The Spanish research network in the current budget crisis

Alice Duhaut, Université Libre de Bruxelles, ECARES *

Sept. 2016

Abstract

In this paper, I look at the impact of inter-university research partnerships on the production of research outputs. Using an original data set of scientific publications, I analyze the network of research in Spain based on the network of Spanish co-authors. I show how the growth in research productivity of Spanish institutions before the crisis was linked to the increase in universities' budgets and in inter-university collaborations. The results show that the size of the university is the key factor to understand universities' productivity, and that, everything else constant, a one percent increase in the number of students in associated universities leads to an average decrease of the research output of the university. The network multiplier is significant and positive, indicating that collaboration has a positive effect. Finally, in the context of the current crisis, I am able to identify the universities that are the least productive, taking into account their own characteristics and the indirect effects of the collaborations. This analysis has clear policy implications, as the least productive universities could be targeted to minimize the impact of further budgets cuts.

Keywords: Higher education and research institutions, Economics of science, Regional policy.

JEL classification: I23, R5, O30

1 Introduction

Since 2008, Spain has experienced a major economic crisis that has led to a severe reduction in funding for the public sector. Subsequent budget cuts were large and affected previous efforts to promote research in Spain. Since the beginning of the present century, government funding of higher education and research has increased, and a large decentralization process had allowed the creation of regional science agencies. This decentralization process changed the way funds were allocated to research, going from a centralized system to a regionalized one. Overall, the impressive growth of the university sector helped Spain build its research capacities and helped secure its position in an increasingly competitive research world. Several features of the Spanish case are particularly interesting and motivate the present paper. First, due to budget increases and the

* *email*: aduhaut@ulb.ac.be. I thank Stephane Bonhomme and Micael Castanheira for their comments and suggestions. All remaining mistakes are mine.

decentralization process, up to 60 % of research projects are conducted in collaboration with other universities. Second, the Spanish higher education system is characterized by a high level of decentralization and heterogeneity in the size and quality of the research institutions, thus leading to asymmetric partnerships. In this context, this paper studies the interplay between university resources, inter-university partnerships, and the universities' research production. Based on the the two features of the Spanish higher education and research system, I argue that research shouldn't be seen as having been conducted by each single university, but by teams that are part of a network of universities. Any change within an institution will have direct and indirect impacts on the other universities connected through the network of partners. This allows us to explicitly take into account the direct and indirect network effects on the single-university approach of research production. The direct effect implies that existing partnerships foster direct changes in the investment of resources and ideas by the partners. The indirect effect indicates that partnerships induce changes in investment decision of ideas and resources in partners and third parties, either to catch up with fellow researchers or to benefit previous investments in capital or experience.

This papers aims to understand how the network structure affects the total re- search output of the country and, in the current crisis context, if further budget cuts can be managed such as to minimize their overall impact. I contribute to the literature by adding to the growing body of empirical work on networks. I provide evidence of the correlation between research output, the universities' budgets and the intensity of inter-university links. Finally, I quantify the strength of the complementarity between universities' research production, and the effect of changes in available university resources on the whole system.

The model used is a model of networked production that explains universities' research output by their own characteristics as well as strategic decision of their partners. It focuses on the short-term behavior of the partners, knowing the existing partnerships. It characterizes the equilibrium research output as being linked to the relative position of each university in the inter-university partnership network as well as its size and budget. This equilibrium can be characterized by the use of a common network metric. Based on this equilibrium outcome, and provided that the structure of the network of partnerships remains unchanged, it indicates which universities are instrumental to maintaining the level of the research output.

The identification strategy is based on the stability of the inter-university partnerships over time and estimates the game played on the pre-existing network of partnerships. Like the existing literature, I start by assuming exogeneity of the network and the quantity of research output produced. I then relax this assumption by allowing the inter-university network to be predetermined before each time period. This assumes that the changes to the network, while having lasting effects on the output, are not chosen simultaneously as the research level. This leaves the strategic decisions behind the changes unexplained. This allows us to understand the decision to allocate resources to internally based projects or, instead, to jointly based projects, knowing the resource constraints. This uses past structure of the network to understand how the shock impacts the research process in the short run, without addressing the question of the changes in alliances. In this setup, the budget shocks are not used for the estimation of the network parameters themselves, but give the context of the network game.

I show that the current budget crisis induced a slowdown in Spanish research, stopping the growth of the research output. I show evidence that joint research projects are a major element explaining the Spanish science sector's production. Using the model, I show that the elasticity of the research output to the number of staff working at the university is positive. The network multiplier is positive and indicates the positive association between one university and its partners' research output. Finally, depending on the value placed on the diversification of the research partners, I provide a list of the least influential universities, those that are either middle range universities in regions dominated by bigger institutions, or weakly connected universities from peripheral regions.

The outline of the paper is the following. Section 2 presents the literature on higher education institutions, particularly in Spain, and the network models upon which this paper is based. Section 3 describes the institutional context for Spanish higher education and research institutions and its impact on university resources and describes the original data on publications and university budgets. This section also provide an empirical assessment of the size of inter-university collaborations and their patterns. It shows that taking into account the network provides information on research production. Section 3.4 provides a description on how the Spanish crisis unfolded, and its impact on the higher education institution. Section 4 presents a simple model of networked research production between universities to explain its implications in terms of welfare and public policy. Section 5 presents the estimates of the model. Section 7 presents results indicating the robustness of the model to various changes in specification. Section 9 analyzes the implications of those results for public policy. Finally, section 10 concludes.

2 Related Literature

While previous literature focused either on networks of co-authors or university productivity as a single element, this paper is an attempt to use network models to link the two fields of study, and broaden the scope of empirical applications of network models, and provide elements for integrated policy design.

One of the first papers to model universities as organizations producing multiple outputs and maximizing an objective function subject to constraints is the work of Becker (1979). The literature has focused on a unique feature of the university, mainly that the consumers are also inputs of its production function (Rothschild and White (1995)). My paper relies on this literature for the understanding of the university production function, and build on previous attempts to understand the cost of teaching and research, as well as the allocation of resources inside universities. With respect to the productivity of the universities themselves, efforts have been made to estimate the cost functions for teaching and research using data envelopment techniques. Those studies view higher education institutions as prestige maximizers, and focus on the optimal size of the teams and departments. Among others, Dundar and Lewis (1995) and Laband and Lentz (2003) who focus on American universities, Longlong et al. (2009) for Chinese universities, and Duch-Brown et al. (2010) and Abramo et al. (2012) for Europe, find economies of scale. Adams and Griliches (1998) and Worthington and Higgs (2011) arrive at a contrasting result. Duch-Brown et al. (2010) analyses the case of Spanish higher education institutions and focus on both the research output and technology transfers. They estimate the costs of producing multiple outputs using a flexible fixed cost quadratic function. While my paper focuses on research produced by universities, I highlight the production of research as the main output and consider the resources spent on educating students as part of the costs. Finally, on a study of inter-university collaboration involving American universities between 1981 and 1999, Adams et al. (2005) find that the most prestigious and richest universities are part of larger teams and that scientific influence increases with team size and participation. This paper is also related to a large body of literature focusing on competition among universities. De Fraja and Iossa (2002) analyzes competition for the best students and research activities in a model with prestige-oriented institutions. Closer to the present paper, Del Rey (2001) models faculty teaching activities as a source of income used to fund research and not something that directly increases their utility. De Fraja and Valbonesi (2012) notes that most universities' administrators are academics themselves, and that career progression is mostly based on research or on mechanical age rules, not on teaching activities, giving hint of the importance of research in the universities' objective function. The approach I follow here is an attempt to combine these two branches of literature by evaluating the impact of joint projects on a university's production function. While the problem of alliances is twofold in the long run - choosing whom to partner with and how much effort to exert - I show in the empirical part of the paper that the network of alliances among universities is stable across time. Based on this fact, I follow most of the empirical literature on networks and focus on the short-term impacts of the crisis, taking the network as given and considering the game of research effort played knowing the long term partnership formed.

The model of networked production I use here is a variation of the classical quasi linear games developed by Ballester et al. (2006). The equilibrium and welfare properties of those games have been described in Ballester et al. (2006) for games with strategic complements, in Bramoullé et al. (2014) for games with strategic substitutes, and in Ballester and Calvó-Armengol (2010)for games that, once transformed, display strategic complementarities. While this model might not be suited to analyzing formation of partnerships and long-term strategies, it provides short-term comparative statistics to evaluate the impact of the network structure on the productivity of universities and to understand how they decide whether or not to spend time on collective projects. I will thus distinguish two things: the network structure itself, given by the share of the project that a university allocates to each partner and that is stable over time, and the investment in collective work.

3 Data description

In this section, I present three broad facts about Spanish universities: First, although all universities engage in research, they are heterogeneous in their research performance as measured by the number of papers published by researchers affiliated with the institution. Second, they are heterogeneous in size and resources. Third, research partnerships, as defined by co-authorships, are essential to the research process.

3.1 The Spanish system of higher education and research

As described in Mora (2001), the system of higher education in Spain consists almost exclusively of universities: 50 of them are public universities, 29 are private universities that account for a small segment of the total number of students.

Until 1996, the Federal State was in charge of funding and regulating the universities. Since 1996, the Spanish regions, the Autonomous Communities, have the jurisdiction over the financing and the organization of higher education, while the Federal State specifies the general rules. This model was designed to insure a stable level of funding regardless of the current economic situation, to ensure the provision of a standard of quality education and to create a better research environment. It created a system in which the salaries of the civil servants working in universities are determined by federal law, but are paid by the regional governments. Tenured university researchers, researchers at the CSCI, and those centers dependent on the Spanish Ministry of Economy and Competitiveness are civil servants; their contracts are designed by the central governments. Untenured personnel cited above have short-term contracts that can vary in terms of workloads and duration(Mora (2001)).

Concerning the funds themselves, around 80 % of the operational budget comes from public funding, while the rest of it comes from private resources, via project-based funding. While the rules for allocating funds to universities varies from one Autonomous Community to the other, the general guidelines come from the *"Ley Organica 6/2001 de Universidades"*. In parallel, in several Autonomous Communities, a system of performance-based funding ensures the production of quality science and competition between research centers. Those rules are centered around performance in both research and teaching (up to 20 % of the budget non based on the number of students) (González, Luis Otero, David Rodeiro Pazos (2007)).

With regard to research, three other types of institutions exist: the *Spanish National Research Council* (*CSIC*)at the federal level, the several regional research councils, and finally, nonprofit private institutions. These nonprofit institutions are mostly private partnerships, and their research funds remain marginal and are mainly targeting applied research. In 2008, 95% of the basic research taking place within the public universities and was financed mainly by the Autonomous Communities.

This description pictures a strong support of public sector for Spanish research centers. The budget crisis of 2008 intervened in this context[1].

[1]The highly decentralized funding system, as well has the non performance-based rule for funding

3.2 Heterogeneous Universities

I collected budget data for all 48 public universities in Spain, using liquidated expenses budgets. These data are unified and made available by the Spanish National Institute of Statistics starting in 2009. I collected the same data for previous years from university budgets and various publications of the universities' union platform, the Conferencia de Rectores de las Universidades Espanolas (García and Chica (2010), Armenteros (2006), Armenteros (2008)). I also collected data on the number of staff and the number of students at various study levels from CRUE and the Spanish Ministry of Education. Spanish universities are heterogeneous on three dimensions: the number of students, the number of teaching and research staff, and budget.

Universities' summary statistics, 2004-2014

	Mean	Std. Dev.	Min	Max
Budget	184,617,909.93	118,415,251.60	32,651,000	624,638,952
Number of teaching and research staff	2,004.42	1,221.08	383	6,448
Number of students	27,474.42	24,089.41	4,255.	167,610

Figure 1: Summary statistics of the Spanish public universities' budget, number of research and teaching staff, and number of students, 2004-2014

Aggregated expenditures have increased from 7 billion in 2004 to 9.8 billion in 2008 and increased slightly between 2008 and 2010 to reach 10.01 billion After 2010, the aggregated expenditure budget decreased to 8.07 billion in 2014, roughly the level of 2007. However, aggregated expenditures mask strong variations in universities' expenditure level:

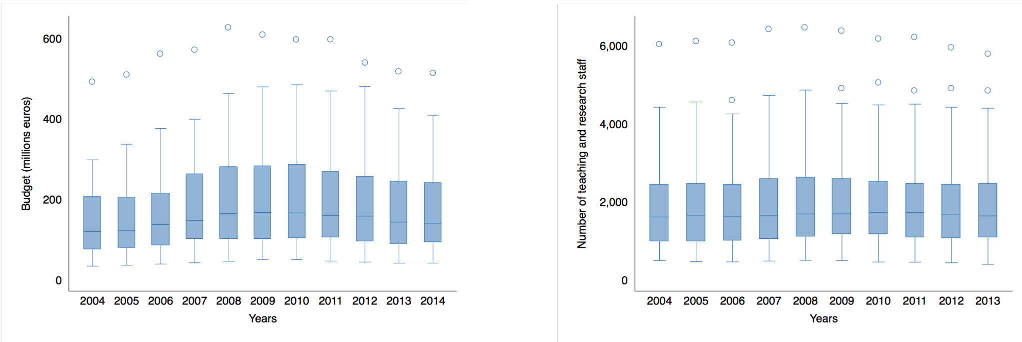
Figure 2 shows the evolution of Spanish public universities' liquidated expenditures between 2004 and 2014. The median level of expenditure, the blue line in the middle of the boxplot, increased from 120 million in 2004 to 165 million in 2009 and decreased to 120 million in 2014. The gap between the richest (the upper bar for each year on the box plot) and the median university also increased between 2004 and 2008..

The decrease in budgets after 2009 is of comparable relative size across universities, as assessed by a test comparing the mean growth for each university[2]. The highest budget

explain the randomness of the cuts happening since 2008. Figure 16 in the appendix shows the moment of the first shock, defined as a negative growth of the budget of any size, for Spanish public universities. There is no clear geographical pattern for the first occurrence of shock, nor is there a strong correlation between the moment of the first shock and the research performance of the university. Universities in Madrid, Andalucía, Cataluña, Castilla y León and Aragón were the first to see their budget reduced in 2008-2009. However, several university budgets in the Autonomous Community of Madrid, Andalucía and Castilla y León started to decrease only after 2010. The maps in Figure 15, in the appendix show the average growth rate of the universities budgets by location, before and after the crisis. Figure16 in the appendix provides the starting points of the crisis for each university, as defined by a decrease in the budget for two consecutive years. See the exploratory regressions at the end of for a discussion of the slow-down in research due to budget cuts

[2] Wilks' lambda, Pillai's trace, Lawley-Hotelling trace and Roy's largest root all reject the null

levels are those of the University Complutense of Madrid and the offsite university, U.N.E.D., with a peak to 620 million in 2008. The lowest budget levels are of the Universidad de La Rioja and range from 32 million in 2004 to 40 million in 2014, with a peak of 49 million in 2009.



mat.png

Figure 2: Evolution of the Spanish public universities' liquidated expenditures budgets, number of students at bachelors and masters level, and staff size, 2004-2014

The maximum number of teaching and research staff is 6,448, which were present at the University Complutense of Madrid in 2008, while the lowest belongs the Universidad de La Rioja in 2013. Comparing the mean of the number of teaching and research staff to the standard deviation between universities gives a coefficient of variation of about 60 %, which indicates heterogeneity among the university staff sizes. Figure 2 shows the evolution across time of the number of teaching staff. The median number of staff did not evolve much in the period considered. Both visual inspection of Figure 2 and multiple group variance comparison show that the variation among universities in staff size remains constant across time.

