Giulia Faggio, Olmo Silva, William C. Strange

Heterogeneous agglomeration

Article (Accepted version)
(Refereed)

Original citation:
Faggio, Giulia, Silva, Olmo and Strange, William C. (2016) *Heterogeneous agglomeration.* Review of Economics and Statistics. ISSN 0034-6535

DOI: 10.1162/REST_a_00604

Heterogeneous Agglomeration*

Giulia Faggio
Westminster Business School,
University of Westminster,
35 Marylebone Road, London, NW1 5LS
and
Spatial Economics Research Centre,
London School of Economics,
Houghton Street, London, WC2A 2AE, UK
Tel.: +44 (0)20 7852 3592
G.Faggio1@lse.ac.uk

Olmo Silva
Department of Geography and Environment,
and
Spatial Economics Research Centre,
London School of Economics,
Houghton Street, London, WC2A 2AE, UK
Tel.: +44 (0)20 7955 6036
O.Silva@lse.ac.uk

William C. Strange
Rotman School of Management,
University of Toronto,
105 St. George St.,
Toronto, ON M5S 3E6, Canada
Tel.: +1 (416)978-1949
wstrange@rotman.utoronto.ca

October 19, 2015

*We thank the Economics and Social Research Council (ESRC) UK, the Social Sciences and Humanities Research Council of Canada, and STICERD-LSE for financial support. We also thank Simona Iammarino, William Kerr and Gianluca Tarasconi for sharing data. Finally, we are grateful for the helpful comments we received from three anonymous referees, Gordon Hanson, Steve Gibbons, Stephan Heblich, Antonio Miscio, Henry Overman, Michael Storper, Jacques Thisse, and seminar participants at the AREUEA National Conference, the CESifo Conference on the Economics of Entrepreneurship, the IEB Workshop on Urban Economics, LSE, Manchester University, the NARSC Annual Meetings, Syracuse University, Tinbergen Institute, and Wilfrid Laurier University. We are responsible for any errors or omissions.

Heterogeneous Agglomeration

Abstract

Many prior treatments of agglomeration explicitly or implicitly assume that all industries agglomerate for the same reasons, with the traditional Marshallian (1890) factors of input sharing, labor pooling, and knowledge spillovers affecting all industries similarly. An important instance of this approach is the extrapolation from one key sector to the larger economy, such as the drawing of very general lessons about agglomeration from the specific case of the Silicon Valley. Another is the pooling of data to examine common tendencies in agglomeration even across very different industries. This paper uses UK establishment-level data on coagglomeration to document substantial heterogeneity across industries in the microfoundations of agglomeration economies. The analysis shows that the Marshallian factors interact with the organizational and adaptive aspects of agglomeration discussed by Chinitz (1961), Vernon (1960), and Jacobs (1969). Our findings highlight the importance of treating Marshall's microfoundations of agglomeration as complements to the analysis of Jacobs and others, rather than as alternatives.

I. Introduction

This paper considers heterogeneity across industries in the microfoundations of agglomeration economies. Marshall (1890) notes the existence of three sources of agglomeration economies: labor pooling, input sharing, and knowledge spillovers. Many subsequent treatments of agglomeration either explicitly or implicitly suppose that all industries agglomerate for the same reasons, with the three Marshallian forces affecting all industries similarly. An important instance of this approach is the extrapolation of individual cases to the larger economy, such as the drawing of very general lessons about agglomeration from the specific case of the Silicon Valley. Another is the pooling of data to examine common tendencies in agglomeration even across industries that theory suggests would agglomerate differently. This paper documents the existence of significant heterogeneity and shows that the pattern of heterogeneity has important implications for our understanding of the nature of agglomeration economies.

The paper's empirical analysis focuses on the relationship between the coagglomeration of industry pairs and Marshallian links between industries. The motivation for this approach is that the variation in the characteristics of industries that co-locate sheds light on the microfoundations of agglomeration economies. Ellison et al. (2010), who developed this approach, show that proxies for labor pooling, input sharing, and knowledge spillovers between an industry pair are positively and significantly related to co-location. Our paper explores how these results vary across industries, guided by classic non-Marshallian analyses of agglomeration. Jacobs (1969) stresses the unplanned nature of the creation of new work in cities, while Vernon (1960) discusses how cities help manage the instability involved in certain production processes. Chinitz (1961) argues for a positive role of small firms in the generation of agglomeration economies – another example of this non-Marshallian research. Similarly, Porter's (1990) influential analysis of industry clusters identifies a positive role for competition. The pattern of heterogeneity that we document is consistent with these non-Marshallian microfoundations that focus on adaptive and organizational factors.[1]

Our analysis makes use of establishment-level data from the UK's Business Structure Database (BSD) covering the years 1997-2008. We initially estimate benchmark models of the relationship between measures of industry links and coagglomeration across all manufacturing industries, as in Ellison et al. (2010). We then consider heterogeneity in ways suggested by Jacobs' and Vernon's notions of adaptation and Chinitz's organizational approach, as well as by more recent research on the role of human capital in the agglomeration process (e.g., Rauch, 1993, Glaeser and Saiz, 2004, Moretti, 2004, and Berry and Glaeser, 2005). Using coagglomeration to look at these aspects of agglomeration is unique in the literature.

[1]The modern theoretical literature offers various formal results that are consistent with the informal treatments in the classics. Duranton and Puga (2001) model cities as "nurseries" for infant industries, while Strange et al. (2006) establish the attraction of cities for businesses facing uncertainty. Helsley and Strange (2001) is another treatment of adaptive and organizational agglomeration economies. See Duranton and Puga (2004) for a survey of this theoretical literature.

Furthermore, we examine the interaction between Marshallian forces and other elements of agglomeration rather than looking at Marshall as a rival to Jacobs and others – which is also unique and in contrast with the previous literature. In this sense, the paper is an attempt to create a détente between Marshall and Jacobs.

