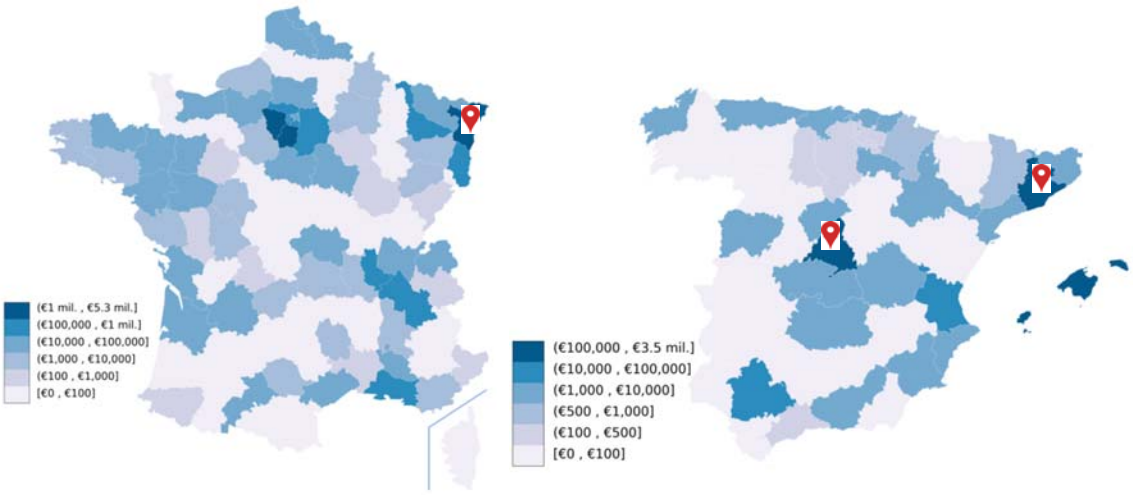
Note: Statistics calculated pooling 2010-2012 for all universities in the sample. Percent within distance bins and within state give the average across all vendors - whether the vendor is within the distance bin or state of the university. In the case of establishments, we calculate whether each establishment is within the distance bin or state of each university and then average across universities. Employment growth is calculated using the standard Davis, Haltiwanger and Schuh (1996) method weighting by the average of employment in *t* and *t-1*. Medians calculated as the mean of the 45th and 55th percentiles.

It was not possible to construct industry linkages for France, but in Spain, the pattern is similar. About one third of the domestic vendors are in high technology: more than 12% of the firms belong to sectors classified as medium or medium-high tech by the Eurostat taxonomy, and they account for about one third of the total amount purchased. There is detailed information on the nature of research purchases from French data: two thirds of the research purchases, both by value and volume were for scientific materials.

There is some initial evidence of the regional clustering of purchases. Figure 1 illustrates the link between the universities in the U.S. sample (primarily Mid-Western universities) and vendor purchases; purchases appear clustered in the Midwest and on the East and West coast. The same pattern of clustering purchases near the university is observable in both France and Spain.

Figure 1: The geographic distribution of vendor purchases in the US (A), France (B), and Spain (C). The figure shows the amount of vendor purchases on grants from the universities in each country (indicated by the pins) to each part of the countries.





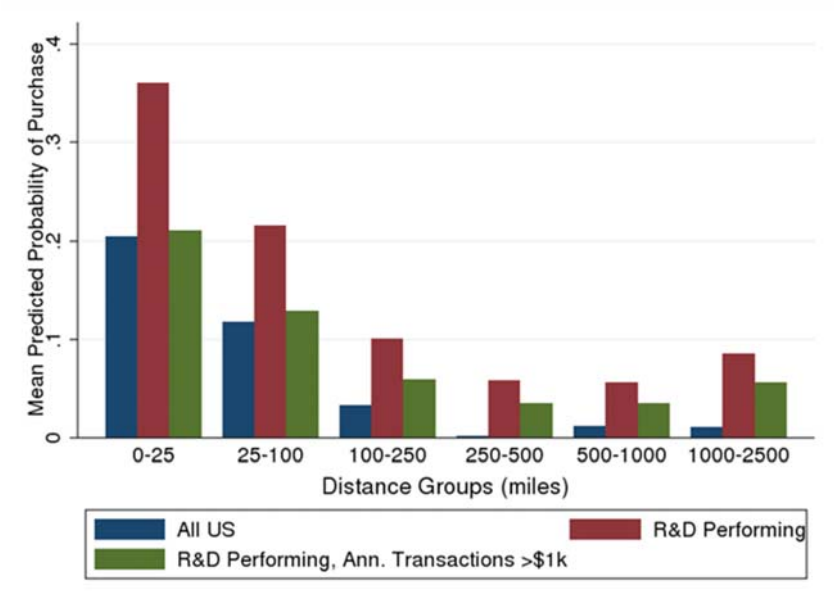
Source: Spanish data by province from Complutense University of Madrid, the Instituto de Ciencias Matemáticas in Madrid and the Centre for Genomic Regulation in Barcelona; French data by district from University of Strasbourg; US data by county from UMETRICS; author's calculations.