Finally, the Spanish public universities are also heterogenous in terms of their number of students. As seen on Figure 2, the maximum number of students is very high and is due to the large distance learning institution, the U.N.E.D.. Comparing the mean and the standard deviation gives a coefficient of variation of about 80 %. This is mainly due to the regional appeal of most universities, as a huge share of the students at the bachelors and masters levels attend their regional university. In 2012, 92 % of the students studying in Cataluña were from Cataluña, and 73 % of those studying in Madrid were from the Community of Madrid (Perez and Serrano Martinez (2012)). These percentages reflect the general trend in Spain. As the budget of universities is determined partly on the basis on the number students, it means that the Autonomous Communities are paying for students who are from the region itself. If we look at the evolution of the number

hypothesis of universities having similar mean budget growth rates across years before 2008, but I cannot reject the null hypothesis of having similar mean growth rate after 2008 or after 2009. These results remain the same even when I consider heterogeneity within budget levels for each university of the students, the median number has not changed over time, nor the variance in size. As is clear from Figure 2, the U.N.E.D, is a clear outlier.

As the budgets, staff numbers, and student numbers vary from university to university, it is useful to assess whether the university is research-oriented by looking at the student-to-staff ratio. This ratio goes from 28 for the offsite university to less than 10 for the most well-staffed universities. While the ratio decreased up to 2008, it increased slightly afterwards. Three universities have seen their ratio of students to staff increase dramatically over time, most notably the University Rey Juan Carlos of Madrid. Some universities, such as the Rovira I Virgili University, have seen their ratio decrease after the crisis. While the variance between observations remains constant over time, the median clearly increases prior to 2008, and then decreases slightly.

3.3 Heterogenous universities' research performance

Looking at the period of interest, Figure 3 shows that when the main fields of publication are aggregated (conference proceedings, articles, and books), medicine is the most popular subject, regardless of the year considered. The number of publications in every field increased until 2012. After 2012, the amount of output in most fields except medicine and engineering stagnated.

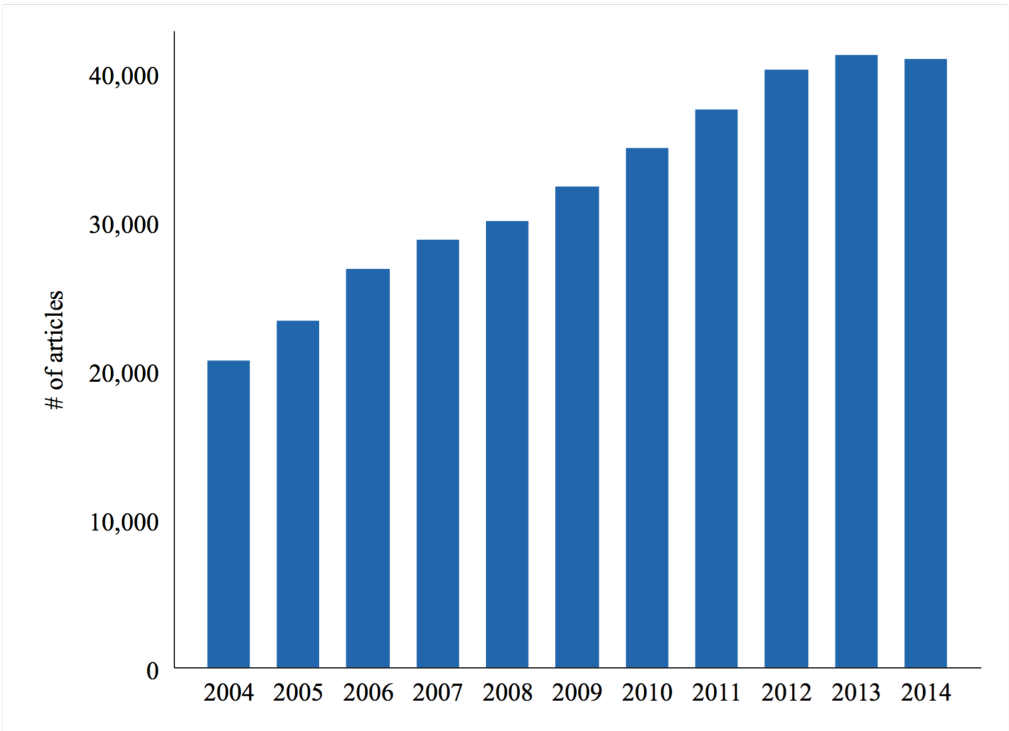


Figure 3: Stock of publications, by subfield, 2004-2014

Using the SCOPUS database, I collected all the publications appearing in scientific

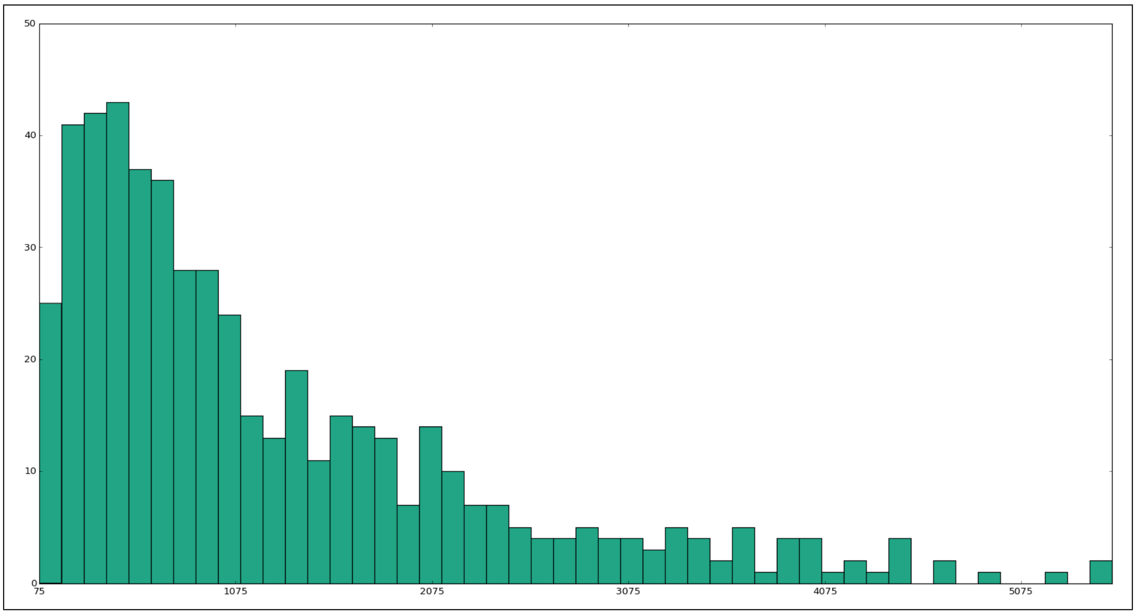


Figure 4: Output per university

journals with at least one co-author mentioning Spain in his or her affiliation between January 2004 and January 2015. I grouped those publications according to four broad topics, following Scopus' classification. The four topics are: physical sciences,[3] life sciences,[4], health sciences [5] and social sciences.[6] The analysis done here is focused on the physical and life sciences sectors, as they have been described as the sectors of focus in the Spanish national plans for scientific and technical research issued since 2004.

Figure 4 displays the histogram of the total output -internal and joint research projects- of the Spanish public universities. No university has a count of zero publications, and the maximum number of published outputs in a year is 5536.

The number of articles and conference proceedings in this sector has dramatically increased since 2008 for all universities, especially for those with the highest levels of productivity The four biggest universities are the University of Barcelona, the Politecníca de Cataluña, the Complutense de Madrid and the Politecníca de Valencia. The Politecníca de Valencia clearly improved its position after 2008. The least research intensive universities saw an increase in the number of articles, books and conference proceedings they

[3]This covers the following fields: Chemical Engineering,Chemistry,Computer Science,Earth and Planetary Sciences, Energy, Engineering, Environmental Science,Materials Science,Mathematics, Physics and Astronomy.

[4]This covers the following fields: Agricultural and Biological Sciences, Biochemistry, Genetics and Molecular Biology, Immunology and Microbiology,Neuroscience,Pharmacology, Toxicology and Pharmaceutics

[5]This covers the following fields: Medicine, Nursing,Veterinary Medicine,Dentistry and Health Professions

[6]6This covers the following fields: Arts and Humanities,Business, Management and Accounting,Decision Sciences,Economics, Econometrics and Finance,Psychology and Social Sciences.

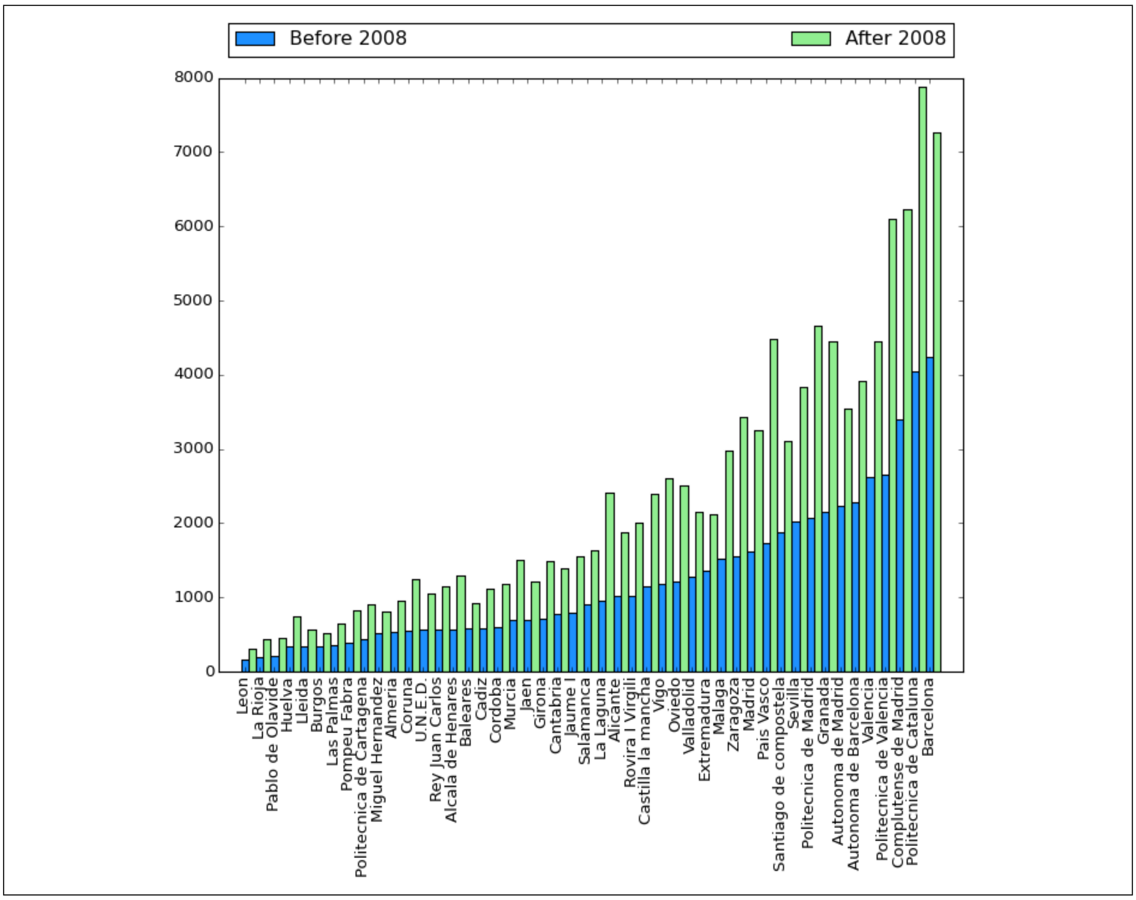


Figure 5: Evolution of the output per university, before and after the crisis

published since 2008, albeit not as strong.

3.4 Inter-university collaborations play a crucial role

I measure inter-university collaborations by the number of joint projects university i has with another institution. This is measured by the number of articles, conference proceedings, and books published by researchers affiliated with university i that have at least one co-author working at another institution. Each time two universities take part in a joint project, they are considered to be linked. This means that the corresponding entry in the 47 x 47 table of all possible interactions, or adjacency matrix, is non null.

One third of the projects are conducted with institutions in other countries, a rate stable across universities and increasing slightly over time. While those projects enter into the count of the number of articles, they do not enter into the mapping of relationships between universities.

For projects combining authors who list only Spanish institutions as their main affiliation, Table 1 shows the share of the total number of collaborative projects by type of partner and by region. The main collaborator was another public university in average 65 % of the time, with a minimum of 51% and a maximum of 84 %. The federal agencies, CSIC and CIEMAT (*Centro de Investigaciones Energ*{ticas, Medioambientales y Tecnol{

gicas) mainly, are the second biggest partners, with on average 28 % of the collaborative outputs made in collaboration with researchers from federally funded agencies. The regional agencies follow at the third place, with the maximum of 11 % for Cataluña, showing the strong regional support for R&D[7]. Private universities, as well as the private sector are negligible for basic research. Other aggregate the mixed type institution, and institutions that couldn't be matched.

Collaboration with (%)	Mean	Std. dev.	Min	Max
Public U	65.45	8.85	51.1	84.9
Federal agencies	20.21	6.48	11.3	35.9
Regional agency	3.44	2.86	0	11.9
Private U	0.82	0.59	0	2
European agencies	0.21	0.25	0	0.8
Private sector	0.29	0.42	0	35.9
Other	5.49	4.65	0	12.8

Table 1: Patterns of regional collaboration in percentage of the total collaborations, 2004-2008

Table 2 provides summary statistics for the joint research production. Across all periods, around 60 % of the research outputs were produced in collaboration with another university, as indicated by the collaboration rate. The average university partners with 35 other universities; its researchers produce 762 research outputs in collaboration with researchers affiliated with other institutions.

[7]Table 15 in the appendix gives detailed regional information on the type of collaborations

Finally, I provide indication on the Katz centrality metric. It is linked to the relative position of a university in the network of partnerships, and linked to the equilibrium concept of the model. While a more formal definition will be given later, the intuition behind is to reflect how important a university is for its partners.

	Mean	Std. Dev.	Min	Max
Collaborative outputs	762.248	667.58	8	3387
Number of partners	35.51	8.89	5	49
Collaboration rate	0.651	0.077	0.404	0.796
Katz centrality	.26	0.067	0 .1	0.35

Table 2: Descriptive statistics of partnerships

In terms of regional patterns, most of the collaborative research before the crisis was taking place within a region, or with universities from the more populated regions, Madrid and Catalunã. Table 16 in the appendix provides more detailed numbers on regional patterns of inter-university partnerships.