The empirical analysis leaves no doubt that agglomeration works differently for different industries. The key empirical results are as follows. First, in a great variety of coagglomeration models, we show the robust predictive power of Marshall's agglomeration forces. This confirms prior work and supports our focus on interactions between Marshallian and non-Marshallian approaches. Second, a quantile regression that differentiates pairs by their tendency to coagglomerate provides results that are consistent with Jacobs' analysis of unplanned knowledge spillovers and labor pooling. Third, differencing by entry and industry age provides robust evidence of an adaptive element to agglomeration, consistent with Vernon. This manifests itself more strongly in labor pooling and in knowledge spillovers than in input sharing. Fourth, differentiation by the sector's technology orientation and workforce education shows that agglomeration is not just a high-technology phenomenon. However, high-technology sectors show stronger evidence of knowledge spillovers, while low-technology industries show stronger evidence of input sharing and labor pooling. These findings are broadly consistent with learning playing an important role in the agglomeration process as suggested by Jacobs and Vernon. Finally, agglomeration effects – in particular those related to input sharing – tend to be stronger when firms are smaller, consistent with Chinitz.

In addition to building on the classics in the agglomeration literature, the paper also builds on more recent econometric work on agglomeration.[2] The line of research closest to this paper examines the relative importance of Marshallian forces using what might be called "horse race" models. For example, Audretsch and Feldman (1996) and Rosenthal and Strange (2001) regress levels of agglomeration on proxies for the presence of labor pooling, input sharing, and knowledge spillovers. Another recent approach is Jofre-Monseny et al. (2011), who estimate count models of new firms as functions of proxies for Marshallian forces. A related body of work consists of papers that separately consider Marshall's three forces. See, among others, Fallick et al. (2006), Almazan et al. (2007) and Serafinelli (2015) on job hopping, Holmes (1999) on input sharing, and Jaffe et al. (1993), Arzaghi and Henderson (2008) and Lin (2012) on patents, networking and learning, and the creation of new work. This body of work presents persuasive evidence that the three Marshallian forces are present. Our paper provides further such evidence and extends this line of research by incorporating theories of organization and adaptation.

While the agglomeration literature has much to say about how agglomeration economies are generated, it has less to say about heterogeneity in microfoundations. Henderson et al. (1995) show that agglomeration economies differ between high- and low-technology industries in an analysis of urban growth.

[2]See Hanson (2001), Rosenthal and Strange (2004), Behrens and Robert-Nicoud (2014), and Combes and Gobillon (2014) for reviews of the agglomeration literature.

However, they consider whether agglomeration economies arise from own-industry activity or from urban diversity, rather than directly considering Marshall's three forces. Together with Glaeser et al. (1992), this paper has spawned a literature that contrasts Marshall vs. Jacobs – rather than studying the interactions between Marshallian and non-Marshallian forces in characterising heterogeneous agglomeration, as we do. More recently, Hanlon and Miscio (2014) estimate a dynamic industry growth model, and establish the importance of input-output linkages and labor pooling. Their results show that smaller firms both benefit from and produce stronger agglomeration effects. Glaeser and Kerr (2009) and Rosenthal and Strange (2010) also consider the idea that agglomeration economies are stronger when there are many small firms. In these papers, agglomeration is organizational. Duranton and Puga (2001), although largely a theoretical exercise, present empirical evidence on location decisions over an industry's life cycle that is consistent with a model of cities as nurseries that tend to young industries. Strange et al. (2006) show a systematic tendency for industries facing more uncertainty in Marshallian dimensions to agglomerate. In both of these papers, agglomeration is fundamentally adaptive. As a group, this literature suggests that there is reason to believe that agglomeration economies are heterogeneous. Our analysis systematically documents the pattern of this heterogeneity and what this implies for our understanding of agglomeration economies.

Taken as a whole, our results on heterogeneity argue for caution in extrapolation from individual cases of agglomeration. This is important because extrapolation from cases is a central part of the justification for cluster policy.[3] Unfortunately, as satisfying as it is to draw conclusions from interesting and highly salient examples of agglomeration such as the Silicon Valley and computers or Detroit and cars, our findings show clearly that different industries respond differently to agglomerative forces. Similarly, our results suggest that one should interpret horse race models on the relative strength of agglomeration effects with care since these specifications do not allow for heterogeneous effects across industries. All of this is consistent with the advice offered by the cluster policy review paper by Chatterji et al. (2013). Policymakers should recognize that agglomeration issues are complex, and there is much to recommend caution in cluster policies. Careful pilot projects have the potential to uncover what works and what does not for particular industries. Policies that are consistent with growth in general are likely to help clusters emerge. Conversely, policies targeting specific industries run the risk of picking losers rather than winners, given the uncertainties associated with heterogeneity in agglomeration economies.

The remainder of the paper is organized as follows. Section II discusses our empirical approaches. Section III presents the baseline Marshallian analysis. Section IV considers Jacobs' unplanned interactions, while Section V presents the analysis where industries are heterogeneous. Section VI concludes.

[3]See, for instance, Porter (1990) and the critique in Duranton (2011).

II. Coagglomeration and agglomeration forces

A. Measuring coagglomeration

Our analysis of microfoundations is based on the tendency of industries to co-locate across metropolitan areas. We use the Ellison and Glaeser (1997) measure of coagglomeration, which is standard in the field. Let N_i denote total employment in industry *i*, and n_{mi} denote employment in metropolitan area *m* and industry *i*. Let $s_{mi} = n_{mi}/N_i$ denote the share of a given industry *i*'s employment in metropolitan area *m*, and let x_m denote the metropolitan area's share of national employment. For two industries *i* and *j*, the Ellison-Glaeser measure of coagglomeration can be written as (Ellison et al., 2010):

$$\gamma^C_{ij} = \frac{\sum_{m=1}^{M}(s_{mi} - x_m)(s_{mj} - x_m)}{1-\sum_{m=1}^{M}(x_m)^2}\,. \tag{1}$$

This measure is related to the covariance of industries across metropolitan areas.