5. The link between research expenditures and regional activity

The clustering of research expenditures near research organizations evidenced in Figure 1 is not sufficient to conclude that there are disproportionate links; some of the universities are located in dense urban areas where there is a great deal of economic activity. Fortunately, the United States data enables a much deeper level of analysis, since it is possible to match data on research vendors to the register of all non-agricultural employer firms and associated establishments in the US.

An analysis of that match makes it clear that establishments physically closer to each university are disproportionately more likely to be vendors to the university's researchers than are other US establishments. Establishments within 50 miles of the university, as shown in Table 2, are more than 7 times likely to be research suppliers than establishments in the US as a whole; establishments 50-100 miles away are 4 times more likely, and 100-250 miles away are considerably less likely. Put another way, Figure 2 shows results from a simple regression of the probability a given US establishment being a vendor to a given university as a function of distance suggests that an establishment within 25 miles of each university has about a .20% chance of being a vendor; that drops to about .10% if the establishment is between 25 and 100 miles away. The probabilities are much higher for establishments associated with R&D performing firms – those that are closest have a .35% chance of doing business with the university, and the probabilities, while remaining higher than for all establishments, still decline with distance. However, in all cases there is an uptick in probability for vendors 1000-2500 miles away, which is consistent with the distance from our MidWestern universities and the East and West Coasts.

Figure 2: Vendor purchases related to distance from the university. The figure shows the probability of purchase from establishments as a function of distance.



Source: UMETRICS and LBD, author's calculations.
Note: Mean predicted probability calculated as the estimated probability from a regression of whether an establishment is a vendor in for a given university as a function of distance, distance to the 2nd, 3rd, 4th, 5th, and 6th power, 2010 population within the establishment's zip code and population squared, year, and university fixed effects with robust standard errors.

The data are so rich that it is possible to control for multiple confounding factors – and the results continue to be robust. In particular, since universities often have restrictive purchasing agreements, we control for university specific purchases to particular vendors across universities by including university fixed effects and we also use year fixed effects to control for year specific changes in purchasing agreements. We also control for density of people per zip code, to capture city agglomeration effects (Chatterji et al. 2013). In addition, since the data are transaction level data, there were many very small purchases and one-time vendors; we can transcend the micro-level activity to only examine larger scale transactions, such as those exceeding $5,000 in a year.

Table 3 reports the results of a sixth-order polynomial regression examining the effects of distance including all controls and with three different outcome measures. The first has a binary dependent variable: whether or not any purchase was made from a vendor. The second is also binary: whether the expenditures exceeded $5,000. Finally, since much of the policy interest is in the effect of research funding on R&D intensive firms, the third uses the binary outcome measure for the subset of R&D intensive firms.

The estimated effects of distance are strongly statistically significant and imply that the probability of a transaction is minimized at a substantial distance from the university, ranging from 763 miles (for establishments owned by R&D performing firms) to 2,003 miles (for annual

transactions exceeding $5000). (The model for all transactions also has a local minimum at 558 miles while that for transactions exceeding $5000 has a local minimum at 563 miles). The results are robust to the choice of polynomial.

Table 3: Physical Distance Regressions (Linear Probability Model of whether establishment is a vendor)

	All Establishments: All transactions	All Establishments: Transactions at least $5000	Establishments owned by R&D Performing Firms: All transactions
x	-0.0603*** (0.000970)	-0.0150*** (0.000493)	-0.0637*** (0.00647)
x^2	0.0111*** (0.000189)	0.00276*** (9.64e-05)	0.0102*** (0.00130)
x^3	-0.000927*** (1.64e-05)	-0.000231*** (8.32e-06)	-0.000829*** (0.000115)
x^4	3.80e-05*** (6.84e-06)	9.49e-06*** (3.47e-07)	3.67e-05*** (4.95e-06)
x^5	-7.39e-07*** (1.34e-08)	-1.85e-07*** (6.79e-09)	-7.88e-07*** (9.97e-08)
x^6	5.42e-09*** (9.92e-11)	1.36e-09*** (7.52e-12)	6.31e-09*** (7.51e-10)
Constant	0.102*** (0.00176)	0.0259*** (0.000898)	0.166*** (0.0144)
Observations	179,373,000	179,355,000	7,986,000

Source: UMETRICS and LBD, author's calculations.
Note: X is Distance between establishment and university (100m); Robust standard errors in parentheses; year and university fixed effects included. Observations are university-establishment pairs. Observation counts rounded.