Figure 6 suggests that the number of collaborative projects is strongly linked to research output, hence making the case for looking at the network of inter-university partnerships. The upper panel shows the evolution of the distribution of research output, while bottom panel shows the distribution of the number of collaborative outputs. During the years of budget growth, the intensity of the partnerships grew as well, indicating a shift towards more collaborative projects with the partners. The distribution become flatter as more and more universities engage in research over time. After 2008, the number of the joint projects stagnated, reflecting the general trend in publications. Hence, collaborations were linked to the growth of the research output, and mirror the current stagnation.

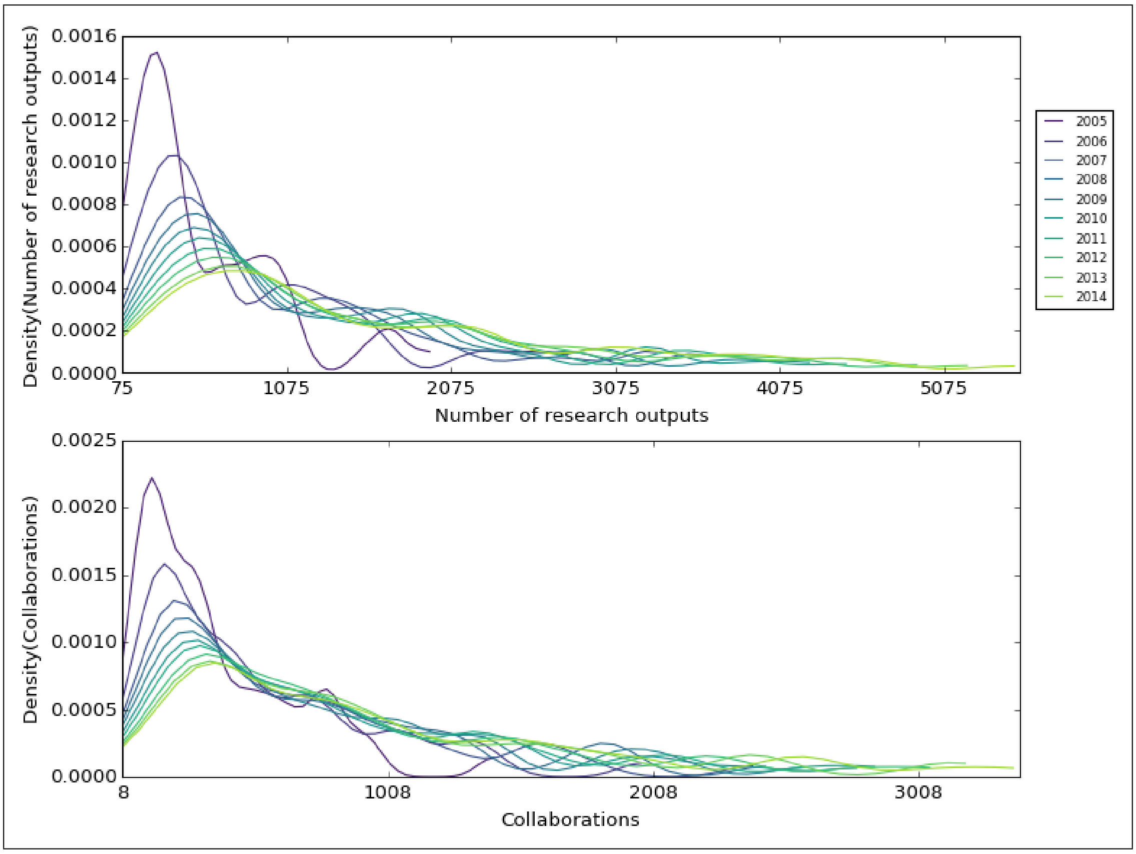


Figure 6: Kernel density estimation of the distribution of the number of articles published and the collaboration rate, 2005-2014

Looking at the institution level, Figure 7 indicates the evolution of the network of partnerships, normalized by the number of publications each year, between 2005 and 2008. For simplicity, the graph displays only the interactions that represent more than1 % of the total of the joint research projects. Each university node's size is proportional to the total number of publications linked to that university, while each link between universities represents the intensity of the research collaboration.

The network is still structured around the same major universities and shows the popularity of universities in Cataluña and the region of Madrid. The major differences between the two figures are the decrease in joint projects between the universities in Northwest Spain (Vigo, Coruña, and Santiago de Compostela) and the reduction in collaboration between the Baleares and universities in Cataluña.

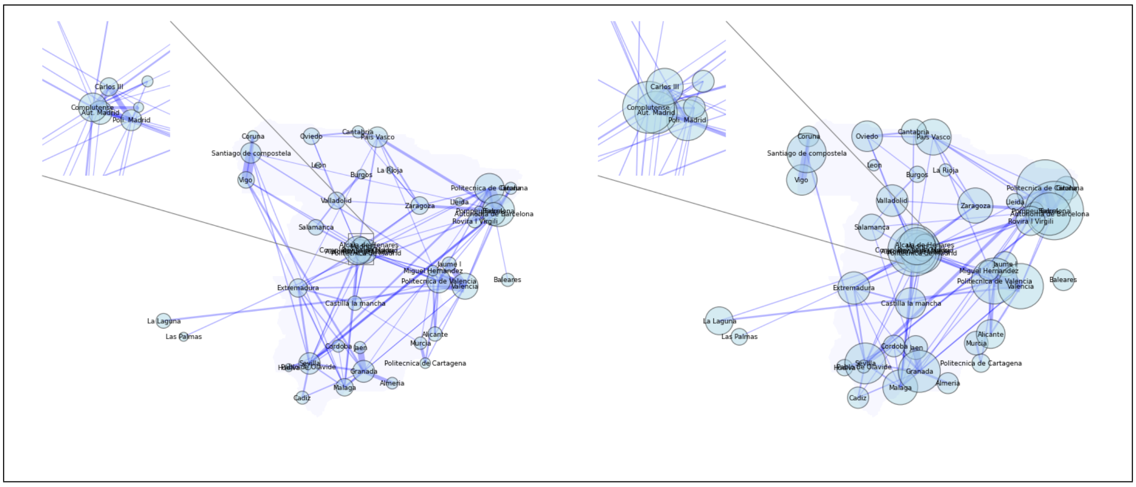


Figure 7: Network of inter-university collaboration, in 2005 and 2008

This suggests that the patterns of the network are stable over time, and not easily modified in the short-term. Partnerships are costly and are built on historical ties between universities that are unlikely to be affected by transitory shocks. Figure 8 shows that, while the more productive universities tend to engage more frequently in research partnerships, the least productive universities have heterogenous attitudes towards the balance between internal and external projects.

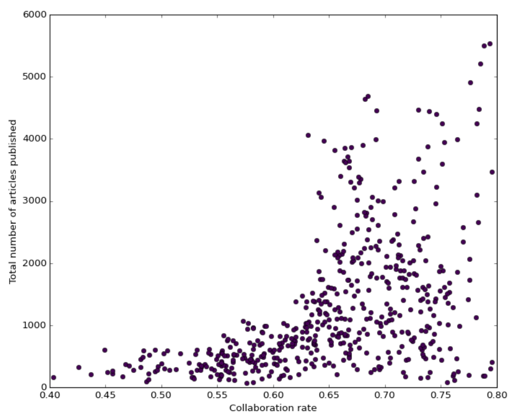


Figure 8: Correlation between number of articles published and the collaboration rate

Finally, Table 3 provides partial correlations between the number of research outputs and various university characteristics. As the theoretical model's equilibrium will be based on some variable characterizing the centrality of the university among the partnership networks, I include a centrality metric on the right hand side. All the models include year and university fixed effects.

	Q	Q	Q
log(Budgets)	167.65	151.70	-308.85
		(169.59)	(195.78)
log(# Students)	170.36***	-825.79***	-731.86***
		(174.59)	(158.46)
log(# Staff)	236.10	-122.00	145.71
		(240.56)	(222.29)
Collaboration rate		1,141.44**	
		(486.38)	
Katz centrality			3,119.14***
			(825.16)
ρ	0.95	0.95	0.97
R^2between	0.32	0.37	0.45
R^2overall	0.06	0.07	0.18

Table 3: Partial correlations of universities' characteristics and research output

In all cases, the variable describing the number of students is significant. After controlling for a partnership measure, either the collaboration rate or the centrality measure, the coefficient become negative. This means that once I have controlled for the relative eminence and collaboration behavior of the university, a change in the number of students is likely to induce a decrease in the research output. Intuitively, more collaborative and more central universities are bigger, as shown when describing the patterns of collaborations.

Centrality absorbs the positive effects of the size, such that adding extra students only increases the teaching burden. The collaboration rate as well as the centrality measure are both significant and correlated to a high research output. The share of the variance due to the fixed effect also increases, as seen in the parameter ρ. Adding the centrality measure increases the overall explanatory power of the model, as seen when comparing the R^2. This shows that the centrality parameter carries some information about the research performance and supplements the non-network approach. The model in the next section, as well as its empirical implementation, builds on this to explain the link between centrality and productivity.

Timing of the budget cuts and possible impact

Most of the reduction in budget devoted to higher education started in 2008 or 2009, along with the contraction of the Spanish economy, and deepened with the sovereign debt crisis in 2010. As stated in a report of the Bank of Spain (Ortega and Peñalosa (2012)), Spain had a fiscal surplus of close to 2% of GDP and a public debt/GDP ratio of 36% in 2007, but promoted pro-cyclical measures before the cris, deteriorating the budgetary situation. The fiscal deterioration due to the crisis was thus bigger than anticipated, not only because the scale of the recession was greater than expected, but also because the extraordinary revenue to which the real estate boom had given rise disappeared (Ortega and Peñalosa (2012)). The deficit grew to an unprecedented rate (almost 10% of GDP in 2009). In 2010, the deepening of the sovereign debt crisis and the emerging doubts over the sustainability of Spanish public finances quickened the need for the fiscal consolidation process.

The budget cuts experienced in the higher education sector intervened in this context. A quick review of the articles in published in the local newspapers at the time show that the first cuts were unanticipated in Madrid in 2008. The Government decided at the beginning of Fall to reduce the budget devoted to the universities of 418 millions (El Pais, 10/18/2008, El Pais, 16/10/2008). Two of the most important universities both in terms of their number of students and of their research output have seen their budget decrease in 2008 and 2009. During the Fall 2009, the federal Government announced the change in the national plan for science and research, that have seen the budget for RD go back to its 2006 level (El Pais, 10/09/2009). In 2010, the Federal Government announced that further cuts were to be implemented (La Vanguardia, 20/09/2009). Most of the universities dealt with the first budget cuts by reducing the investment levels for maintenance and construction works. This means that the highly indebted universities faced the shock quicker than their financially sane peers. Cataluña was spared until 2009-2010, when the Government deciding to cancel the 62 millions of planned funds for the Pompeu Fabra university and it announcing the reduction of the planned fundings for the universities. As the salaries are the biggest share the budget, the next step was to reduce the number of contractual staff in teaching and research and not to replace retired faculty .

Figure 16 shows the moment of the first shock, defined as a negative growth of the budget no matter the size of it, for the Spanish public universities. There is no clear geographical pattern for the first shock to happen, nor there is a strong correlation between the moment of the first shock and the research performance of the university. Universities from Madrid, Andalucía, Cataluña, Castilla y León and Aragón were the first to see their budget reduced in 2008-2009. However, several universities' budgets in the Autonomous Community of Madrid, Andalucía and Castilla y León started to decrease only after 2010.

Tables 17 and 18, in appendix, analyses the timing and impact of the budget cuts . Table 17 reproduces the same the same analysis as Table 3 but using the format of an event study. The time period are the year to and after the budget shock. While the effect of the main variables characterizing the universities is similar to the previous analysis, the interest lays in the time fixed effects. A simple test of parameters shows that, for each university, the output grew till two years after the budget shock, but stagnated afterwards. Table 18 shows, based on a test on the parameter of the years fixed effects, that the research output grew till 2012 and has since stagnated.

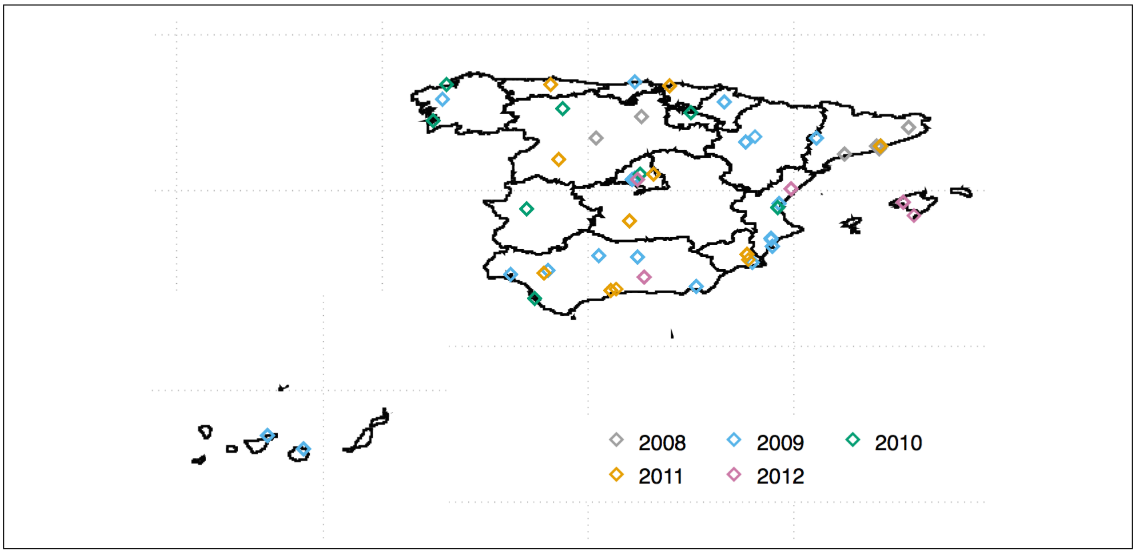


Figure 9: Year of first shock

4 Network model of inter-university spillovers

Each region determines its total budget for higher education. Within a region, each university budget depends on its relative share of the total number of students in the region, plus a transfer. The literature usually assumes that the universities seek to maximize their research output and, accordingly, allocate to teaching the minimum level of resources required to achieve a level of quality that is fixed by the government, which is measured by the student-to-teachers ratio.

To model the interactions between the public universities, I use a variation of the network model developed in Ballester et al. (2006). Each university can conduct projects either internally or in collaboration with another university, by investing into research projects those resources leftover after teaching. When deciding to work jointly on projects, this investment gives rise to a network of inter-university partnerships. In a network setting, each university i, $i = \{1, ..., N\}$ is then said to be a node in the network of universities and collaborations $\boldsymbol{G}$.

The focus of the present paper is the impact of a rapid change in budget after the shock to the original network structure. While long-term changes can be expected, previous partnerships involved building trust, common interests, and possibly common research facilities. Thus, while the manner in which a network arises is undoubtedly another important question, I consider it as given here.