We construct measures of coagglomeration of UK manufacturing industries using data from the Business Structure Database (BSD) for the period 1997 to 2008. The data come from administrative records covering 99 per cent of economic activity in the UK. We use BSD data at the local unit (i.e., plant or establishment) level including both single- and multi-plant enterprises. For each local unit, information is available on employment, industrial activity, year of birth (start-up date) and death (termination date), and postcodes. We use this detail to assign each local unit to a Travel-to-Work Area (TTWA, see below). The raw data include approximately three million local units every year. After a series of data cleaning procedures, our dataset comprises more than two million plants annually over 12 years.[4]

To quantify coagglomeration we focus on three-digit industries of the UK Standard Industry Classification (SIC) 1992 and restrict our attention to manufacturing (SIC151-SIC372).[5] In line with the literature, we do not consider other sectors (such as services) because measuring the extent of labor pooling, input sharing and, especially, knowledge spillovers in those industries is challenging. After excluding and recombining sectors that present a limited or erratic evolution in the number of plants and/or their employment, we are left with a final sample of 94 manufacturing three-digit industries. This gives a total of 4,371 unique pairs a year over twelve years (1997-2008) for an overall count of 52,452 observations.

The level of geographical aggregation we use is the Travel-to-Work Area. TTWAs are geographical entities defined so that at least 75% of the resident population works in the area and 75% of the people working in the area reside there. TTWAs were devised to delineate areas that are as self-contained labor markets and economically relevant aggregates. As of 2007, there were 243 TTWAs in the United Kingdom.

[4]We use data from England, Scotland and Wales but drop Northern Ireland because of poor data coverage. See the Web Appendix for further detail (available at http://personal.lse.ac.uk/silvao/HA_WebAppendix.pdf).
[5]The UK SIC is a system for classifying industries by a five-digit code similar to the US SIC that was used prior to the introduction of the six-digit North American Industry Classification System (NAICS code) in 1997.

In our analysis, we focus on 84 urban TTWAs with population in excess of 100,000 residents. In some extensions, we also consider rural TTWAs and use other levels of aggregation such as regions.

To measure coagglomeration, we compute Ellison et al.'s (2010) γ^C measure based on the total employment shares of the selected 94 three-digit industries contained in the 84 urban TTWAs. Descriptive statistics are presented in Table 1. The mean and median of γ^C are centered at zero, with a standard deviation of 0.005, a minimum of -0.028, and a maximum of 0.107. Relative to Ellison et al. (2010), UK coagglomeration displays less dispersion, although it is similarly skewed towards positive values.

Table A1 in the Appendix lists the fifteen most coagglomerated sectors alongside the three top TTWAs where most of the coagglomeration takes place. Textiles, products of clay and ceramic, and basic metals are the most recurrent sector groupings. Some geographic patterns emerge. Textile-related industries tend to co-locate in Bradford, Leicester, Manchester and Nottingham. Basic-metal activities cluster around Birmingham and Sheffield. Clay- and ceramic-product manufacturers group around Stoke-on-Trent. The publishing and printing sectors tend to locate in London.

In some extensions to our core analysis, we use variants of γ^C. In particular, we calculate: (i) a measure of coagglomeration constructed using the number of plants rather than their employment; (ii) a version of γ^C that excludes London; (iii) a measure that includes single-plant companies only; (iv) a version that includes both urban and rural areas; and (v) a measure that excludes publishing (SIC221) and printing and reproduction of media (SIC222). Descriptive statistics of these alternative measures are very similar to those presented in Table 1. Furthermore, their correlation with our main measure is always high, between 0.76 (when only including single-plants firms) and 0.99 (when considering both urban and rural areas).

B. Marshallian agglomeration forces

Marshall attributed the spatial concentration of industry to three forces: labor pooling, input sharing, and knowledge spillovers. In this section, we discuss the variables we use in order to measure the flow of goods, people and ideas across industrial pairs.[6] Our proxies are deliberately very similar to those used in Ellison et al. (2010), which we consider to be 'best practice' given available data.[7]

To assess the potential for labor pooling, we use UK Labour Force Survey (LFS) data between 1995 and 1999. The LFS is a representative survey of households living in the UK. The data report a worker's industry and Standard Occupation Classification (SOC) 1990. The UK SOC categorizes occupations on the basis of skill level and content. We use the 331 occupation groups defined by the three-digit SOC

[6]More details are provided in the Web Appendix.
[7]The only other proxy that might be relevant is Fallick et al's (2006) measure of job hopping as an alternative to the occupational similarity proxy for labor pooling employed by Ellison et al (2010). This, however, cannot be computed with the data currently available to us.

classification in conjunction with the 94 three-digit manufacturing industries to calculate $Share_{io}$ and $Share_{jo}$. These measure the shares of employees of occupation *o* in industry *i* and *j*, respectively. Using this information, we measure the similarity of employment in industries *i* and *j* by computing the correlation between $Share_{io}$ and $Share_{jo}$. Descriptive statistics are presented in Table 1. The mean value is 0.237 with a standard deviation of 0.188.

To assess input sharing, we use the ONS Input-Output Analytical Tables for 1995 to 1999.[8] We calculate the shares of inputs that each industry within a pair buys from the other as fractions of their total intermediate inputs and the shares of outputs that they sell to each other as fractions of their total output, excluding direct sales to consumers. We then construct three different proxies for input-output linkages. First, we consider the maximum between the share of inputs that sector *i* is buying from sector *j*, and vice versa. Next, we recover the maximum between the share of output that sectors *i* is selling to sector *j*, and vice versa. These capture upstream and downstream linkages, respectively. Finally, we consider the maximum of these two proxies as a synthetic measure of the linkages between pairs. The mean for all three proxies is close to zero (see Table 1), suggesting that most industries do not share intermediate goods to an important degree. In fact, 30% of the sector pairs do not share any input or output, while 75% of the pairs share less than 0.005.