6. The source of regional ties

The literature has hypothesized that network connections are an important source of regional ties. The data make it possible to examine that possibility, because we have detailed information on all grants and all people working on all grants. As such, it is possible to go beyond a simple binary analysis of whether a business is a vendor to a research project, and examine the nature of the relationship. We examine both: whether the vendor does repeat business with senior personnel on a given grant (persistence) and the role of physical propinquity.

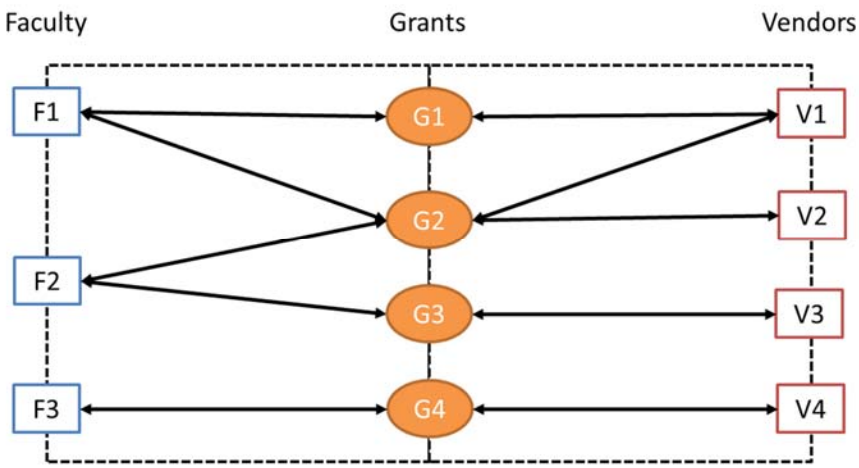
Constructing network measures

The construction of the analysis draws on the fact that the data consist of each purchase from each vendor for each grant for each year, so it is possible to track persistence in purchases between vendors and grants. In addition, because the data include information on all individuals supported on each grant for each transaction period (typically monthly), it is possible to link faculty connections across grants. Measuring these connections requires constructing network units of observation based on grant-vendor relationships. The reason for this is that it is possible to determine from the data that a purchase from a vendor was made by a grant, but, since there can be several faculty on a grant, it is not clear which of the faculty members on the grant made

the purchase decision. We thus construct network measures where the nodes are grants, and the edges are defined where the grants share ANY faculty in common.

Formally, we define an annual grant-to-grant network, illustrated in Figure 3 as follows. Each node in the network represents a grant in a given year. Edges are drawn between two nodes if at least one faculty member is paid on both grants in a given year. From this grant-to-grant network we can then calculate the shortest path length between each pair of grants. This measure represents a measure of how distant those grants are with the social dimension of the payroll transaction data. We hypothesize that selling to a grant in one year increases the probability that a vendor sells to that grant in the next year. Figure 3 also illustrates this hypothesis about social distance, namely that V3 is more likely to sell to G2 in the next year than to G1 (because V3 is 1 step away from G2, via F2 but 2 steps away from G1). Additionally, V3 is more likely to sell to G2 in the next year than is V4 because V4 and G3 are not connected at all. Figure 3 provides a visual description of the relationships knowable from the purchase matches. In this example, it is clear that there has been a purchase from V2 using funds from G2 but it is not known whether the purchase decision was made by F2 or F1.

Figure 3: An illustrative mapping of faculty, grants and vendors



In particular, let $Y_t(G_i, V_j)$ denote whether or not a purchase is made between grant *i* to vendor *j* at time *t*. Further, let $D_t(G_i, G_k)$ denote the shortest path length between grants *i* and *k* at time t. For example, in the figure above $D_t(G_1, G_2) = D_t(G_2, G_1) = 1$ and $D_t(G_1, G_3) = 2$. Let $V_t^j(G_i; d)$ denote whether any grants distance *d* from grant *i* purchased from vendor *j*. In this case, $V_t^j(G_i; d) = max_{k|D_t(G_i,G_k)=d} Y_t(G_k, V_j)$.

We estimate:

$$Y_t(G_i, V_j) = \beta_0 + \beta_{-1} Y_{t-1}(G_i, V_j) + \beta_1 V_{t-1}^j(G_i; 1) + FE(U, t) + \epsilon_{ijt}$$

In this formulation:
β_{-1} denotes the effect of lagged activity between grant *i* and vendor *j*.
β_1 denotes the effect of activity among grants that are distance 1 from grant *i* with vendor *j* last period.