At the beginning of each period of time, universities choose simultaneously which universities to collaborate with according to mutually beneficial links. If two universities collaborate on a joint project, a link exists between the two nodes in the network of inter-university partnerships.

In network notation, these joint projects are summarized by the adjacency matrix $\boldsymbol{G}$.

While other and less decentralized alliances could form in the long run by changing the cost of research, this is beyond my time frame. Given the network and its own ability to produce research, each university can decide whether or not to invest in common projects. Based on this, I am able to study the impact of budget shocks on the total productivity of universities and the mitigating or reinforcing effects of collaborations that can be seen as a way of reducing the cost of conducting research.

The utility university i has in t from the research production depends on its own research, y_{it}, and on the benefits it has from research conducted by its partners in the network, y_{jt} , $j \neq i$. This production function can be written as:

$$U(y_{it}, y_{jt}\Big|G_t) = (a_{it} + \nu_t + \epsilon_{it})y_{it} + \frac{1}{2}y_{it}^2 + \phi\sum_{j=1}^{N} g_{ijt}y_{it}y_{jt} \tag{1}$$

Where $\phi\ > 0$ and $\epsilon_{it} \sim (\sigma^2, 0)$. In this model, the universities are ex ante heterogenous in terms of observable and unobservable characteristics : $a_{it} = \mu_i + \sum_{m=1}^{M}\beta_m x_{it}^m$. This allows each university's objective and production function to depend on the m type of resources involved in the production of research, x_{it}^m, and on its unobservable characteristic, μ_i.

This university effect encompasses unchanging characteristics, such as the geographical position or the date of creation, but also unobservable intrinsic qualities. The heterogeneity between universities, in term of size and ability mainly, is taken has given ex ante and is an element of the production function. ν_t iis a period specific effect, allowing for each year to have a common impact on all universities. The idiosyncratic shock, ϵ_{it} is assumed to be known by the agents when deciding of their investments decisions but are unobserved by the econometrician.

I define α_i as $a_{it} + \nu_t + \epsilon_{it}$. The first two terms of (1), $\alpha_i y_i - \frac{1}{2}y_i^2$ give the benefits and cost for i of investing in y projects in t.

The last term of the right hand side, $\phi\sum_{j=1}^{N} g_{ijt}y_{it}y_{jt}$ gives the benefits that university i has from the investments of its peer j, $j \neq i$. If the cross derivative ϕ is positive or null, the game displays strategic complementarity in research efforts. If the cross derivative ϕ is negative or null, the game displays strategic substitutability in research efforts. Here, as a share of the output of university j will also be output of i, strategic complementarity in the production of research is a natural assumption. Thus, ϕ is assumed to be positive or null. Finally, the level of production is considered to be unbounded.

This simultaneous research production decision game is characterized by the following best reply function for university i:

$$y_{it}^* = \alpha_i + \phi\sum_{j=1}^{N} g_{ij,t}y_{jt} \tag{2}$$

$$\boldsymbol{y_t}^* = \boldsymbol{\alpha} + \phi\boldsymbol{G}\boldsymbol{y_t}^* \tag{3}$$

In matrix notation, aggregating the first order conditions of all the universities, it becomes:

The equilibrium effort of such a game is closely related to the specific centrality measure characterizing each university's relative position in the network. Katz (1953) and Bonacich (1987) define the following measure: Given a vector of weights $\boldsymbol{w} > 0$ and $\phi \geq 0$, the weighted Katz-Bonacich centrality of a node is given by:

$$\boldsymbol{b_w}(\boldsymbol{w}, \phi) = (\boldsymbol{I} - \phi \boldsymbol{G})^{-1} \boldsymbol{w} = \sum_{k=0}^{+\infty} \phi^k \boldsymbol{A}^k \boldsymbol{w}$$

The intuition behind the Katz-Bonacich centrality measure is simple: a university is central in the network of universities if its partners are also central. The exponential discounting indicates that the universities further away matter less for a university's centrality than its closest partners, and that the closest partners have to be well connected themselves to matter.

This leads to the following proposition about the characterization of the Nash equilibrium of the game with $y^*_{it} \geq 0$ Let λ_{min} denote the lowest eigenvalue of the adjacency matrix $\boldsymbol{G}$. Then, This game displays a unique and interior Nash equilibrium in pure strategies if $|\lambda_{min}(\phi \boldsymbol{G})| < 1$. In this case, the optimal effort level is given by:

$$\boldsymbol{y_i t}^* = \boldsymbol{b_\alpha}(\phi, \boldsymbol{A}) = \mathbf{M}(\phi, \boldsymbol{G}) \alpha \sum_{k=0}^{K} (\boldsymbol{I} - \phi \boldsymbol{G}^k) - 1 \quad (4)$$

(Ballester et al. (2006)) As noted in Jackson and Zenou (2014), the condition on the lowest eigenvalue of the network stipulates that the complementarities in production between universities must be small enough compared to its own concavity. If the condition does not hold, the feedback effects create a loop that can prevent the existence of a finite equilibrium.

The optimal effort level is closely related to the Katz-Bonacich centrality measure defined above: for a given vector of intrinsic abilities $\boldsymbol{\alpha}$, the optimal effort level is given by the Katz-Bonacich centrality weighted by the abilities' vector $\boldsymbol{\alpha}$. The matrix $\mathbf{M}(\phi, \boldsymbol{G})\alpha$ characterizing the equilibrium has an important interpretation: It aggregates all direct and indirect links among universities using an attenuation factor k that penalizes the contribution of collaborations between distant partners.

First, let us note that the utility of one university increases in its own valuation and ability to produce research α. When ϕ is positive, the utility of one university also increases in its partner research provision, as it benefits from a share of it.

Intrinsic abilities In order to understand the meaning of that equilibrium effort, let us look at a very simple example of the effect of the strategic interaction parameter ϕ with a simple network of $N = 5$. The figures below illustrate the impact of ϕ in the case with intrinsic abilities for various values of ϕ.

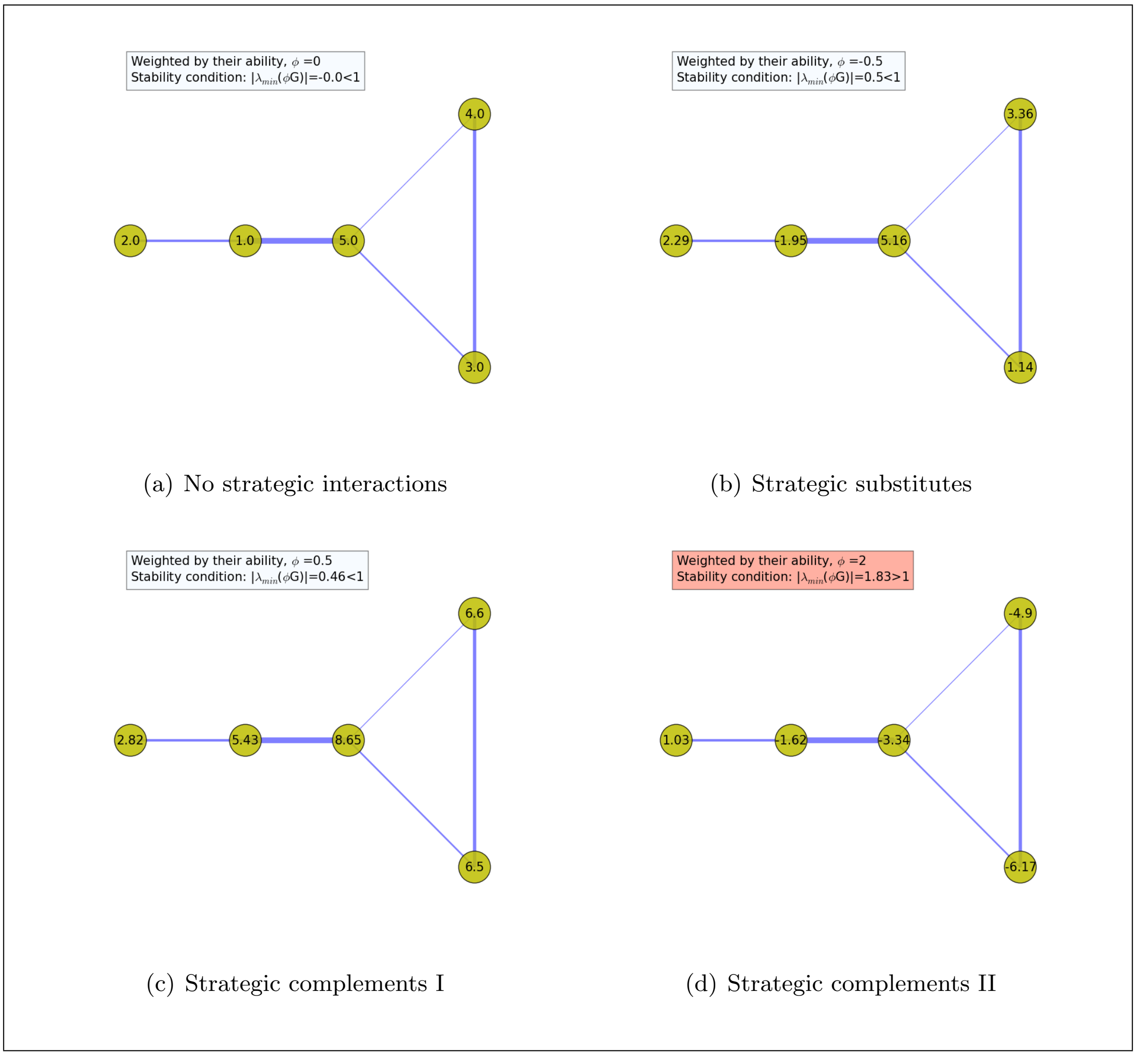


Figure 10: Equilibrium efforts of the game with intrinsic abilities for various ϕ

When $\phi = 0$, all players exert the amount of effort determined by their intrinsic abilities, as the actions of their neighbors do not impact them.

When $\phi < 0$, the game displays strategic substitutability, and the effort of one of the universities reduces the effort of its neighbors. The lower the ϕ, the stronger the effects of universities further away form university i. Thus, if the stability of the equilibrium condition holds as in Figure 2.9(b), *5* and *1*, are negatively affected by the efforts of the players further away from them. As *3* and *4* are also linked and display a lower effort level, *5* provides more than *1*.

When $\phi > 0$, the game displays strategic complementarity and the effort of one of the universities increases the effort of its partners. The higher the ϕ, the stronger the effects of universities further away from university i. If the stability condition is respected, as in Figure 2.9(c), the actions of the universities do not spiral out of control. Thus, the two universities that appear to be more central, *1* and *5*, are positively affected by the efforts of the universities further away from them. As *3* and *4* are also linked and display a higher effort level, *5* provides less effort than *1*. When the stability condition does not hold, as in Figure 2.9(d), the efforts of the games are out of control, and the centralities displayed are not the stable equilibrium effort levels.

4.1 Comparative statics

From this very simple model, comparative statics can be drawn from the analysis of changes in ϕ and $\boldsymbol{G}$. If the strategic complementarity parameter ϕ increases, the effort level of every university in the network will increase as well. On an aggregate level, the denser the local complementarities, the higher the aggregate research outcome. Intuitively, if the marginal return to joint work increases, then the number of joint projects increases, leading to an increase in each university's production, everything else held constant.

Considering an alternative social network $\boldsymbol{G'}$ such that $\boldsymbol{G} \subset \boldsymbol{G'}$. The network $\boldsymbol{G}$ is nested in the network $\boldsymbol{G'}$ such that all the links that exist in $\boldsymbol{G}$ exist in $\boldsymbol{G'}$, and $\boldsymbol{G'}$ has at least one more link between two universities than $\boldsymbol{G}$. In this case, as the network becomes more dense, for ϕ constant, the effort level of each university increases as well as the overall production.

4.2 Key player and intercentrality measure

Finally, let us consider the question faced by the social planner who would like to make budgets cuts while having the lowest impact possible on total production and welfare. Several policy suggestions exist in the literature stemming from Ballester et al. (2006)'s key player analysis and aimed at targeting the group with the highest impact. I use a variation of this measure, intending to have the smallest possible impact.

First, let us define the key player as the university that, once removed from the inter-university network, induces the largest drop in total production due to direct and indirect effects. This provides only a short-term evaluation of the problem, as the network does not react immediately to that deletion. Formally, the problem of the social planner is finding university i such that

$$\max_i \{y^*(\boldsymbol{G}) - y^*(\boldsymbol{G}_{-i})\} \tag{5}$$

where $\boldsymbol{G}$ is the existing network and $\boldsymbol{G}_{-i}$ is the network remaining after university i and its links have been removed. The solution of this problem is given by a specific centrality measure, the intercentrality parameter. Consider a network $\boldsymbol{G}$ characterized by the squared adjacency matrix $\boldsymbol{A}$ and a scalar s such as $\boldsymbol{M}(\boldsymbol{G}, s) = [I - s\boldsymbol{G}]^{-1}$ is well-defined and non negative. The inter-centrality of a node of parameter s in the network $\boldsymbol{G}$ is given by :

$$c_i(\boldsymbol{G}, s) = \frac{b_i^2(\boldsymbol{G}, s)}{m_{ii}(\boldsymbol{G}, s)}$$

Ballester et al. (2006) show that the solution to the problem of equation 5 is given by the university with the highest inter-centrality measure.

Here, as the goal is to minimize the impact of budget cuts, I aim to find the university for whom a reduction in funding would have the least effect on the production of the higher education and research system. This exercise, in turn, suggests possible ways of minimizing the impact of targeted budget cuts, i.e., designing them such that they have the least effect on the whole production of science. While science production is not the only goal of universities, and reducing funding for research is not the only concern, I show in the empirical part how targeted cuts may reduce the impact on research.

5 Empirical analysis

5.1 Empirical specification

In order to implement the key player policy, I first need an estimate of the network effect in the inter-university networks, ϕ, and of the vectors of universities' effects, α_i. In this section, I introduce the econometric equivalent of the model presented above. Adding subscript t for time to 3, its expression is given by :

$$y_{it} = \alpha_i + \boldsymbol{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \eta_t + \epsilon_{it} + \sum_{j=1}^{n} g_{ijt} y_{jt} \tag{6}$$

where y_{it} is the best reply for university i at time t, x_{nt} is the n x k matrix of observed exogenous characteristics of the university i in t, α_i is the university fixed effect, and η_t is the time fixed effect. ϵ_{it} is $i.i.d$ across n and t with mean 0 and variance σ^2.