To construct a proxy for knowledge spillovers, we track patent citation flows using information on UK inventors contained in the European Patent Office (EPO) data for the years 1997 to 2009. Approximately 144,000 patents were filed by UK inventors over this period generating more than 77,000 citations. Using this information, we measure the extent to which patents associated with industry *i* cite patents associated with industry *j* and vice-versa. The main difficulty lies with creating a mapping between sectors and patents – which are categorized using technological classes rather than a standard industrial classification. Following the literature, we adopt two approaches and use: *(1)* a probabilistic mapping based on the Industry of Manufacture (IOM); and *(2)* an alternative probabilistic mapping based on the Sector of Use (SOU). After applying these procedures, we investigate the number of citations that a patent in sector *i* is receiving from patents in sector *j*, and the number of patents in sector *j* that a patent in sector *i* is citing. These measures are analogous to the input sharing proxies described above. Our two indicators consider the maximum patent-citation flow between sector *i* and sector *j* – normalized by total citations in that industry – using either the IOM or the SOU probabilistic mapping. Descriptive statistics in Table 1 show that the average knowledge spillover shares are 0.012 (SOU) and 0.016 (IOM). Both distributions are highly skewed with median values in the order of 0.003/0.004, and 75% of the industries having citation flows below 0.011/0.013.

In addition to Marshallian agglomeration forces, we also control for access to resources and

[8]We refer to this as "input sharing" in line with prior usage even though there are both upstream and downstream elements to our measure, as in Krugman (1991).

infrastructure that might impact location choices. Using the ONS 1995-1999 I-O Tables, we gather information on industries' use of primary resources and other non-manufactured inputs in order to quantify industry-pair similarity in these respects.[9] Specifically, we build a measure of the share of inputs that an industry is purchasing from the seven I-O primary "natural resource" industries (including agriculture, forestry and fishing, and mining and quarrying). We also control for usage of water and energy by separately considering the share of inputs bought from water-related service companies, and from energy-related industries (both electricity and gas). Further, we consider the share of inputs bought from transport-related sectors (including railways, air, water and other land transport) to control for the importance of transport costs. Finally, following Overman and Puga's (2010) analysis of labor pooling, we create a proxy for access to business services by considering the share of inputs bought from this sector.[10] We construct our proxies for the dissimilarity of industry pairs by measuring (one half of) the absolute value of the difference in the shares of these various inputs used by the pair. Descriptive statistics are presented at the bottom of Table 1.

C. Beyond Marshall: adaptive and organizational aspects of agglomeration

As discussed in the Introduction, there has been considerable empirical research on Marshall's three forces. The literature on non-Marshallian aspects of agglomeration is much less developed. Within this literature, the approach that has received the most attention is Jacobs (1969), who focuses on adaptation, specifically on the unplanned nature of "the creation of new work." It is common to treat Jacobs as proposing an alternative to Marshall, as in the Glaeser et al. (1992) and Henderson et al. (1995) papers on urban growth. There is a natural sense in which this is true. Marshall sees increasing-returns forces as promoting the spatial concentration of industry. Jacobs, in contrast, focuses primarily on knowledge spillovers, and sees the creation of new work as being enhanced by local diversity. There is another sense, however, in which Jacobs and Marshall ought not to be presented as polar opposites. Jacobs' analysis of knowledge is certainly in the spirit of Marshall, and she clearly mentions the labor and input market aspects of the creation of new work.

In this spirit, this paper examines complementarities between Marshall and Jacobs and offers a novel approach to investigating these issues. Previous work has studied the 'Marshall vs. Jacobs' dichotomy by regressing measures of local productivity, growth, or wages on measures of local specialization or diversity – typically proxied by a Herfindahl index of industrial concentration. While the specialization measure is

[9]Ellison et al. (2010) address this issue by using the US spatial distribution of natural resources, transport costs and labor inputs to predict coagglomeration stemming only from differences in resource costs. This approach cannot be replicated in the UK because the geographical scale of the country makes the spatial distribution of resources and "natural infrastructure" – e.g. access to the sea – much more homogeneous, and because differences in the cost of resources – such as gas, oil, water and electricity – are negligible due to regulatory constraints.

[10]This group includes, among others, computer services, R&D activities, legal consulting, accounting services, market research and management consulting, and advertising. Results do not depend on the inclusion of this variable.

tightly tied to Marshall's ideas, the diversity variable is only loosely linked to Jacobs. Her intuition is that a diverse city offers opportunities for unplanned, unpredictable, or otherwise unusual interactions between different industries, leading to increased creation of new work arising from these unexpected connections.

The idea behind our approach is to focus on differences between industry pairs that agglomerate frequently and those that do not. While there are many factors that determine whether agglomeration is more or less common, the unplanned, unpredictable, or unusual interactions at the heart of Jacobs's analysis are more likely to be found among industry pairs that are infrequently co-located. Conversely, planned or otherwise predictable interactions that arise from strategic migration decisions and entrepreneurial survival are likely to be found among industry pairs that co-locate frequently. Following this logic, we investigate heterogeneity in the response to Marshallian forces between more- and less-coagglomerated industry pairs using quantile regressions to identify Jacobs-type agglomeration economies.

Jacobs is not Marshall's only important successor in the study of agglomeration. In Chinitz (1961), New York differs from Pittsburgh because its industry is organized in a less-concentrated fashion, making it a friendlier environment for startups and innovation. Porter (1990) similarly argues that competition is healthy for a business cluster. Vernon (1960) writes about the importance of "instability" for increasing-returns industries, arguing that newer industries with more entry are the ones that benefit more from locating in a large city. Others have emphasized the importance of human capital (e.g., Rauch, 1993, and Glaeser and Saiz, 2004, Moretti, 2004) and creativity (Florida, 2003), both of which are related to a city's adaptive capacity. As with Jacobs, we believe these approaches to agglomeration should be seen as complements to Marshall rather than as substitutes or alternative explanations. This intuition informs our empirical work.