Examining Persistence

In 2012 each of the 8 universities has, on average, about 1,200 grants and more than 6,500 vendors, which yields more than 76 million possible university-grant-vendor combinations. The chance that a given grant purchases from a particular vendor is slight: 0.14% in 2012. Yet, as is shown in Table 4, there is substantial persistence in purchases. In particular, if funds from a given grant were used to make a purchase from a given vendor in 2011, the chance of a purchase being made from the same grant in 2012 is increased by approximately 40%; if in 2010 by approximately 15%. The social dimension created by the relationships between faculty and grants also impacts the probability of purchasing from a given vendor. If a purchase was made in 2011 the grant-vendor distance in 2011 is necessarily zero, otherwise the distance is positive and dependent on the network structure.[3] Given that there was no purchase in 2011, if the shortest network distance was 1 in 2011 then the probability of purchase in 2012 increases by 3.2%; if distance of 1 in 2010 approximately 1.2%.

The data permit an even deeper examination, since we can subset to only examine vendors associated with R&D performing firms. There are some 12 million possible university-grant-vendor combinations for such vendors; the persistence in purchases is similar. Our analysis also suggests that the impact of the network dimension is slightly higher for the R&D sample, with a 4.2% increase in the probability of purchase for one-step-away vendors in 2011. Both the lagged relationship and overlap in teams are strong predictors when included together. In other words, persistence effect and shared faculty are important independent factors related to vendor purchases.

[3] Note that using shortest distance does not capture the possibility that, for a given grant, a vendor may interact with a number of grants at varying distances to that grant. Future research could explore this dimension in more detail.

Table 4: The probability that a purchase is made from a vendor in 2012

	All Establishments			Establishments in R&D performing firms		
Purchase made in 2011	39.43***	34.55***	34.54***	41.08***	35.74***	35.78***
	(0.160)	(0.161)	(0.161)	(0.340)	(0.345)	(0.345)
Purchase made in 2010		14.88***	14.59***		15.54***	15.12***
		(0.155)	(0.155)		(0.333)	(0.332)
Grant-vendor shortest path of 1 in 2011			3.157***			4.241***
			(0.0408)			(0.114)
Grant-vendor shortest path of 1 in 2010			1.209***			1.296***
			(0.0431)			(0.112)
Constant	0.0886***	0.0810***	0.0838***	0.122***	0.111***	0.0944***
	(0.000341)	(0.000388)	(0.000398)	(0.000994)	(0.000985)	(0.000925)
Observations	76,071,000	76,071,000	76,071,000	12,338,000	12,338,000	12,338,000
R-squared	0.140	0.153	0.156	0.151	0.165	0.169

Source: UMETRICS and LBD, author's calculations.
Note: The regressions include university-vendor fixed effects to control for university contracts with specific vendors. Observations are university-vendor(establishment)-grant triplets. Observation counts rounded.

The role of propinquity

If propinquity is indeed important to creating persistent business relationships, firms should be aware of it and should be more likely to locate near universities. The US data permits an examination of this possibility, because the LBD has data on all establishments owned by all firms in the United States. It is thus possible to examine the opening of all new establishments by all firms, including those which have business relationships with university.

An analysis of the 2012 LBD data supports the notion that firms pay attention to previous business relationships in making location decisions. Opening a new establishment within 50 miles of the 8 universities in the sample is a rare event – only 0.24% of all multi-unit firms in the US did so. However, 5.85% of firms who did business with one of the universities in 2011 opened an establishment within 50 miles of that university in 2012; 4.42% of R&D performing firms did so. Among the 860 R&D performing firms that did business with the 8 universities in 2011, 38 firms opened one or more establishments within 50 miles of the universities. These 38 firms averaged transaction revenues of over $462,000 in 2011 and $368,000 in 2012 with the university near their new establishment. By contrast, 822 R&D vendor firms did not open an additional establishment; the mean value of the transactions of those 822 firms was $28,990 in 2011 and $28,960 in 2012. Thus, being a vendor and the size of the relationship is strongly related to where firms open new establishments. Table 5 presents a regression analysis of the probability that a firm with multiple establishments in 2011 opens a new establishment within 50 miles of a university in 2012 based on whether the firm was a vendor in 2011 and the amount of the 2011 transaction.