The only difference between the theoretical and the econometric model is that I allow the network links g_{ij} to potentially vary over time, and I use the log trans- formed variables rather than the raw numbers. The adjacency matrix used is the full adjacency matrix. Formally:

$$Full\ adjacency_{ij,t} = \begin{cases} w\ if\ \ i\ works\ with\ j\ in\ t, t-3 \\ 0\ otherwise \end{cases}$$

This is a natural candidate for the adjacency matrix, as it summarizes all the interactions between universities for a given period of time. I come back to this specification in the identification section and discuss its implications for identification[8]. In vector-matrix form, I can write equation 6 as

$$\boldsymbol{y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{X}_t + \eta_t \boldsymbol{l}_n + \phi \boldsymbol{G}_t \boldsymbol{y}_t + \boldsymbol{\epsilon}_t \tag{7}$$

[8]Other adjacency matrices and results are presented as robustness checks in Section 7

where $\boldsymbol{y}_t = (y_{1t}, y_{2t}, ..., y_{nt})^{\mathrm{T}}$, $\boldsymbol{X_t} = (\boldsymbol{x}_{1t}, \boldsymbol{x}_{2t}, ..., \boldsymbol{x}_{nt})^{\mathrm{T}}$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)^{\mathrm{T}}$, and $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \epsilon_{2t}, ..., \epsilon_{it}, ..., \epsilon_{nt})^{\mathrm{T}}$ are n-dimensional vectors and $\boldsymbol{l}_n$ is a n-dimensional vector of ones. $\boldsymbol{G}_t = [g_{ijt}]$ is the n x n column-normalized adjacency matrix of the network in t.

For the T periods, 7 can be written as:

$$\boldsymbol{y} = \boldsymbol{\alpha} \otimes \boldsymbol{l}_T + \boldsymbol{\beta X} + \boldsymbol{l}_n \otimes \boldsymbol{\eta} + \phi \; diag\{\boldsymbol{G}\}\boldsymbol{y} + \boldsymbol{\epsilon} \tag{8}$$

where $\boldsymbol{y} = (\boldsymbol{y}_1, ..., \boldsymbol{y}_T)$, $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_T)$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$, $\boldsymbol{l}_T$ is a T-dimensional vector of ones, and $\boldsymbol{\epsilon_t} = (\epsilon_1, ..., \epsilon_T)$ is the $n \times 1$ vector of the idiosyncratic error . All vectors are of dimension $(nT \;\; \times \;\; 1)$, where T is equal to 11 for the 11 years of our panel and n is equal to 47 for the 47 public universities described above. $diag\{\boldsymbol{G}\}$ is the $(nT \;\; \times \;\; nT)$ block diagonal matrix containing the network weights at each time:

$$\text{diag } \{\mathbf{G}\} = \begin{bmatrix} \boldsymbol{G}_1 & 0 & ... & 0 & 0 & 0 \\ 0 & \ddots & 0 & ... & 0 & 0 \\ 0 & 0 & \boldsymbol{G}_t & 0 & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & .... & ... & \boldsymbol{G}_T \end{bmatrix}$$

I use the number of articles, proceedings, and books to capture the research output, y_{it} and the number of students, the number of teaching and research staff, and the total budget to capture the universities' characteristics, $\boldsymbol{x}_{it}$. The network in t is the weighted network constructed on the basis of the past three years co-publications, as presented above. As the spread of the data is likely to lead to heteroskedadesticity, I transform the variables to logs.

5.2 Identification and IV estimation

The approach here is structural in the sense that I estimate directly the first order condition of the universities' maximization problem. The econometric model corresponding to the first order condition is comparable to a spatial autoregressive model where the spatial weight matrix is given by the adjacency matrix of the network. As in the spatial regressive model, the spatial lag $\boldsymbol{G}_t\boldsymbol{y}_t$ is endogenous: the production of university i takes the production of j into account, and vice versa. I use an IV approach to tackle this problem, and instrument the partners' research level by their production factors, $\boldsymbol{G}_t\boldsymbol{x}_t$. This takes the network as given at a particular point of time, and estimates the game played on it, knowing the universities exogeneous variables.

This identification strategy is valid if $\boldsymbol{G}_t \boldsymbol{x}_t$ is independent of the error term in t. If this is plausible for $\boldsymbol{x}_t$ in a context of very stable institutional background before the crisis and rigid labor market rules, this is a strong assumption for $\boldsymbol{G}_t$. As commonly argued in the literature on networks, the fixed effect somehow mitigates the extent of the problem, as any time-invariant or individual-invariant element that would influence both the error term and the network structure will be swept away.

To avoid the incidental parameter problem when estimating the model in equation 8, I follow Lee and Yu (2014) and König et al. (2014) and I remove the time and the cross sectional means of the observations to eliminate the time and individual effects.

Formally, the time stacked model becomes

$$\boldsymbol{J}\boldsymbol{y} = \boldsymbol{J}\beta\boldsymbol{X} + \phi\, \boldsymbol{J}\boldsymbol{G}\boldsymbol{y} + \boldsymbol{J}\boldsymbol{\epsilon}$$

where $\boldsymbol{J} = \boldsymbol{J}_T \otimes \boldsymbol{J}_n$ and $\boldsymbol{J}_T = \boldsymbol{I}_T - \frac{1}{T}\boldsymbol{u}_T\boldsymbol{u}_T^{\mathrm{T}}$ and $\boldsymbol{J}_n = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{u}_n\boldsymbol{u}_n^{\mathrm{T}}$ where $\boldsymbol{I}_T$ is the $(T \times T)$ identity matrix, $\boldsymbol{u}_T$ is a $(T \times 1)$ vector of one, $\boldsymbol{I}_n$ is the $(n \times n)$ identity matrix, and $\boldsymbol{u}_n$ is a $(n \times 1)$ vector of ones.

Based the identification strategy described above, I use, $\boldsymbol{Q} = [\boldsymbol{JX}, \boldsymbol{JGX}]$ as the set of instruments. Let us define $\epsilon(\theta) = \boldsymbol{J}(y - \boldsymbol{Z}\theta)$, where $\boldsymbol{Z} = [\boldsymbol{GY}, \boldsymbol{X}]$ and $\theta = [\phi, \beta]$. According to the identification assumption, I have the following moment condition $\boldsymbol{Q}'\epsilon(\theta) = 0$. The standard errors are cluster robust.

5.3 Exogeneity of G

The identification strategy above relies on the assumption of exogeneity of $\mathbf{G}$ and the error term at each period of time. I address this assumption in the following analysis. I compare the estimation with the assumption of the exogenous network to the estimation based on the assumption of a predetermined network. In this case, current errors can be correlated to future values of the network variable.This estimation strategy can be seen as estimating the second step of a game of network update and research decisions[9]. In this version of the game, the agents are randomly selected and allowed to update one of their links in the network in the first stage. In the second stage, the game used in this paper is played on the preexisting network structure. While in this version the strategic selection of partners is not accounted for, it allows for some rewiring of the network of partnerships[10]. The intrinsic ability component is a bit less easy to interpret compared to the exogenous case and the key player analysis as described above should be understood as taking place between two re-wiring of the network. The key player policy will be based on the observed characteristics and the average of the prediction error, as is the case in the rest of the network literature using cross-sectional data, as reviewed in Zenou (2014).

To be more precise, not only $\boldsymbol{y}$ is endogenous, but now $\boldsymbol{GX}$ is also predetermined, i.e. that the past realizations of the network and characteristics of the partners are

[9] This is a version of the " morning afternoon" game as described in Liu et al. (2012)

[10] This clearly leaves aside the question of the joint estimation of the endogenous network formation process and the outcome, which seems to be the next challenge for this branch of literature

correlated to the future productivity shocks. Formally, $E(\boldsymbol{\epsilon}_s|\boldsymbol{GX}_t) \neq 0$ if $s < t$ while $E(\boldsymbol{\epsilon}_s|\boldsymbol{X}_t) = 0 \;\forall\; s, t$.

The literature of panel data with predetermined variables suggest using the Helmert transformation, as Arellano and Bover (1995), to sweep away the time fixed effects while preserving the orthogonality between the variables[11]. Intuitively, this is a procedure used to remove the means of all future observations. The operator that produces this transformation is given by the $(T \times T-1)$ matrix F :

$$F = diag\left[\frac{T-1}{T}, ..., \frac{1}{2}\right]^{\frac{1}{2}} \begin{bmatrix} 1 & -(T-1)^{-1} & -(T-1)^{-1} & ... & -(T-1)^{-1} & -(T-1)^{-1} \\ 0 & 1 & -(T-2)^{-1} & ... & -(T-2)^{-1} & -(T-2)^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & .... & 1 & -1 \end{bmatrix}$$

Instead of equation 8, I stack the vector of equation 7 as:

$$\boldsymbol{y}^* = \boldsymbol{l}_T \otimes \boldsymbol{\alpha} + \beta \boldsymbol{X}^* + \boldsymbol{\eta} \otimes \boldsymbol{l}_n + \phi \boldsymbol{G}\boldsymbol{y}^* + \boldsymbol{\epsilon}^*$$

Where $\boldsymbol{y}^*$ is a $(n \times T)$ matrix given by $[\boldsymbol{y}_1^{\mathrm{T}}, ..., \boldsymbol{y}_T^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{Gy}^*$ is a $(n \times T)$ matrix given by $[\boldsymbol{Gy}_1^{\mathrm{T}}, ..., \boldsymbol{Gy}_T^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\epsilon}^* = [\boldsymbol{\epsilon}_1^{\mathrm{T}}, ..., \boldsymbol{\epsilon}_T^{\mathrm{T}}]^{\mathrm{T}}$. $\boldsymbol{X}_k$ is equal to $[\boldsymbol{X}_{1,k}^{\mathrm{T}}, ..., \boldsymbol{X}_{T,k}^{\mathrm{T}}]^{\mathrm{T}}$ for the k variables encompassing universities characteristics. Then, the $n \times (T-1)$ matrix $\boldsymbol{y}^{**} = [\boldsymbol{y}_1^{\mathrm{T}}, ..., \boldsymbol{y}_{T-1}^{\mathrm{T}}]^{\mathrm{T}}$ is defined as .

$$\boldsymbol{y}^{**} = \boldsymbol{y}^* F = \boldsymbol{\beta X}^* F + \phi\; \boldsymbol{Gy}^* F + \boldsymbol{\epsilon}^* F$$

This removes the individual effects. Note that this formulation implies that the network interaction term as a whole, $\boldsymbol{GX}$, is considered as predetermined. I transform the model further to remove the time fixed effects, using $\boldsymbol{J}_n$ as defined above:

$$\boldsymbol{Jy}^{**} = \boldsymbol{J}\beta \boldsymbol{X}^{**} + \phi \boldsymbol{JGy}^{**} + \boldsymbol{J\epsilon}^{**}$$

The matrix of instruments now contains the predetermined variable up till $t-1$, and the exogenous variables in t, $\boldsymbol{Q}_{bis} = [\boldsymbol{GX^{**}}_{t-1}, \boldsymbol{X}^{**}{}_t]$, or $\boldsymbol{Q}_{ter} = [\boldsymbol{JGX^{**}}_{t-1}, \boldsymbol{JX^{**}}_{t_1}]$ if $\boldsymbol{X}$ is also consider as predetermined. The moments conditions are similar to the ones defined above using $\boldsymbol{Q}_{bis}$. I use a finite moment approach, as above and the standard errors are cluster robust.

5.4 Interpretation and average effects

Finally, it is worth noting that the introduction of a spatial lag changes the interpretation of the coefficients. The reduced form equation of 7 is given by:

$$\boldsymbol{y}_t = (\boldsymbol{I} - \phi \boldsymbol{G})^{-1}(\boldsymbol{\alpha} + \boldsymbol{\beta X}_t + \eta_t \boldsymbol{l}_n) + (\boldsymbol{I} - \phi \boldsymbol{G})^{-1}(\boldsymbol{\epsilon}_t) \quad (9)$$

[11]This procedure has also been used in Lee and Yu (2014) to address the issue of dynamic spatial panel data. However, the network itself is assumed to be invariant and exogenous.

$(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}$ can be rewritten as $(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}=\boldsymbol{I}+\phi\boldsymbol{G}+\phi^2\boldsymbol{G}+\phi^3\boldsymbol{G}$.... The direct effects of an explanatory variable, $\boldsymbol{X}$, are the diagonal elements of $(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}\beta$, while the indirect effects are the out-of-diagonal elements of the matrix. In this case, the betas reported are the first order direct effects, without feedback, the first order indirect effects are given by $\phi\boldsymbol{G}*\beta$, ... Hence, the total direct effects are greater or equal to the reported coefficient.

If the matrix of network weights is row normalized, the network effect ϕ can be interpreted as the average effect of the partners on the university. If the matrix is column normalized, then the network effect can be interpreted as the impact of a change in the exogenous variable on the dependent variable of all other universities.

I use the full matrix of weights, so this interpretation no longer holds. As the literature suggests, I report the simulated average total effect, the average of the average marginal effects[12]. In this case, the average total effect of a variable is computed as:

$$ATE_k = T^{-1}n^{-1}(\boldsymbol{I}-\phi\boldsymbol{G}))^{-1}\mathbf{1}\beta_k \tag{10}$$

Where $\mathbf{1}$ is a $n\ x$ 1vector of ones. They can be interpreted as the average total of indirect and direct effects of a change in a variable.

6 Estimation results

Table 4 presents the results for the models with exogenous and predetermined network matrices before the crisis[13]. It is clear that, among the own elements of the production function, the main explanation of current production is the university size as represented by the number of teaching and research staff and . This coefficient is significant whatever the type of normalization applied to the network. The coefficient for the budget is seldom significant. This is unsurprising, as all the models include university and time fixed effects, so a the similar growth of all universities before the crisis is absorbed by the fixed effects.

The elasticity of the total research output with respect to the number of own staff (in ten thousand) directly and without feedbacks ranges between 0.311 and 0.355 for all models[14]. For the budget, this elasticity is negative and between -0.111 and -0.138.

The network multiplier itself is always significant, and considering it as predetermined does not change the coefficient. The effect is positive, for all models, and bigger when considering the column normalized model, which takes into account the relative weight of the partners.

Finally, I report the J-stat for overidentification restriction. As the null hypothesis is that the instruments are jointly valid, not rejecting it is usually interpreted as a sign of

[12]The simulation procedure consists of computing the standard errors of the effects based on draws of the parameters and the covariance matrix, as in Lesage (2008).

[13]The first-stage regression results are found in Table19 in the appendix. The F stat indicates that the instruments are not weak (Kleibergen and Paap (2006)).

[14]As discussed before, the total and total indirect effects are reported in the next section

the exogeneity of the network. Here, the null hypothesis can be rejected at 5% casting doubt on the joint validity of the instruments.

	Log(Total research), Full adj.		Log(Total research), colum. norm. adj.	
	Exog.	Predet.	Exog.	Predet.
Network	0.0000525 ***	0.0000568 ***	0.084 ***	0.091***
	(0.0000177)	(0.0000189)	(0.011)	(0.014)
log(# students)	-0.011	-0.04	-0.005	-0.011
	(0.116)	(0.128)	(0.112)	(0.123)
log(# staff)	0.345**	0.311**	0.355***	0.318**
	(0.136)	(0.145)	(0.119)	(0.127)
log(# budget)	-0.111*	-0.118*	-0.126**	-0.138**
	(0.066)	(0.07)	(0.056)	(0.060)
Constant	-0.667***	-0.696***	-0.691***	-0.725***
	(0.036)	(0.041)	(0.023)	(0.025)
Observations	460	414	460	414
R^2	0.941	0.844	0.958	0.889
J	7.095	6.928	7.281	7.758
(pval)	(0.029)	(0.031)	(0.026)	(0.021)

Year fixed effects not reported
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimates of the static network collaboration models, total research produced

As the total production of articles counts joint research papers, I also use articles produced by teams entirely composed of researchers within the same university as the dependent variable. This gives an idea of the choices made by each university for their internal production, when taking into account their existing collaborations.