In order to test these non-Marshallian mechanisms, we examine the pattern of heterogeneity in Marshallian agglomeration effects using a sectoral breakdown that captures the various non-Marshallian approaches. To begin, we use information collected by the OECD in 1997 to classify sectors as 'high-' or 'low-technology.' Next, we gather data on the share of college graduates in each industry using the LFS and classify sectors as 'high-' or 'low-education' according to whether this share is above or below the median (at 0.078). Finally, we use information gathered from within the BSD to split our sample along the following dimensions: *(a)* sectors where the first year of opening of currently operating plants is above or below the median across all years and sectors (at 1967). These industries are labelled 'new' and 'old', respectively;[11] *(b)* sectors where the share of entrants – i.e. the incidence of new firms at time *t* in the total number of firms in that year – is above or below the median across all years and industries (at 0.10). These are labelled 'dynamic' and 'steady' sectors; *(c)* sectors where the average size of the entrants – i.e. firms operating at

[11]We rank industries by the age of the oldest *currently operating* plant, not the age of the industry itself. We believe that this captures the degree to which an industry's operations are settled. However, given that our data refers to UK manufacturing from 1997 to 2008, it is hard to detect the births of new industries.

time *t* that did not exist at time *t-1* – is above or below the median size across all years and sectors (at 8.59). We label these as 'large entrant' and 'small entrant' sectors; and *(d)* sectors where the average size of the incumbents – i.e. firms operating both at time *t* and *t-1* – is above or below the median size across all years and sectors (at 18.95). These are labelled 'large incumbent' and 'small incumbent' sectors, respectively.

Given that the level of observation in our dataset is the industry pair, we use this information to classify combinations where both sectors belong to one group (e.g., both high-technology or both low-technology) and mixed pairs where the two sectors belong to different groups (e.g., one high- and one low-technology). More details about the construction of these groupings, number of observations in each block, and further descriptive statistics are presented in Table A2 in the Appendix.

We then study the pattern of heterogeneity in the intensity of the Marshallian forces across these groups to shed light on non-Marshallian approaches to agglomeration. In particular, we analyze heterogeneity along the 'new' vs. 'old' and 'dynamic' vs. 'steady' dimension to provide evidence about adaptive aspects of agglomeration, as in Vernon and Jacobs. We focus on the 'high tech' vs. 'low tech' and 'high education' vs. 'low education' spectrum to quantify the importance of related ideas about human capital and adaptive capacity. Finally, we study heterogeneity along the dimensions of entrants' and incumbent's size to shed light on the organizational aspects of agglomeration, as in Chinitz.

III. Coagglomeration and Marshallian microfoundations: UK Evidence

A. Univariate and multivariate OLS regression analysis

In this section we study the microfoundations of agglomeration economies by linking the proxies for the three Marshallian forces discussed above to industry-pair coagglomeration. Our results come from regressions of the following kind:

$$\gamma^{C}_{ijt} = \alpha + \beta_{LP}\, LP_{ij} + \beta_{IO}\, IO_{ij} + \beta_{KS}\, KS_{ij} + \Sigma_{k=1}^{5}\, \lambda_k\, Diss_{ij}^{k} + \varepsilon_{ijt}, \tag{2}$$

where γ^{C}_{ijt} is the Ellison et al. (2010) measure of coagglomeration between sectors *i* and *j* at time *t*; LP_{ij}, IO_{ij} and KS_{ij} are proxies for labor pooling (LP), input sharing (IO) and knowledge spillovers (KS) between sectors *i* and *j* averaged over the relevant years (see Section II.C for details); and $Diss_{ij}^{k}$ is one of the five measures of dissimilarity between sectors *i* and *j* in terms of use of primary resources and non-manufacturing inputs. Finally, ε_{ijt} is an error term uncorrelated with all other variables. We allow for an arbitrary degree of correlation in the shocks of sector pairs over the years and cluster standard errors at this level. The dataset consists of 4,371 unique combinations of 94 manufacturing sectors over 12 years, giving a total of 52,452 observations.[12] Throughout the analysis, we standardize our variables to have unitary standard deviation.

As Ellison et al. (2010) note, the motivation for this approach is that the characteristics of industries that frequently co-locate can shed light on the microfoundations of agglomeration economies. For instance, if industries that frequently buy from and sell to each other coagglomerate to a large degree, this suggests that input sharing is an important agglomeration force. This in turn requires that coagglomeration is related to the strength of the agglomeration economies operating within the industry pair.

In addition to being intuitively appealing, Ellison et al. (2010, Mathematical Appendix) prove this property formally in the context of a specific model of agglomeration with industries partitioned into groups that must co-locate in order to have positive profit. With sequential location choices, in this all-or-nothing agglomeration model industries that benefit from coagglomeration will coagglomerate. The authors note that it is likely that this result would hold in weaker form with somewhat weaker agglomeration economies. It is worth noting, however, that there is a fundamental coordination problem in the determination of city composition (Helsley and Strange, 2014), and it is possible that coagglomeration fails to occur even when it would be mutually beneficial or that coagglomeration does occur when it is not. Nonetheless, there are good reasons to believe that equilibrium coagglomeration does increase when the strength of the agglomeration effect is stronger. First, there is a robust empirical relationship between proxies for agglomeration forces within an industry pair and equilibrium coagglomeration. In addition to Ellison et al. (2010), a number of papers (Kolko, 2010; Jacobs et al., 2013; and Gabe and Abel, 2013) find evidence of this sort. Thus, it seems that the selection among the multiple equilibria noted by Helsley-Strange is skewed in favor of a positive relationship between the benefits of coagglomeration and the coagglomeration that occurs in equilibrium. Second, O'Sullivan and Strange (2015) use an agent-based model to select from multiple-equilibrium city compositions. They also show a positive relationship between the strength of the spillovers within an industry pair and equilibrium coagglomeration.