Table 5. Probability of a Multi Establishment Firm Opening an Establishment Near a University

	All	R&D Performing	Those Opening New Establishments	Those R&D Performing Opening New Establishments
Vendor in 2011	2.065***	-0.321	6.643***	-1.742
	(0.480)	(0.704)	(1.559)	(2.337)
2011 Amount ($1M)	3.358**	5.710**	5.575**	6.445*
	(1.548)	(2.585)	(2.789)	(3.393)
Constant	0.104***	0.243***	0.832***	1.632***
	(0.00880)	(0.0566)	(0.0703)	(0.371)
Observations	1,183,000	59,000	147,000	9,000
R-squared	0.252	0.370	0.246	0.364

Source: UMETRICS and LBD, author's calculations.
Note: The regressions include university fixed effects and firm fixed effects to control for firm growth patterns and regional differences. Observations are university-firm pairs. Observation counts rounded.

An analysis of the results presented in the first column suggests those that are vendors to a specific university are over 2% more likely to open an establishment near that university with the probability being 3.6% higher for each $1 million increase in the amount of 2011 transactions.[4] The same analysis subset for R&D performing firms, suggests being a vendor in its own right is not important, but that the size of the relationship in dollar terms is (at 5.7% per $1M). This result is consistent with the notion that R&D performing firms are considerably larger than other firms, and often have more specialized business processes.

The results are even more compelling when, in columns 3 and 4, we subset the sample to consider only the firms that opened an establishment in 2012. Here we are interested in whether that new establishment is more likely to be near a university with which it is a vendor. Although the coefficients are not directly comparable to those in columns 1 and 2, the estimates in columns 3 and 4 show that the relationship between being a vendor and $1M in transactions are similar whereas for R&D performing firms, only transaction amounts are related to the probability of opening an establishment. These results suggest that firms that participate in university via the supply of the inputs to the research enterprise are more likely than other firms to collocate near the university, potentially strengthening the ties between university research and local economic activity.

[4] These results are robust to including controls for the firm's size and the geographic distribution of a firm's establishments in prior years instead of firm fixed effects.

7. Summary

Although all the results described here are descriptive rather than causal, they document striking facts that directly relate research purchases to regional economic activity at a detailed micro level. We show that grant funds are more likely to be expended at businesses physically closer to universities. In addition, if a vendor has been a supplier to a grant once, that vendor is subsequently more likely to be a vendor on the same or related grants. Firms behave in a way that is consistent with the notion that propinquity is good for business; if a firm supplies a research grant at a university in a given year it is more likely to open an establishment near that university in subsequent years than other firms. The results are consistent with the notion that research funding has a stimulative regional economic impact, particularly in the high tech and R&D performing sectors of the economy.

Of course, much more can be done. A key point for future research will be to examine the persistence in vendor faculty relationships over a faculty member's career rather than focusing on just individual grants. The results do not show causal effects, so it will be necessary to examine the impact of exogenous shocks, such as faculty moving from one university to another, new large scale grants to university based research centers, or changes in state economic development policies.

We hope this paper provides the stimulus for a new body of research on the economic and knowledge-transmission role of university research purchases, particularly since more universities are joining the UMETRICS program. These data will be made available for scholarly research purposes (at least for the U.S.) by the newly-formed Institute for Research on Innovation and Science (iris.isr.umich.edu) and the Census Bureau's Federal Statistical Research Data Centers ((http://www.census.gov/fsrdc).

8. References

Andersson, R., Quigley, J.M. & Wilhelmson, M., 2004. University decentralization as regional policy: the Swedish experiment. *Journal of Economic Geography*, 4(4), pp.371–388.

Caragliu, A. & Nijkamp, P., 2016. Space and knowledge spillovers in European regions: the impact of different forms of proximity on spatial knowledge diffusion. *Journal of Economic Geography* , 16 (3), pp.749–774. Available at: http://joeg.oxfordjournals.org/content/16/3/749.abstract.

Cassiman, B. & Veugelers, R., 2006. In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management science*, 52(1), pp.68–82.

Charlot, S., Crescenzi, R. & Musolesi, A., 2015. Econometric modelling of the regional knowledge production function in Europe. *Journal of Economic Geography* , 15 (6), pp.1227–1259. Available at: http://joeg.oxfordjournals.org/content/15/6/1227.abstract.

Chatterji, A., Glaeser, E.L. & Kerr, W.R., 2013. *Clusters of entrepreneurship and innovation*, National Bureau of Economic Research.

Davis, S.J. et al., 2007. *Measuring the dynamics of young and small businesses: Integrating the employer and nonemployer universes*, National Bureau of Economic Research.

Duranton, G. & Puga, D., 2004. Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics*, 4, pp.2063–2117.

Fellegi, I.P. & Sunter, A.B., 1969. A Theory for Record Linkage. *Journal of the American Statistical Association*, 64, pp.1183–1210.