In this case, as shown in Table 5, the number of staff of the university and the budget are the main explanatory components[15]. The elasticity of the total research output with respect to the number of own staff (in ten thousand) directly and without feedbacks ranges between 0.428 and 0.468 for all models[16]. For the budget, this elasticity is negative and between -0.221 and -0.207, meaning that a decrease in budget is linked to an increase in the internal research produced. Again, this may look counterintuitive, but the research production hasn't collapsed during and after the crisis, and even continued to grow. For both variables, the elasticity is bigger in absolute value than in the total

[15]The first-stage regression results are found in Table 19 in the appendix. The F stat indicates that the instruments are not weak (Kleibergen and Paap (2006)).

[16]As discussed before, the total and total indirect effects are reported in the next section

research case, indicating that internal research is more sensitive to changes in the factor of production when controlling for network effects.

The network multiplier itself is always significant, and considering it as predetermined does not change the coefficient. The effect is positive, for all models, and bigger when considering the column normalized model, which takes into account the relative weight of the partners. The network multiplier is still strong and close to the level of the models using the total research effort. In the predetermined case, the elasticity of internal research projects with respect to the partners research projects (without feedback effects) is 0.0000536 for the full adjacency matrix and 0.075 for the column normalized matrix.

	Log(Internal research), Full adj.		Log(Internal research), colum. norm. adj.	
	Exog.	Predet.	Exog.	Predet.
Network	0.0000536**	0.0000583**	0.075***	0.089***
	(0.0000263)	(0.0000249)	(0.012)	(0.014)
log(# students)	0.298	0.284	0.318	0.324
	(0.241)	(0.230)	(0.234)	(0.223)
log(# staff)	0.468***	0.443***	0.467***	0.428**
	(0.161)	(0.171)	(0.158)	(0.167)
log(# budget)	-0.210**	-0.207*	-0.210**	-0.221**
	(0.103)	(0.109)	(0.094)	(0.100)
Constant	-0.737***	-0.765***	-0.744***	-0.775***
	(0.052)	(0.058)	(0.039)	(0.043)
Observations	470	424	470	424
R^2	0.898	0.763	0.905	0.779
J	3.078	3.319	1.198	1.515
(pval)	(0.215)	(0.190)	(0.549)	(0.469)

Year fixed effects not reported
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimates of the static network collaboration models, research produced internally

Finally, let's note that in the case of this case, the J-stat suggests that the instrument are jointly valid, giving credibility to theses results.

6.1 Average total and marginal effects

As discussed in Section 5.4, the betas cannot be interpreted as marginal effects due to the presence of the spatial lags. I report here the average total effects, that includes direct and indirect effects of any order.

	Full adj.		Column normalized adj.	
Internal res.	2.5291***	2.0849***	0.0621***	0.0723***
(std e)	1.7	1.3	0.0	0.0
Total res	9.1646***	3.1249***	0.0744***	0.0794***
(std e)	1.7	2.7	0.0	0.0

Table 6: Average total effects

In the full adjacency matrix case, the elasticity internal research to partner's researcg considering the network multiplier and the feedback, is equal to 9.16 for the full adjacency case and 0.074 in the column normalized case[17]. This mean that, when taking into account the feedback and adjustments, an increase in a university's budget will have a positive impact on the research produced internally by its partners. It takes into account the indirect effects of the partners weighted by the network multiplier and the feedback effects.

This indicates that diversification matters, as well as the identity of the partners, to explain the sensitivity of the reaction. These parameters are taken into account with the key player, describing both the spread of the partnership and the characteristics of the partners.

7 Robustness checks

7.1 Dynamic model estimates

As the models tested might be dynamic due to the high persistence of the research level, I run regressions including the lag dependent variable, using a dynamic model. While the coefficients did not change as much when comparing exogenous to pre-determined network, the inclusion of past production significantly reduces greatly the impact of the network variable, and makes the overidentification restrictions less convincing.

[17]As explained earlier, effects in the column normalized case can be thought of a weighted average out, considering what each university does with its partners, so averaging again doesn't bring additional information.

	Log(Total research), Full adj.		Log(Total research),,Col. norm adj.	
Exog.	Dyn.	Exog.	Dyn.	
Network	0.0000536***	0	0.084***	0.020***
	(0.0000263)	(0)	(0.011)	(0.006)
log(# of students)	-0.011	-0.121***	-0.005	-0.116***
	(0.116)	(0.031)	(0.112)	(0.037)
log(# of staff)	0.345**	-0.005	0.355***	-0.000
	(0.136)	(0.044)	(0.119)	(0.049)
log(budget)	-0.111*	0.037	-0.126**	0.002
	(0.066)	(0.036)	(0.056)	(0.037)
Lag(log(Total. Res.))		0.875***		0.803***
		(0.032)		(0.034)
Constant	-0.667***	0.021***	-0.691***	0.019***
	(0.036)	(0.006)	(0.023)	(0.005)
Observations	460	368	460	368
R^2	0.941	0.938	0.958	0.940
J	7.095	4.077	7.281	3.826
(pval)	(0.0288	0.1302	0.0262	0.1477

Year fixed effects not reported

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Comparison of the static and dynamic network collaboration models, total research produced

Results in Table 7, show that, for the dynamic models for total research produced, while the coefficient of the network variable is still significant when looking at the impact of the weighted share of the budget the partner allocates to common projects, it is not when considering the full matrix of collaborations. As in the static case, the J-stat indicates that we are likely to reject the null hypothesis of the validity of identification restrictions and, hence, the validity of the identification approach.

	Log(Internal research), Full adj.	
	Exog.	Dyn.
Network	0.000**	0*
	(0)	(0)
log(# of students)	0.298	-0.138***
	(0.241)	(0.039)
log(# of staff)	0.468***	0.062
	(0.161)	(0.059)
log(budget)	-0.210**	0.033
	(0.103)	(0.044)
lag(Log(Internal res.))		0.871***
		(0.24)
Constant	-0.737***	0.016**
	(0.052)	(0.008)
Observations	470	378
R^2	0.898	0.919
J	3.078	6.094
(pval)	0.2146	0.0475

Year fixed effects not reported
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Comparison of the static and dynamic network collaboration models, research produced internally

Results in Table 8, show that, for the dynamic models for internal research produced, the results are robust. The coefficient of the network variable is still significant, even if negligible in size in the full adjacency case. The J-stat indicates that the joint validity of instruments can be rejected for the dynamic case at 5 %.

7.2 Other matrix normalization

As the matrix normalization used here is somewhat unusual with respect to the existing literature that favors row-normalization, I also compare with the standard normalizations. Tables 9 and 10 reproduce the analysis of the static models using alternative interaction matrices. The first two columns of the tables show the binary network matrix, i.e. G's elements in period t are given by

$$Binary\ adjacency_{ij,t} = \begin{cases} 1\ if\ \ i\ works\ with\ j\ in\ t, t-3 \\ 0\ otherwise \end{cases}$$

I thus use as well a normalization of the adjacency matrix that links the outcome of one university to the partner's relative diversification:

$$Column-normalized\ adjacency_{ij,t} = \begin{cases} \frac{w_{ij}}{\sum_{i=1}^{I} w_{ij}}\ if\ \ i\ works\ with\ jint, t-s \\ 0\ otherwise \end{cases}$$

In the latter case, the weights between universities, each entry of G in t, are determined by the number of projects in common divided by the total number of projects the partner university, university j, has in t. This captures the element of the relative intensity of i and j's partnership, by putting the emphasis on the resources allocated by j to joint projects with i

While the normalization used in the previous sections allows for different partners to have a different weights as determined by the strength of the collaboration, this case considers all links and partnerships as having the same impact. Finally, the last two columns show the impact of what is referred to in the literature as the local average of neighbors effort, as it uses a row normalization for the network matrix.This makes the outcome of a particular university depend on the average of what its research partners are doing. In this case, G's elements in period t are given by

$$Row - normalized\ adjacency_{ij,t} = \begin{cases} \frac{w_{ij}}{\sum_{j=1}^{J} w_{ij}}\ if\ \ i\ works\ with\ j\ in\ t, t-3 \\ 0\ otherwise \end{cases}$$

where w_{ij} is the number of projects i and j have completed in common. Clearly, comparing these models does not give precise information about the coefficients themselves, as they use different networks. However, the binary interaction matrix can be seen as the minimal network effect in my analysis: if, regardless of the identity or the strength of the partner and partnership, conducting research in partnership with another university is useful, the network coefficient should still be positive. This is the case here, indicating that taking into account the network effects should always be considered when explaining the research production of universities.

	Log(Total research), row norm. adjacency		Log(Total research), binary. adjacency	
	Exog.	Predet.	Exog.	Predet.
Network	0.057	0.065	0.002***	0.002***
	(0.153)	(0.169)	(0.001)	(0.001)
log(# students)	0.012	-0.035	-0.013	-0.078
	(0.115)	(0.126)	(0.112)	(0.129)
log(# staff)	0.309**	0.280**	0.333***	0.312**
	(0.122)	(0.130)	(0.116)	(0.123)
log(# budget)	-0.037	-0.049	0.009	0.004
	(0.063)	(0.070)	(0.056)	(0.062)
Constant	-0.692***	-0.724***	-0.623***	-0.624***
	(0.104)	(0.121)	(0.041)	(0.046)
Observations	460	414	460	414
R^2	0.938	0.837	0.941	0.846
J	0.087	0.152	1.343	1.402
(pval)	(0.958)	(0.927)	(0.511)	(0.496)

Year fixed effects not reported
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Estimates of the static network collaboration models, total research produced with the alternative interaction matrices

Table 9 shows the estimates for models using the total research produced as the dependent variable. Again, the log of the number of staff at the university is, among the variables describing the internal elements of the university production function, the main explanatory element. The budget is no longer significant, nor is the network effect in the case of row-normalized adjacency. For the binary matrix, the network effect coefficients are higher and the number-of-students coefficients larger in absolute value.

	Log(Internal research), row norm. adjacency		Log(Internal research), binary adjacency	
	Exog.	Predet.	Exog.	Predet.
Network	0.149	0.191	0.002*	0.002*
	(0.333)	(0.335)	(0.001)	(0.001)
log(# students)	0.351	0.308	0.288	0.256
	(0.234)	(0.226)	(0.240)	(0.229)
log(# staff)	0.405**	0.391**	0.438***	0.411**
	(0.159)	(0.169)	(0.155)	(0.166)
log(# budget)	-0.144	-0.154	-0.108	-0.097
	(0.108)	(0.110)	(0.096)	(0.100)
Constant	-0.755***	-0.778***	-0.688***	-0.690***
	(0.122)	(0.133)	(0.065)	(0.075)
Observations	470	424	470	424
R^2	0.889	0.730	0.900	0.767
J	1.997	1.745	0.291	0.440
(pval)	(0.368)	(0.418)	(0.865)	(0.802)

Year fixed effects not reported
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Estimates of the static network collaboration models, research produced internally and alternative interaction matrices

Table 10 shows the estimates for models using research produced internally as the dependent variable. The log of the number of staff at the university is the main explanatory element. The budget is not significant, nor is the network effect in the case of row-normalized adjacency. For the binary matrix, the network effect coefficients are higher and the number-of-students coefficients larger in absolute value. For both tables, the J-stat suggests the validity of the instruments for the case of the binary matrix.

The binary adjacency matrix can be seen at the minimal network, as any link between two partners has exactly the weight. In this minimal case, the results show that the network effects are still to be controlled for when evaluating research production.

7.3 Length of collaboration

As the model presented above depends crucially on the interaction matrix used, I also perform tests of the impact of the network effect when changing its specification. Table 11 presents the results when considering the previous four years of collaborations as the meaningful interactions. As a robustness check, I consider the case with the adjacency matrix constructed with t to $t-4$ data.

	Log(Internal research), Full adjacency		Log(Internal research), c. norm. adjacency	
	Exog.	Predet.	Exog.	Predet.
Network	0.000 *	0.0000185*	0.064***	0.066***
	(0.000)	(0.0000192)	(0.013)	(0.013)
log(# of students)	0.387*	0.371**	0.363*	0.366**
	(0.209)	(0.183)	(0.205)	(0.178)
log(# of staff)	0.388***	0.328**	0.398***	0.328**
	(0.133)	(0.133)	(0.134)	(0.133)
log(budget)	-0.171*	-0.188*	-0.188**	-0.201**
	(0.094)	(0.097)	(0.086)	(0.089)
Constant	-0.874***	-0.928***	-0.850***	-0.902***
	(0.047)	(0.050)	(0.036)	(0.038)
Observations	470	423	470	423
R^2	0.939	0.871	0.944	0.881
J	3.369	3.303	3.483	3.342
(pval)	(0.186)	(0.192)	(0.175)	(0.188)

Year fixed effects not reported, Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Estimates of the static network collaboration models, research produced internally and 4 years of past collaborations

Tables 5 and 11 both consider the dependent variable as research produced internally by a university. The coefficient of the network effect is similar for the predetermined models in Table 5 and Table 11, indicating that the predetermined models are robust to a change in the time period considered. Finally, note that most of models' J-stats suggest the validity of the overidentification restrictions.

8 Monte Carlo Simulations

In order to evaluate the finite sample performance of the proposed estimator, I conduct a Monte Carlo simulation. The data generating process (DGP) is given by the reduced form equation,

$$y_t = (I - \phi G_t)(X_t \beta + \alpha + \epsilon_t) \tag{11}$$

where α is a $i.i.d.\, N(0,1)$ fixed effect for each university, $\beta_1 = 0.3$, $\beta_2 = -0.4$,$\beta_3 = 0.3$, $\phi = 0.00005$, and ϵ_t is generated from a normal distribution with mean equal to zero and variance equal to 8.

In order to test how the estimators behave under different specifications, I use two different networks: an exogenous valued network and a predetermined valued network. In each case, three explanatory exogenous variables are generated as correlated normals

$$x_i \sim N(0, V) \forall i \in \{1, 3\} \tag{12}$$

with the following covariance V

$$V = \begin{bmatrix} 1 & .83 & 0.95 \\ 83 & 1 & 0.85 \\ 0.95 & 0.85 & 1 \end{bmatrix} \tag{13}$$

The network links $g_{i,j,t}$ are generated following a zero inflated poisson process. The value of each link is given by :

$$g_{ij,t} \sim P(\mu) \tag{14}$$

where μ depends on the distance between the exogenous variables of the two partners, a shock parameter that depends on the time period, and an error term that is specific to each pair i, j.

$$\mu_t = \exp(\gamma_1 + \gamma_{21} * |x_{1t,i} - x_{1t,j}| + v_{t,i,j}) \tag{15}$$

where $\nu \sim U(0, 1)$. When the network is exogenous, the error terms ϵ_t and v_{t+1}'s are generated as $i.i.d.$ $N(0, 1)$. When the network is predetermined, the error terms ϵ_t and v_t's are generated as bivariate normal random variables such that $(\epsilon_{i,t}, v_{i,j,t+1}) \sim N(0, \Sigma)$ where Σ is defined as

$$\Sigma = \begin{bmatrix} 8 & \sigma_n \\ \sigma_n^{'} & I_{n-1} \end{bmatrix} \tag{16}$$

where σ is a $n-1$ vector of covariances $\sim N(0, 1)$ and I_{n-1} is a $(n-1 \times n-1)$ identity matrix. Table 12 indicates the point estimates, the variance and the mean square error (MSE) for the two valued network specifications for 500 runs.