Focussing on coagglomeration, rather than on the cross-sectional pattern of industry clustering (as in Audretsch-Feldman, 1996, and Rosenthal and Strange, 2001), has additional advantages. First, this approach looks directly at links between industry pairs and thus sheds light on the mechanisms of agglomeration in a way that looking at the concentration of industries cannot. Second, studying the links between coagglomeration and pair-wise Marshallian forces helps dealing with unobservables that could bias the results when the unit of observation is the industry – but are less likely to be important when the analysis is

[12] Our proxies for the Marshallian forces are measured at the beginning of the observation window and have no time variation. If we collapse γ^{C}_{ijt} to its average across all years and run regressions that exploit variation over 4,371 observations only, we find identical results to those reported here. This is expected as the two approaches produce identical point estimates and significance levels with clustered standard errors. The reason why we keep the dataset at the year × industry-pair level is that in some robustness checks we stagger and modify our observation window.

carried out at the industry-pair level.[13] Of course, the emphasis we put on these advantages does not imply that we consider coagglomeration as the only valid approach to studying the microfoundations of agglomeration economies. We simply argue that it is a valid approach and one we can flexibly use to study heterogeneous patterns by neatly characterising the nature of industry pairs.

The first set of results is presented in Table 2. Columns (1) and (2) tabulate results from univariate regressions where we consider only one Marshallian force at the time (and include dissimilarity controls in Column 2). The results show that labor pooling has the largest and most significant association with coagglomeration. A one standard deviation increase corresponds to 19% of a standard deviation increase in γ^C. For input sharing and knowledge spillovers, the corresponding increases are 14% and 10%, respectively. This pattern is consistent with the results by Ellison et al. (2010) who also document weaker agglomerative effects from knowledge spillovers. Interestingly, controlling for the dissimilarity proxies does not change in any meaningful way the three Marshallian coefficients, suggesting that access to natural resources and non-manufacturing industries does not bias the results in simple models without the additional controls.

Columns (3) and (4) present coefficients from the multivariate regressions. We still find labor pooling to have the strongest relationship with coagglomeration with an estimated effect of approximately 0.16 of a standard deviation. On the other hand, the coefficients on input sharing and knowledge spillovers decline to 0.082 and 0.024-0.031, respectively. All in all, our findings are comparable to Ellison et al. (2010), with all three of Marshall's forces showing a positive relationship with coagglomeration.[14]

We carry out a number of robustness checks.[15] First, we study whether upstream linkages are more important than downstream connections. We find that the effect of input sharing is twice as large as the effect of output sharing, but this distinction is not significant and does not affect the other coefficients. Second, we investigate whether focusing on a specific year in our sample changes the overall picture. To do so, we run regressions for 1997, 2002 and 2008 separately. We find a slight attenuation in the effects of LP, IO and KS as we move towards more recent years, but the differences are not substantial. Third, we investigate whether using the proxy for knowledge spillovers based on the sector-of-use (SOU) probabilistic mapping affects the findings. The conclusions reached so far still hold: all three Marshallian forces matter, though the effect of labor pooling seems somewhat stronger. We also find that the link with knowledge spillovers is stronger with this proxy, while the effect of input sharing is weaker. Since the SOU mapping is partly based on the technology (and the related patents) contained in goods bought and sold as intermediates

[13]We will return to this point in Section III.B.

[14]Note that we follow Ellison et al. (2010) and do not correct γ^C for differences in the variance of the area-industry employment shares. We assess the robustness of our findings against this issue by including in our specification industry *i* and industry *j* dummies. When we do this, we find very similar labor-pooling effects; slightly smaller, but still significant input-output effects; and larger and more precisely estimated knowledge-spillover effects.

[15]Results are not tabulated for space reasons but can be accessed through our Web Appendix.

across industrial sectors, it incorporates some of the linkages stemming from input sharing and attenuates the effect of IO. Given this issue, our preferred proxy is the one based on the industry-of-manufacture (IOM) mapping which we will use throughout the rest of the paper.[16]

Finally, we check that our results are not affected if we change our measure of coagglomeration to: *(a)* be based on number of plants as opposed to total employment; *(b)* be based on local units belonging to single-plant enterprises only; *(c)* be based on both urban and rural areas. We also experiment with excluding publishing (SIC221) and printing and related activities (SIC222) since these sectors are classified among services in the US industrial classification. None of these robustness checks affect our findings.[17]

B. Addressing endogeneity concerns

The literature on the microfoundations of agglomeration economies has put forward two sources of possible bias in OLS estimation: *(i)* reverse causation; and *(ii)* sorting. In this section, we discuss these issues and provide a set of robustness checks and instrumental variable (IV) estimates to address them.

The reverse causation argument is laid out in Ellison et al. (2010). Firms in industries with strong Marshallian links could choose to locate together in order to benefit from those links. Alternatively, firms that locate together for other reasons could later forge Marshallian links. In contrast to Ellison et al. (2010), we see the reverse phenomenon of coagglomeration leading to productive links as being itself a type of agglomeration economy. For instance, if two firms realize after choosing locations that they can hire from the same labor market, then they benefit from labor pooling. Similarly, if two firms learn from each other *ex post*, then the resulting technological improvement is an instance of knowledge spillovers. These agglomeration economies are in fact in the spirit of Jacobs (1969), who gives numerous examples of accidental agglomeration economies. Even so, we describe below a strategy to address this issue and arrive at estimates that capture the effect of Marshallian links on agglomeration, rather than the reverse.

As for sorting, the main concern is that agglomeration – which increases productivity in possibly unobservable ways – might be correlated with coagglomeration. To clarify matters, consider two industries, e.g. apparel and printing/publishing, which are agglomerated for historical reasons in London. Assume that both industries are highly productive because of some advantages connected to this location. Further assume that more productive industries are able to use a wider range of workers because they are better at spotting

[16]One related concern is that input-output linkages partly capture knowledge spillovers because our KS proxy measures the latter imprecisely. To investigate this issue, we run specifications where we include two-way and three-way interactions of the Marshallian forces. The only significant interaction is the one between IO and KS, with a negative and significant coefficient of -0.009. The effects of IO and LP remain very similar, while the effect of KS rises to around 0.100 (significant). This suggests that the input-sharing and knowledge proxies do *not* capture similar effects. This pattern is consistent with the sectoral heterogeneity presented later in the paper, where we show many instances in which pairs that significantly respond to knowledge spillovers are less affected by input-output links.