Freeman, C. et al., 1968. Chemical process plant: Innovation and the world market. *National Institute Economic Review*, pp.29–57.

Gertler, M.S., 2003. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of economic geography*, 3(1), pp.75–99.

Greenia, N., Husbands Fealing, K. & Lane, J., 2008. *Studying Innovation In Businesses: New Research Possibilities*, Available at: http://www.irs.gov/pub/irs-soi/08rpinnovbusgreenia.pdf.

Hausman, N., 2012. *University Innovation, Local Economic Growth, and Entrepreneurship*, Available at: http://ideas.repec.org/p/cen/wpaper/12-10.html.

Von Hippel, E., 2005. *Democratizing innovation*, MIT press.

Jarmin, R.S. & Miranda, J., 2002. *The Longitudinal Business Database*, Center for Economic Studies, U.S. Census Bureau. Available at: http://ideas.repec.org/p/cen/wpaper/02-17.html.

Kantor, S. & Whalley, A., 2014. Knowledge spillovers from research universities: evidence from endowment value shocks. *Review of Economics and Statistics*, 96(1), pp.171–188.

Lane, J.I. et al., 2015. New linked data on research investments: Scientific workforce, productivity, and public value. *Research Policy*. Available at: http://www.sciencedirect.com/science/article/pii/S0048733315000025 [Accessed February 23, 2015].

Liu, S., 2015. Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, 87, pp.25–41.

Nelson, R.R., 2005. *Technology, institutions, and economic growth*, Harvard University Press.

Owen-Smith, J. & Powell, W., 2004. Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Communit. *Organization Science*.

Saha, S., Staudt, J. & Weinberg, B., 2015. Estimating the Local Productivity Spillovers from Science. , 039347(July), pp.1–39.

Saxenian, A., 1996. *Regional advantage*, Harvard University Press.

Simard, C. & West, J., 2006. Knowledge networks and the geographic locus of innovation. *Open innovation: researching a new paradigm*, pp.220–240.

Stephan, P., 2014. The Endless Frontier: Reaping What Bush Sowed? In *The Changing Frontier: Rethinking Science and Innovation Policy*. University of Chicago Press. Available at: http://www.nber.org/chapters/c13034.

Stephan, P., 2007. Wrapping It Up in a Person: The Mobility Patterns of New PhDs. In J. Lerner & S. Stern, eds. *Innovation Policy and the Economy, Volume 7*. MIT Press.

Syverson, C., 2011. What Determines Productivity? *Journal of Economic Literature*, 49(2), pp.326–365.

Whittington, K.B., Owen-Smith, J. & Powell, W.W., 2009. Networks, propinquity, and innovation in knowledge-intensive industries. *Administrative Science Quarterly*, 54(1), pp.90–122.

Zolas, N., Goldschlag, N., Jarmin, R.S., et al., 2015. Wrapping it up in a person: Examining employment and earnings outcomes for PhD recipients. *Science*, 360(6266), pp.1367–1371.

Zolas, N., Goldschlag, N., Jarmin, R., et al., 2015. Wrapping it up in a person: Tracing flows from funded research into the economy. *Science*.

Zucker, L.G. & Darby, M.R., 1996. Star scientists and institutional transformation: patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences of the United States of America*, 93(November), pp.12709–12716.

Appendix: Data Construction

The vendor data used in this report are derived from the financial transactions data of eight research organizations in the United States, three in Spain and one in France

1. United States

The U.S. data are derived from the UMETRICS[5] program - an effort parallel to the federally supported STAR METRICS program. The UMETRICS vendor data used in this report are drawn from the financial transactions associated with federal research grants awarded to researchers at the Universities of Indiana, Iowa, Michigan, Minnesota, Ohio State, Purdue, Penn State, and Wisconsin. The data do not, of course, cover the universe of all research grants. However, these 8 CIC institutions account for more than 10% of federal university R&D expenditures.

Both federal and nonfederal funding is covered in the data. There are a total of 749,149 transactions, with payments made to 62, 086 vendors

The vendor data are matched to Census Business Register (BR) and Longitudinal Business Database (LBD) data. The BR consists of the universe of U.S. non-agricultural firms and their associated establishments and is the ultimate source of all other Census economic data. [6] The LBD is the longitudinally linked employer version of the BR, providing a database that allows us to track firm performance, births and deaths over time. It combines administrative records and survey-based data for all nonfarm employer business units in the United States and hence provides information about the dynamics of firm growth. Key data elements include industry classification, geographic data, employment measures, payroll, and firm age.[7]

---

[5] UMETRICS is a university-led initiative to build a scientific framework that will inform research management, enable evidence-based decision making, and support credible advocacy. Universities participating in the UMETRICS initiative submit quarterly micro-data on university payroll, vendor, subaward/subcontract, and overhead expenditures from federal and non-federal grants and contracts. The data submitted are transactional in format and aggregated for all analyses. In the rest of this report, subaward/subcontract data are referred to as subaward.