		Mean point estimate	Variance	MSE
	ϕ = .00005	0.000045	0.000000	0.000000
Predet. network, treated as exog.	$\beta 1$ = .3	0.2826	0.1220	0.1220
	$\beta 2$ = -.4	-0.3642	0.1290	0.1300
	$\beta 3$ = .3	0.2728	0.1849	0.1852
	ϕ = .00005	0.000046	0.000000	0.000000
Predet. network, treated as predet.	$\beta 1$ = .3	0.2830	0.0840	0.0841
	$\beta 2$ = -.4	-0.3591	0.0801	0.0816
	$\beta 3$ = .3	0.2765	0.0837	0.0840
	ϕ = .00005	0.000047	0.000000	0.000000
Exog. network, treated as exog.	$\beta 1$ = .3	0.2838	0.0163	0.0165
	$\beta 2$ = -.4	-0.3862	0.0166	0.0168
	$\beta 3$ = .3	0.2860	0.0140	0.0142
	ϕ = .00005	0.000045	0.000000	0.000000
Exog. network, treated as predet.	$\beta 1$ = .3	0.2778	0.1013	0.1016
	$\beta 2$ = -.4	-0.3850	0.0712	0.0712
	$\beta 3$ = .3	0.2809	0.0924	0.0926

Table 12: Monte Carlo simulation of the estimators under different network specifications

The simulations' findings can be summarized as follow: In every case, the network parameter has little to no bias. In the case of the other variables, the estimates are biased upward.The bias is slightly higher when considering a predetermined network, especially for the negative coefficient.

9 Budgets cuts and optimal policy

As I explained in Section 3, post-2008 budget cuts were not anticipated by universities. As Spain now experiences the first effects of the funding cuts on research, this section focuses on demonstrating the importance of integrated cuts: by identifying the key players of the university system according to their ability to produce research, I am able to suggest targeting cuts to universities to minimize their total impact. Based on the estimates of the models, I use the formula of the key player as defined in the theoretical discussion to compute the inter-centrality for each university as described in the theoretical discussion.

Table 13 shows the 10 universities with the lowest intercentrality measure in 2008 (based on total research conducted), i.e. the non key players for 2008. For the full adjacency matrix, the lowest ten are fairly similar, with the exception of Alcala de

Henares and the Autonoma de Barcelona. Unsurprisingly, the smallest and more isolated universities are that that group of ten, including Leon, la Rioja, Huelva, Lleida, Oviedo and Burgos. The results for the column normalized adjacency matrix are found in the last two columns. As in the case of column normalization, the network is a linear combination of the resources the partners allocate to the common projects, it is unsurprising that middle range universities working with bigger partners or isolated universities are in this group. Again, the use of the predetermined or the exogenous network produce similar lowest ten

	Full adjacency
Static, exogenous	Static, predetermined
Leon	Las Palmas
La Rioja	Alcala de Henares
Las Palmas	Leon
Huelva	La Rioja
Burgos	Miguel Hernandez
Lleida	Pablo de Olavide
Oviedo	Burgos
Pablo de Olavide	Huelva
Autonoma de Barcelona	Alicante

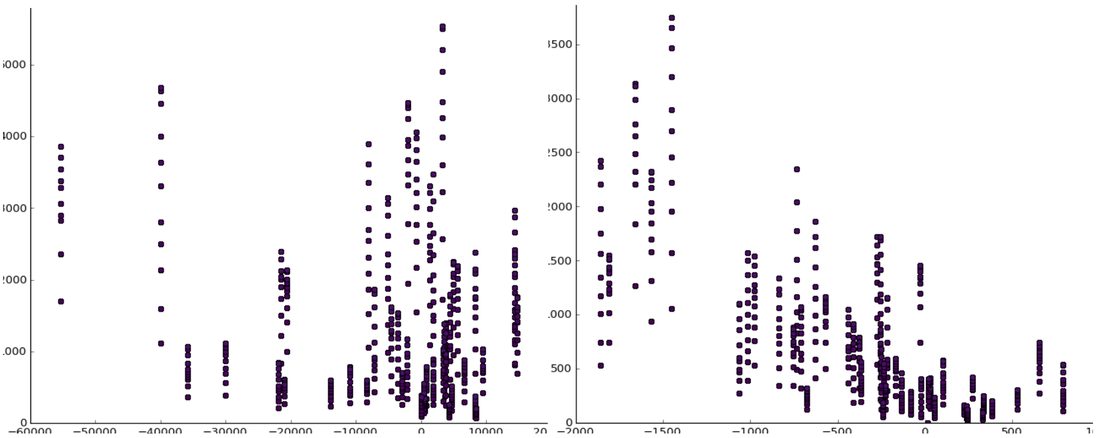
Table 13: 10 lowest intercentral universities, total research produced

Table 14 lists the ten universities with the lowest intercentrality measure in 2008 with the internal research as the dependent variables. The rankings are stable and fairly intuitive, and the universities with the lowest participation in total research are the same if we consider their internal research only or the internal and joint research projects. For the column normalized adjacency matrix, the lowest ten are fairly similar, with the exception of Cordoba, Almeria, Rey Juan Carlos and Pompeu Fabra. They are also the same as the lowest intercentral universities from the perspective of the total research produced, with the exception of Huelva. Again, I find that the smallest and more isolated universities, such as the universities de Leon, La Rioja, Huelva, Lleida, Oviedo and Burgos, are the less intercentral universities, meaning that the reduction in their budgets would lead to the lowest possible impact on the total of the universities research. The results for the full adjacency matrix are in found in the first two columns. They are consistent with the ones obtained when using the total research produced and should be understood in a similar fashion. Again, the use of the predetermined or the exogenous network produces a similar list, even if some changes appears in the ranks.[18]

[18]For an economist, the presence of the Pompeu Fabra is somehow surprising. However, the focus is here on life and physical sciences, in which the Pompeu Fabra is not specialized.

Static, exogenous	Full adjacency Static, predetermined
Leon	Alcala de Henares
Las Palmas	Las Palmas
La Rioja	Leon
Huelva	La Rioja
Burgos	Alicante
Lleida	Miguel Hernandez
Pablo de Olavide	Burgos
Oviedo	Pablo de Olavide
Autonoma de Barcelona	Huelva
Rovira I Virgili	Baleares

Table 14: 10 lowest intercentral universities, research produced internally

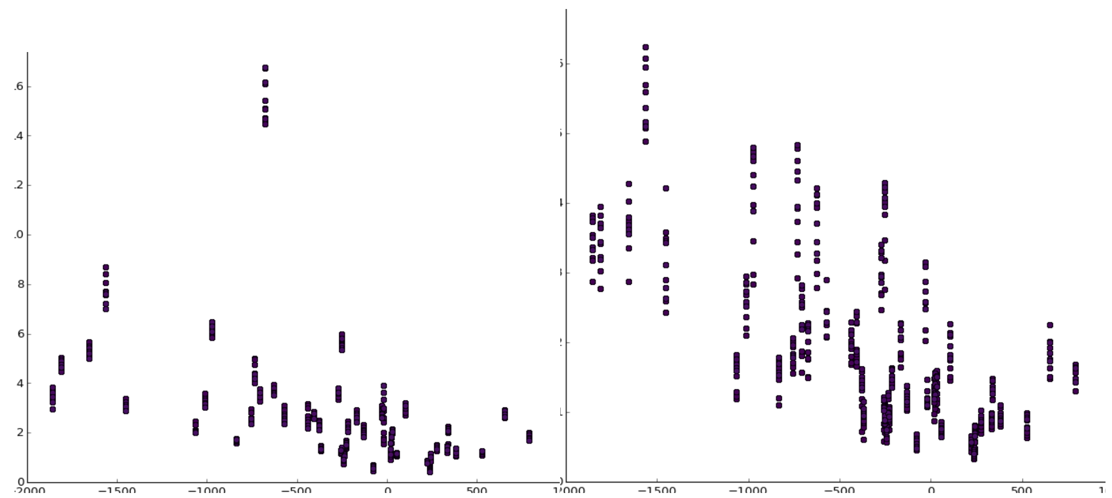


(a) Total research output

(b) Internally conducted research output

Figure 11: Correlation between the intercentrality measure and research output

Figure 12 shows the correlation between the intercentrality measure for the predetermined model with the full adjacency matrix and the research output. No clear pattern emerges for the link between the intercentrality parameter and total research output, as in Figure 11(b). For the internal research output only, there is a slight and negative correlation between the number of articles, proceedings and books and the intercentrality of a university, as in Figure 11(a). These correlations indicate that the intercentrality indeed takes into account both the direct productivity of a university, and its indirect effects on its neighbors.



(a) Correlation of the intercentrality measure and the number of students

(b) Correlation of the intercentrality measure and the budget

Figure 12: Correlation between the intercentrality measure and the number of students and budget

Figure 2.11 shows the correlation between the intercentrality measure for the predetermined model with the full adjacency matrix and various variables of the universities' production function. There is a negative correlation between the intercentrality measure and the total number of students (Figure 12(a)) and with the budget (Figure 2.11(b)).

10 Conclusion

This paper is one of the first attempts to use the literature on networked production to study the welfare implications of policies affecting the system as a whole. I show the importance of inter-university collaborations for the production of research (articles, books, or conference proceedings) and the importance of a systematic approach when designing higher education policies. I use an original dataset of publications and universities' characteristics to understand how the growth of the Spanish universities took place before the crisis, and how the crisis impacted the system of universities as a whole. While this type of literature usually focuses on the production function of each university as a single unit, this paper is an attempt at taking into account the direct and indirect impact of inter-university collaborations. When factoring in those dimensions, the local budget cuts may be compensated or have indirect effects. I suggest using a network model of production to understand how the universities use their resources to teach students and produce research.

I show how the budget cuts affected Spain (without taking into account the network aspect) and I suggested that analysis of the change in production attributable to budget-induced change in the system as whole could be based on the literature on network production. By using various network weights, I depart from the existing literature based on a binary matrix and show that, for most of the specifications of the network considered, the network effect is to be taken into account. This suggests an integrated approach to any public policy prescription which departs from the highly decentralized funding mechanism in Spain. Among the characteristics of the production function of the universities, the number of research and teaching staff at the university and the budget size are the main explanatory elements. When considering indirect and feedback effect, the budget of partners has a strong and positive effect on research produced.

Finally, I suggest an application the key player policy to minimize the impact of any funding cuts on the overall system, starting in 2008. While the intercentrality measure is sensitive to the model specification, it is still insightful as it shows the importance of some small universities for the research system as a whole. It also shows that the biggest universities are instrumental for the volume of research produced, but not necessarily the ones that, on a relative basis, are the more impactful for the total research produced. Those small but active universities, such as Valencia or Granada and Cordoba, are not to be forgotten in favor of helping the the biggest universities in Barcelona, the Complutense the Madrid or the polytechnic universities. Let us keep in mind that the models have been proved quite sensitive to the introduction of a dynamic element, reducing the impact of the network effects. In a situation with inertia, understanding the raise of the biggest universities is thus crucial, as research production displays inertia. However, such a dynamic decision process involves further network modeling that is beyond the scope of this paper.

The model leaves unexplained the competition between universities for resources, taking the funding rule as given. As the budgets are determined mainly by the number of students, the heterogeneity on size of the universities could be integrated in this framework, the competition for students balancing the partnerships for research. Such a model would merge the literature in economics of education and network models, and allow for a deeper understanding of the possible change of funding rules on the size of universities, but also on research.

Bibliography

Abramo, Giovanni, Tindaro Cicero, and Ciriaco Andrea D'Angelo, "Revisiting size effects in higher education research productivity," *Higher Education*, 2012, *63* (6), 701–717.

Adams, James D and Zvi Griliches, "Research productivity in a system of universities," *Annales d'Economie et de Statistique*, 1998, pp. 127–162.

_ , Grant C Black, J Roger Clemmons, and Paula E Stephan, "Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999," *Research policy*, 2005, *34* (3), 259–285.

Arellano, Manuel and Olympia Bover, "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 1995, *68* (1), 29–51.

Armenteros, J Hernández, "La Universidad española en cifras. 2006. observatorio Universitario," Technical Report 2006.

_ , "La universidad española en cifras: información académica, productiva y financiera de las universidades públicas españolas. Año 2008. Indicaores," Technical Report 2008.

Ballester, Coralio and Antoni Calvó-Armengol, "Interactions with hidden complementarities," *Regional Science and Urban Economics*, 2010, *40* (6), 397–406.

_ , _ , and Yves Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, sep 2006, *74* (5), 1403–1417.

Becker, WE, "Professorial behavior given a stochastic reward structure," *The American Economic Review*, 1979, *69* (5), 1010–1017.

Bonacich, Phillip, "Power and centrality: A family of measures," *American journal of sociology*, 1987, pp. 1170–1182.

Bramoullé, Yann, Rachel Kranton, and Martin D'amours, "Strategic interaction and networks," *The American Economic Review*, 2014, *104* (3), 898–930.

De Fraja, Gianni and Elisabetta Iossa, "Competition among universities and the emergence of the elite institution," *Bulletin of economic Research*, 2002, *54* (3), 275–293.

_ and Paola Valbonesi, "The design of the university system," *Journal of Public Economics*, 2012, *96* (3), 317–330.

Del Rey, Elena, "Teaching versus research: a model of state university competition," *Journal of Urban Economics*, 2001, *49* (2), 356–373.

Duch-Brown, Néstor, Martí Parellada-Sabata, and Jose Polo-Otero, "Economies of scale and scope of university research and technology transfer: a flexible multi-product approach," Technical Report 2010.

Dundar, Halil and Darrell R Lewis, "Departmental productivity in American universities: Economies of scale and scope," *Economics of Education Review*, 1995, *14* (2), 119–144.

fei Lee, Lung and Jihai Yu, "Efficient GMM estimation of spatial dynamic panel data models with fixed effects," *Journal of Econometrics*, 2014, *180* (2), 174–197.

García, JAP and JH Chica, "La universidad española en cifras: información académica, productiva y financiera de las universidades españolas (2008)," Technical Report 2010.

González, Luis Otero, David Rodeiro Pazos, Alfonso Rodríguez Sandiás, "La financiación de la investigación en las universidades públicas españolas:¿ una salida financiera?," in "El comportamiento de la empresa ante entornos dinámicos: XIX Congreso anual y XV Congreso Hispano Francés de AEDEM Asociación Española de Dirección y Economía de la Empresa" 2007.