[17]Results are not tabulated for space reasons, but are available upon request.

the ‘right types’ in a large agglomerated market. Conversely, think of two other industries, e.g. wood laminate and manufacturing of furniture, that operate in a small city, are low productivity and not efficient at sharing workers. Estimating the effect of labor pooling on coagglomeration by comparing these pairs would bias the results by conflating the true effect of LP with a productive advantage arising because of urbanization economies enjoyed by firms locating in more agglomerated places. Although this argument is logically correct, the unobservables that would give rise to these patterns would need to have a particular structure and imply that agglomeration is correlated to both coagglomeration *and* the strength of the linkages between sectors measured by our proxies. One of the advantages of the method developed by Ellison et al. (2010) is that, by studying the relation between co-location and industry-pair links, the approach deals with a number of unobservables that are not easily related to pair-specific linkages. Therefore, we believe the arguments brought forward in the literature are not sufficiently strong to undermine our findings. Nevertheless, we next provide a number of additional results that lend support to our conclusions.[18]

A first set of regression results is presented in Table A3 in the Appendix. Column (1) mitigates reverse causation by staggering our regressions and considering the effect of the three Marshallian forces measured up to 1999 on coagglomeration γ^C for the years 2000-2008. This check confirms our previous results. Columns (2) to (5) investigate whether any correlation between agglomeration and coagglomeration has the potential to bias our findings. To begin with, we exclude London – the biggest agglomeration in the UK – from the calculations of γ^C and re-estimate our empirical models. Although we find that the effect of KS is attenuated, our broad conclusions are unaffected. In Columns (3) and (4) of the table we include proxies for the extent of agglomeration of the areas where the two sectors in the pair are operating. In particular, we include: *(i)* the mean population density of all the TTWAs in which the sectors are operating, averaged across the pair (Column 3); and *(ii)* the mean employment density of all the areas where the sectors are operating, averaged across the pair (Column 4).[19] Employment density is calculated as total employment across all sectors in a TTWA divided by the area size expressed in square kilometers, so this proxy captures general urbanization economies – much as population density – stemming from operating in a larger market.[20] Adding these controls to our regressions has little effect on our estimates. Finally, in Column (5) we add to our specification the average Herfindahl index across the sector pair to check whether industrial concentration (as opposed to urbanization economies) affects our findings. Once again, we find no evidence that our results are sensitive to these considerations and confirm our previous conclusions.

[18]The model in Davis and Dingel (2013) also predicts equilibria that are consistent with coagglomeration in the absence of sector-pair linkages. This is because in their framework labor supply reflects urbanization economies and sectors with similar skill intensities exhibit similar relative employment shares. Our robustness checks address this possibility.

[19]Controlling for the dissimilarity of employment/population density, instead of the mean, does not affect our results.

[20]The correlation between the two urbanization proxies and coagglomeration is small and negative at -0.148 for population density and -0.082 for employment density. These numbers shrink to zero and 0.007 if we exclude London.

To conclude this section, we discuss a number of IV regressions where we instrument the three Marshallian forces using proxies constructed using US data. This approach follows Ellison et al. (2010). We instrument LP using a measure of the correlation between sector pairs in their use of different types of workers as categorized by the National Industrial-Occupation Employment Matrix (NIOEM) published by the Bureau of Labor Statistics. We instrument IO with an identical measure obtained using the 1987 Benchmark Input-Output Accounts published by the Bureau of Economic Analysis (BEA). Finally, we instrument the flows of patent citations among UK inventors as recorded by the EPO using the flows of citations among US inventors as tracked by the NBER Patent Database. More information is provided in the Web Appendix.[21] The validity of this approach relies on thoroughly controlling for co-location that is driven by natural advantages and shared use of non-manufacturing resources. Hence in all our IV specifications we control for the proxies for sector dissimilarities.

Results are presented in Table 3. Following Ellison et al. (2010), we exclude pairs where the two three-digit industries fall in the same two-digit group and a number of sectors that were aggregated in the data construction process. Columns (1) and (3) show that OLS results do not change as a result of these exclusions. Column (2) presents IV regressions where we include and instrument one Marshallian force at the time. The IV coefficients are very close to their OLS counterparts in Column (1). Column (4) presents multivariate IV regressions where we enter and instrument all three Marshallian forces simultaneously. We find positive and significant effects for LP and IO. The size of the associations is similar to the OLS counterparts (see Column 3). However, KS loses its significance and turns slightly negative. We believe this is due to collinearity between measures that makes instrumented knowledge spillovers hard to disentangle from labor pooling and input sharing. A similar argument is put forward by Ellison et al. (2010) who report in their Appendix weak results when instrumenting KS. To partly address this issue, in Columns (5) to (7) we enter the proxies for Marshallian forces two at the time. In Column (5), we include LP and IO and confirm that both have a positive and significant association with γ^C. In Column (6), we consider IO and KS. We find that both measures are positively and significantly associated with coagglomeration and that the KS estimate is very similar to the one documented using OLS (see Column 3). Finally, in Column (7) we instrument LP and KS and find that both are positively associated with coagglomeration. Although only LP is significant at conventional levels, the coefficient on KS points in the right direction and is reasonably sizeable – at about half of its OLS counterpart. All in all, the evidence in Table 3 confirms our previous findings and supports our claim that endogeneity is unlikely to significantly bias OLS results.

[21]We are extremely grateful to William Kerr for sharing his data and codes with us.