[6] The key source data elements in the Business Register are (i) the SS-4, by which a new business tells the IRS whether it is beginning as a sole proprietorship, partnership, corporation, or personal service corporation; the State or foreign country in which it is incorporated; and whether it is applying because it is a new entity, has hired employees, has purchased a going business, or has changed type of organization (specifying the type) and (ii) the 1120S K-1 series which provides information on corporate shareholders.(Greenia et al. 2008)

[7] Non-employer businesses, which constitute the majority of businesses in the United State (although only 4% of sales and receipts), have no paid employees. Our ability to track business from the non-employer to the employer stage allows us to identify startups that may not succeed as well as the transition path.

As a benchmark, we use a subsample of the LBD that includes all establishments associated with R&D performing firms. R&D performing firms include all firms that report non-zero expenditures in R&D in any given year between 1976 and 2012. The firm identifiers and R&D expenditures are collected from two separate surveys collected over two separate time periods. The R&D data from 1976 to 2007 are collected from the Standard Industrial Research and Development Survey (SIRD) and the R&D data from 2008 until 2012 are collected from the updated version of this survey called the Business Research and Development and Innovation Survey (BRDIS)[8]. Both surveys are jointly administered by the US Census Bureau and the National Science Foundation and represent a national sample of firms beginning in 1992. All firms that report conducting R&D in one year are retained to the next year, with additional firms sampled (based on survey weights). R&D performing firms make up a small share of all firms in the United States. Of the 5M+ firms in existence in the United States in 2012, fewer than 12,500 report conducting R&D. These firms are also known to significantly differ from the typical U.S. firms on a number of dimensions including being much larger and much more likely to engage in international trade and multinational activity (Davis et al. 2007).

To match the university vendor and the Business Register (BR) data, we use fuzzy name and location matching and rule based block matching techniques(Fellegi & Sunter 1969). Once BR matches are identified, we create a longitudinal university-vendor-year panel using the LBD. The LBD embodies a number of cleaning algorithms, including the retiming of births and deaths around Economic Census years (Jarmin & Miranda 2002).

The university vendor data is annualized by calculating the sum and the mean of the vendor payment amounts for each vendor for each transaction year and is then combined with the

These data are merged with the LDB to add latitude and longitude fields for distance calculations between vendors and their associated universities. The LBD also provides firm age and cleaned versions of BR attributes

[8] https://bhs.econ.census.gov/bhs/brdis/about.html

2. France

The French data are derived from an analysis of the human resource and financial records data from University of Strasbourg which is a major research university located in the Bas Rhein region about 300 miles from Paris.

The records contain information about non-permanent workforce and goods & services directly supported by expenditures associated with grants from publicly funded sources made to 65 laboratories and active during at least one year in the period 2011-2014.[9] Data are available from about 500 grants, expending almost 40 million euro on over 65,000 transactions. Most of those are salary and personnel payments; in this paper we focus on over 19,000 transactions that generated almost €15 million in purchases from a total of 1,291 vendors.

An advantage of the French data is that each transaction is classified with an object code (a rubrique) which can be aggregated into the nature of the expenditure – research materials, meetings, operating costs and other expenses. More than two thirds of total expenditures were spent for research material (e.g. scientific equipment), while about 10% is the average share for the other groups. It is worth noting that the share for research material is always at least half of the total amount, although there is disciplinary variation. Some applied research laboratories (e.g. genetic engineering) require very expensive equipment and, as a consequence, the share of research material on the total purchases can easily reach values above 90%; social sciences and humanities are much less.

Domestic companies represent over 75% of the total population of vendors, and they account for about 90% of the total purchases, both by value and volume. A similar pattern holds if one looks inside the different types of the expenditure and regardless of the funding source. By way of example, about 70% of vendors of research material are located in France, and receive approximately 90% of the total purchases. As for the non-domestic suppliers, we conclude that only a negligible proportion is located outside the EU, and they account for a very small share of the amount expended.

[9] In this document we distinguish between three main funding sources: (i) the national funding agency ANR (Agence Nationale de la Recherche), (ii) the European Research Council (ERC) , and (iii) a category that we label 'other' which comprehends disparate public sources (e.g. regional funding agencies).