Jackson, Matthew O and Yves Zenou, "Games on networks," *Handbook of game theory*, 2014, *4.*

Katz, Leo, "A new status index derived from sociometric analysis," *Psychometrika*, 1953, *18* (1), 39–43.

Kleibergen, F and R Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometricsconometrics*, 2006, *133* (1), 97–126.

König, Michael, Xiaodong Liu, and Yves Zenou, "R&D networks: Theory, empirics and policy implications," 2014.

Laband, David N and Bernard F Lentz, "New estimates of economies of scale and scope in higher education," *Southern Economic Journal*, 2003, pp. 172–183.

Lesage, JP, "An Introduction to Spatial Econometrics," *Revue d'économie industrielle*, 2008.

Liu, Xiaodong, Eleonora Patacchini, Yves Zenou, and Lung-Fei Lee, "Criminal networks: Who is the key player?," 2012.

Longlong, Hou, Li Fengliang, and Min Weifang, "Multi-product total cost functions for higher education: The case of Chinese research universities," *Economics of Education Review*, 2009, *28* (4), 505–511.

Mora, José-Gines, "The academic profession in Spain: Between the civil service and the market," *Higher Education*, 2001, *41* (1-2), 131–155.

Ortega, Eloísa and Juan Peñalosa, "The Spanish economic crisis: key factors and growth challenges in the euro area," *Banco de Espana Occasional Paper*, 2012, (1201).

Perez, Francisco and Lorenzo Serrano Martinez, "Universidad, universitarios y productividad en Espana," 2012.

Rothschild, M and LJ White, "The analytics of the pricing of higher education and other services in which the customers are inputs," *Journal of Political Economy*, 1995, pp. 573–586.

Worthington, Andrew C and Helen Higgs, "Economies of scale and scope in Australian higher education," *Higher Education*, 2011, *61* (4), 387–414.

Zenou, Yves, "Key Players," Technical Report, CEPR Discussion Papers 2014.

A Impact and moment of the crisis

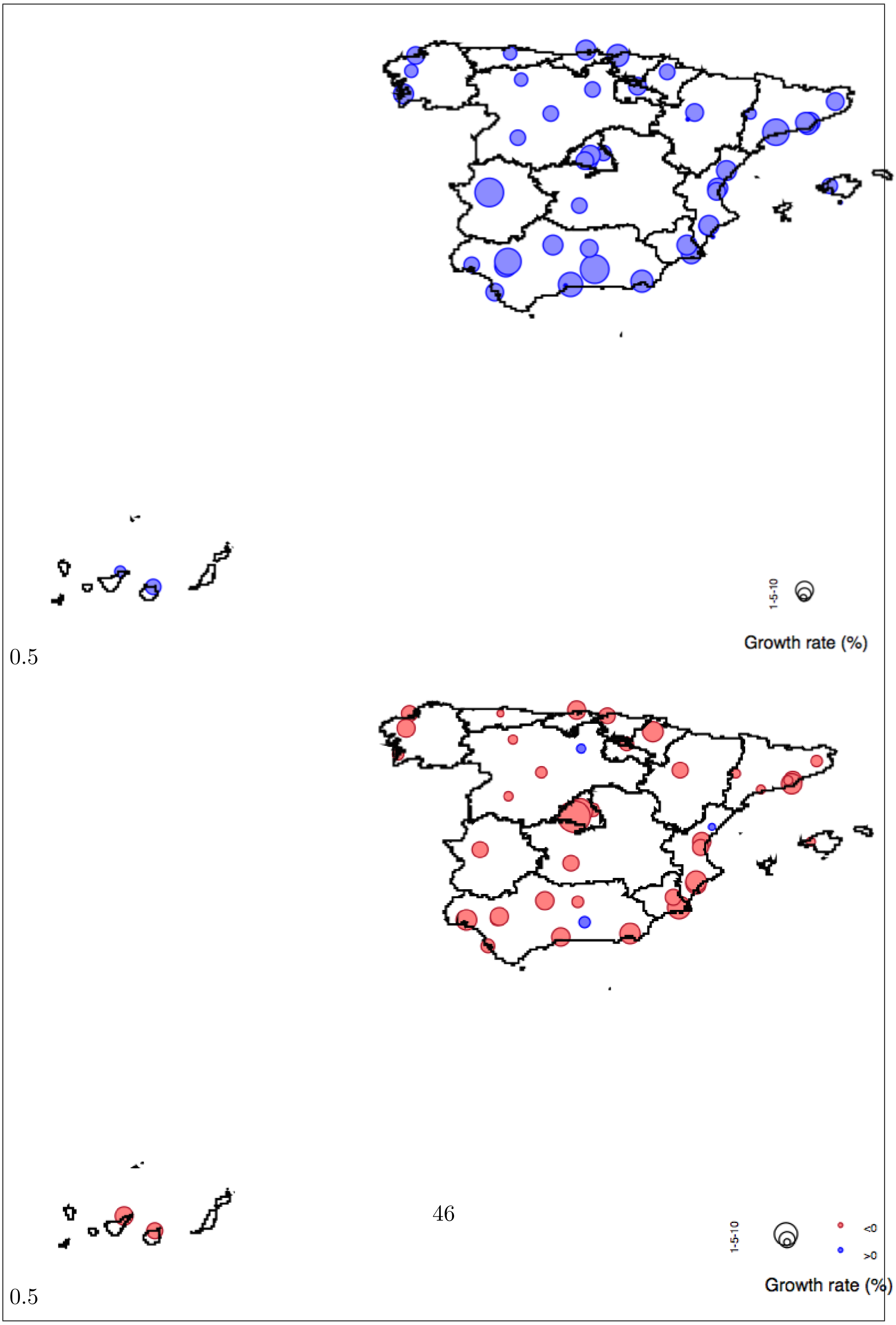


Figure 15: Universities' budgets growth rates

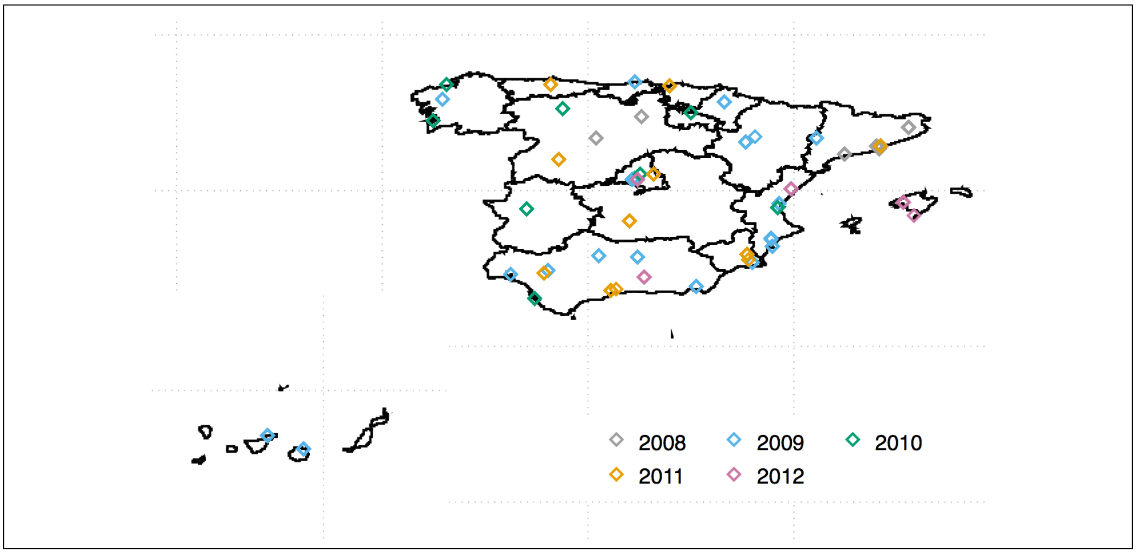


Figure 16: Year of first shock

B Patterns of collaboration

Region \ Partner	Public university	Regional agency	European agency	Federal agency	Private sector	Foundation	Private university	Other
Andalucía	68.1	1.7	0.1	18.2	0.4	0.4	2.0	9.1
Aragon	64.9	2.7	0	18.2	0.2	2.2	1.6	10.2
Asturias	68.3	1.3	0.1	17.0	0	0.3	0.2	12.8
Baleares	62.8	6.3	0	22.1	0	1.5	0.3	7.0
Canarias	51.1	0.7	0.8	35.9	0	1.9	0.7	9.0
Cantabria	67.9	0.5	0.2	20.3	0	0.5	1.2	9.5
Castilla-La Mancha	73.0	0.9	0.2	15.7	1.2	0.3	0.1	8.6
Castilla y León	66.0	1.0	0.2	18.3	0.1	4.3	0.3	9.8
Cataluña	52.8	11.9	0.2	19.5	1.5	1.3	1.0	11.9
Extremadura	68.5	1.9	0.1	16.5	0.4	1.9	0.6	10
Galicia	74.6	0.6	0.1	13.5	0.2	0.2	1.3	9.5
La Rioja	59.3	0.4	0	31.0	0	0.4	0.9	8.0
Madrid	57.4	1.8	0.8	27.3	0.5	0.8	1.4	10.1
Murcia	75.2	0.2	0.1	11.3	0.1	0.8	0.6	11.6
NP	64.5	0.1	0.3	27.9	0.1	0.4	0.1	6.5
Navarra	84.9	0	0	13.2	0	0	0	2.0
País Vasco	51.1	2.6	0	20.9	0.5	12.2	1.3	11.4
Valencia	67.8	1.9	0.6	17.1	0.1	0.9	1.3	10.4

Table 15: Scientific collaborations with each type of institutions, per region

Region \ Partner	Public university	Regional agency	European agency	Federal agency	Private sector	Foundation	Private university	Other
Andalucía	68.1	1.7	0.1	18.2	0.4	0.4	2.0	9.1
Aragon	64.9	2.7	0	18.2	0.2	2.2	1.6	10.2
Asturias	68.3	1.3	0.1	17.0	0	0.3	0.2	12.8
Baleares	62.8	6.3	0	22.1	0	1.5	0.3	7.0
Canarias	51.1	0.7	0.8	35.9	0	1.9	0.7	9.0
Cantabria	67.9	0.5	0.2	20.3	0	0.5	1.2	9.5
Castilla-La Mancha	73.0	0.9	0.2	15.7	1.2	0.3	0.1	8.6
Castilla y León	66.0	1.0	0.2	18.3	0.1	4.3	0.3	9.8
Cataluña	52.8	11.9	0.2	19.5	1.5	1.3	1.0	11.9
Extremadura	68.5	1.9	0.1	16.5	0.4	1.9	0.6	10
Galicia	74.6	0.6	0.1	13.5	0.2	0.2	1.3	9.5
La Rioja	59.3	0.4	0	31.0	0	0.4	0.9	8.0
Madrid	57.4	1.8	0.8	27.3	0.5	0.8	1.4	10.1
Murcia	75.2	0.2	0.1	11.3	0.1	0.8	0.6	11.6
NP	64.5	0.1	0.3	27.9	0.1	0.4	0.1	6.5
Navarra	84.9	0	0	13.2	0	0	0	2.0
País Vasco	51.1	2.6	0	20.9	0.5	12.2	1.3	11.4
Valencia	67.8	1.9	0.6	17.1	0.1	0.9	1.3	10.4

Table 16: Regional patterns of collaborations

C Budget crisis and research output: Some empirical evidence

	(1)	(2)
log(Budget)	167.65	166.13
		(168.21)
log(# Students)	170.36***	-761.88***
		(170.15)
log(# Staff)	236.10	-124.16
		(237.92)
Collaboration rate		1,053.85**
		(494.93)
T-7	64.74	69.36
	(166.49)	(165.85)
T-6	252.92	267.71*
	(161.76)	(161.28)
T-5	369.09	367.66**
	(160.46)**	(159.84)
T-4	510.99	506.55***
	(161.13)***	(160.52)
T-3	637.30	628.12***
	(167.45)***	(166.85)
T-2	727.77	713.15***
	(172.49)***	(171.96)
T-1	832.19	813.90***
	(182.08)***	(181.58)
Shock	962.38	948.27***
	(176.73)***	(176.17)
T+1	1,060.09	1,042.63***
	(176.13)***	(175.64)
T+2	1,178.10	1,157.39***
	(173.72)***	(173.32)
T+3	1,271.90	1,251.78***
	(173.38)***	(172.97)
T+4	1,310.75	1,295.58***
	(172.73)***	(172.21)
T+5	1,345.31	1,325.81***
	(174.76)***	(174.32)
ρ	0.95	0.95
R^2between	0.32	0.28
R^2overall	0.06	0.04

Table 17: Budget crisis and research output:Event study

	(1)	(2)
log(Budget)	169.63	151.70
		(169.59)
log(# Students)	175.14***	-825.79***
		(174.59)
log(# Staff)	239.03	-122.00
		(240.56)
Collaboration rate		1,141.44**
		(486.38)
2005	402.12	400.25***
	(66.78)***	(66.46)
2005	468.87	460.06***
	(76.19)***	(75.91)
2006	560.13	543.28***
	(87.17)***	(87.04)
2007	682.88	671.88***
	(88.39)***	(88.08)
2008	774.17	760.05***
	(90.20)***	(89.96)
2009	909.32	892.38***
	(85.45)***	(85.33)
2010	1,009.12	986.34***
	(77.71)***	(77.93)
2011	1,093.04	1,077.21***
	(72.79)***	(72.74)
2012	1,106.48	1,088.36***
	(71.69)***	(71.76)
2013	1,118.69	239.95
	(500.48)**	(623.09)
2014		
ρ	0.95	0.95
R^2between	0.42	0.37
R^2overall	0.08	0.07

Table 18: Budget crisis and research output: Panel approach

D First-stage regressions results

	Full adjacency		Column. norm. adjacency	
	Exog.	Predet.	Exog.	Predet.
log(# of students)	-108.634**	-15.91	-0.297**	-0.198
	(47.093)	(47.269)	(0.13)	(0.175)
log(# of staff)	103.220**	173.351***	0.138	0.184
	(50.572)	(62.622)	(0.143)	(0.183)
log(budget)	-55.849	-108.183	0.016	-0.105
	(41.792)	(64.575)	(0.133)	(0.163)
Network*log(# of students)	1.461***	0.656	1.201***	1.122***
	(0.318)	(0.486)	(0.229)	(0.24)
Network x log(# of staff)	-2.976***	-3.083***	-2.476***	-2.300***
	(0.131)	(0.157)	(0.097)	(0.09)
Network x log(budget)	2.146***	2.336***	2.760***	2.592***
	(0.224)	(0.398)	(0.204)	(0.198)
Constant	86.358***	-41.293*	-0.206***	-0.275***
	(18.838)	(22.062)	(0.043)	(0.057)
Observations	460	414	460	414
Partial R^2	0.9743	0.9858	0.9329	0.9155
F-stat (KP)	3855.54	2241.16	391.72	348.19

Year fixed effects not reported, Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: First-stage results, total research