Although most of the vendors are not recurrent (only about 8% has frequent interactions with the university labs, defined as more than 20 economic transactions during the full time considered) the recurrent vendors account for a high share of the total amount expended (more than 50%), while vendors who supply a single time account for a negligible proportion. The cumulative amount expended to each vendor is on average about 11,500€, but the distribution is highly skewed since the median value is 1,200€. The picture is also sensitive to the nature of the expenditure. For instance, vendors of scientific material shows an average cumulative amount of purchases of more than 16,500€, with median which exceeds the value of 2,000€. There are finally few "leading suppliers" with a cumulative amount of more than 500,000€.

The street address associated to each domestic vendor allowed us to add latitude and longitude fields for distance calculations between vendors and the associated university.[10] The geographic dimension of expenses shows that purchases appear primarily clustered in the Alsace region (where the University of Strasbourg is located) and in the area which surrounds Paris. To a lesser extent, it is possible to observe a third cluster in the southeast area around Lyon, although a more deep investigation reveals that only 5% of vendors are located in that zone.

The clustering of vendors is evident also according the different types of expenses. However, we observe some interesting discrepancies across categories. By way of example, about 55% of vendors of scientific material are located around Paris, whereas such percentage drops to about 20% for vendors in the category mobility & meeting. A possible interpretation is that the type of goods and services which are required to undertake specific research projects can force labs to sourcing where specialized vendors are located, and this can be particularly true for customized scientific equipment. Although vendors for research material are primarily clustered around Paris, we notice that the number of transactions to vendors of research material located near the university exceeds the one for vendors in the Paris cluster. And this is also the case for all the other categories but 'other costs'.

About two thirds of the research purchases, both by value and volume, were for scientific materials. About 30% of these were made within 50 miles of the University of Strasbourg, compared with about half of purchases for meetings and travel, and almost 90% of other services (which include construction) were within 50 miles.

---

[10] Geographical coordinates were obtained through an in-house algorithm that extracted latitude and longitude from Google Maps APIs for the full vector of domestic vendors. The distance between companies and associated university was calculated with a standard Haversine formula.

3. Spain

The Spanish data are derived from the same source: financial records covering the year 2010 from the Complutense University of Madrid (UCM), and 2012 from the Instituto de Ciencias Matemáticas in Madrid (ICMAT) and the Centre for Genomic Regulation in Barcelona (CRG). Data are available for about 300 grants, expending almost 8 million euros on almost 11,000 transactions from 1,006 vendors (the unique number of vendors may be less, since the three universities may use the same vendor for some purchases).

The Spanish data can be matched to industry information and compared to US data. About one third of the domestic vendors belong to Wholesale Trade industries (2-digit NACE code 46): among these firms more than 60% are in Wholesale of Pharmaceutical Goods (NACE code 46.46), Wholesale of Computers, Computer Peripheral Equipment and Software (46.51), Wholesale of Other Machinery and Equipment (46.69), Wholesale of Chemical Products (46.75). For example, more than 12% of the domestic firms belong to sectors classified as high or medium-high tech by the Eurostat taxonomy, and they account for about one third of the total amount purchased. Just over 3% of vendors are in Scientific Research and Development Services (2-digit NACE code 72), but these firms account for about 90% of the total amount purchased from high-tech companies and more than 25% of the total.

Domestic companies represent over 50% of the total population of vendors, and they account for about 85% of the total purchases, both by value and volume.[11] A similar pattern holds if one disaggregates the data by research organization. In line with the conclusions we draw from the French scenario, we observe that only a negligible proportion of suppliers is located outside the EU, and they account for a very small share of the amount expended.

Most of the vendors are not recurrent, indeed only about 8% have frequent interactions with the university labs - more than 20 economic transactions during the full time considered. Nevertheless these vendors account for a high share of the total amount expended (almost 70%), and this is also the case for high-tech suppliers (albeit with a lower share, at 26%).

The street address associated to each domestic vendor allowed us to add latitude and longitude fields for distance calculations between vendors and the associated university.[12] At the aggregate

[11] Note that the data do not provide the geographical location for approximately 25% of vendors.

[12] Geographical coordinates were obtained through an in-house algorithm that extracted latitude and longitude from Google Maps APIs for the full vector of domestic vendors. The distance between companies and associated university was calculated with a standard Haversine formula.

level the geographic dimension of expenses reveals the presence of two big clusters of vendors around Madrid and Barcelona, and a similar conclusion holds looking at the disaggregated figures. However, the number of vendors located near by the associated university is definitely larger: about 70% of the firms are within 50 miles, and they account for about 60% of the total purchases. Moreover, a focus on high-tech vendors shows that more than 90% of the amount of purchases is within 50 miles.