IV. Heterogeneous agglomeration: Jacobs meets Marshall

This section begins the presentation of results that allow for heterogeneity across industries. Specifically, it takes a new approach towards examining Jacobs' (1969) analysis of how new work is created by exploring complementarities between Jacobs and Marshall. The approach has at its core a simple idea: the coagglomeration of industries that only rarely co-locate is different from the coagglomeration of industries that are often found together. It is the former that captures the sorts of unplanned, unpredictable, or unusual interactions that Jacobs has in mind. To test this idea, we estimate the Marshallian models discussed in the previous section without constraining the Marshallian forces to have the same effect for all industry pairs. More precisely, we estimate equation (2) in a way that allows the effects to vary between the most and least coagglomerated pairs. This estimation is carried out using quantile regressions that simultaneously include all three Marshallian forces as well as controls for natural advantages. Figures 1a–1c present the results for labor pooling, input sharing, and knowledge spillovers, respectively. The confidence intervals on the figures come from bootstrapped standard errors clustered on industry pairs.[22]

It is immediately clear that the pattern of aggregate results in Table 2 conceals considerable variation across industry pairs. Figure 1a presents results for labor pooling. There is clear heterogeneity across pairs according to their coagglomeration. While labor pooling has a positive and statistically significant contribution to industry-pair coagglomeration across the board, the effect is much larger for the less coagglomerated pairs. Labor pooling has an association of around 0.22 and 0.16 (both significant) for industry pairs in the two bottom deciles, declining to around 0.06-0.08 (significant) in the top half of the coagglomeration distribution. Figure 1b shows a pattern for input sharing that is exactly opposite. For this force, the association is larger for the most coagglomerated pairs. The input-sharing coefficient increases from approximately 0.03 (insignificant) for the bottom decile to 0.15 and 0.23 (both significant) for the top two deciles. Figure 1c shows yet another pattern, with the effect of knowledge spillovers positive and significant up to the 60th percentile. The coefficient becomes smaller over the range, declining from 0.03 in the bottom decile to 0.02 at the median, and the estimation becomes increasingly imprecise for the most coagglomerated industries, for which the effect is no longer significantly different from zero.

The pattern of heterogeneity has interesting implications for Jacobs' ideas. The input sharing results in Figure 1b are contrary to Jacobs. They suggest that input sharing primarily is associated with the co-location of pairs that coagglomerate extensively. There is little to be gained from links between industries that are not very coagglomerated. In other words, there is strong evidence that interactions that are most

[22] We performed a similar analysis where we investigate heterogeneity in the effect of the Marshallian forces by considering the quantiles of the unconditional coagglomeration distribution. This approach yields similar results. This is not surprising given that the effects of the three Marshallian variables are not significantly affected by the inclusion of controls as shown in Table 2. See Frolich and Melly (2010) for a related discussion.

typical and are likely to be planned have the largest association with coagglomeration. It is useful to consider an example from Marshall. He writes that:

> "Many cutlery firms [in Sheffield] for instance put out grinding and other parts of their work, at piece-work prices, to working men who rent the steam power which they require, either from the firm from whom they take their contract or from someone else." (Marshall, 1890, 8th ed., p. 172).

Our data show that Sheffield continues to be a center of cutlery production today. Moreover, cutlery and manufacturing of basic iron and steel is one of the most highly coagglomerated pairs. It is entirely understandable that a cutlery maker would deliberately plan its location in a way that secures its metal input supply. See Table A1 in the Appendix for some other highly coagglomerated industry pairs, such as spinning of textiles and textile weaving – also discussed by Marshall in a similar fashion.

On the other hand, the results on labor pooling and knowledge spillovers are much more in the spirit of Jacobs. Regarding knowledge, the effects are not even significant for highly coagglomerated pairs. As for labor pooling, the effects diminish drastically as coagglomeration increases. In other words, both of these sorts of interactions between industries have a larger effect when the industries co-locate less frequently and interactions are more likely to be the sort of unexpected connections on which Jacobs focuses.[23]

Three issues are worth considering. First, working at the TTWA level of aggregation could affect our findings. [24] As noted previously, TTWAs are defined by commuter flows, which depend on the scale of labor markets. This could make it more likely to find a relationship between coagglomeration and labor pooling than one between coagglomeration and input sharing, since input-output linkages could take place over greater distances. To address this issue, we replicate the analysis using 18 macro-regions (as defined in the BSD data). Our results are presented in Figure A1 in the Appendix and fully confirm our findings. We also check whether our results change if we maintain the original TTWA geography, but focus only on the 28 biggest cities and conurbations out of 84, i.e. the top 30%. We find this is not the case.

Second, the labor pooling proxy is based on the correlation between the two industries' occupation mixes, while the input sharing variable is constructed using maximum flows in the sector pair. This could imply that our approach is skewed towards picking up significant input-output linkages only for highly coagglomerated industries, while the labor pooling measure could be more significant in other parts of the distribution. We believe this issue does not affect our conclusion since we find that knowledge spillovers behave very much like labor pooling – despite being measured in the same way as input sharing (i.e. as the

[23]There is an interesting parallel here to the Duranton and Puga (2001) nursery-city phenomenon: certain interactions have greater effects with less frequent co-location; others have greater effects with more frequent co-location.
[24]Rosenthal and Strange (2003) provide evidence that the Marshallian forces have different effects on agglomeration depending on the geographical scale of analysis.

maximum flow of patent citations across pairs). This suggests that the pattern we observe is not mechanically driven by the way our proxies are constructed.

Third, we do not observe whether or not coagglomeration has arisen from unplanned or unpredictable accidents, as in Jacobs. Instead, we only observe the frequency of an industry-pair coagglomerating. Although this is likely to be related to whether coagglomeration is planned or unplanned, other forces also contribute to whether an industry pair coagglomerates. For instance, pollution controls could make it more difficult for heavy industries to co-locate. Agglomeration/dispersion forces such as these introduce noise into the process determining coagglomeration and create a sort of measurement error in the mapping between the planned/unplanned nature of agglomeration and co-location frequency. Presumably, this kind of measurement issues would make it more difficult to obtain the striking pattern that we see in the quantile models – in particular the remarkable difference between labor pooling and input sharing. This suggests that our findings might understate the heterogeneity of Marshall's agglomeration forces along the planned/unplanned dimension.

V. Heterogeneous agglomeration: Adaptation and organization

A. Non-Marshallian approaches

This section presents our empirical results on the adaptational and organizational aspects of agglomeration economies. We extend the traditional Marshallian approach by examining how the patterns of coagglomeration depend on the interaction between Marshallian forces and the nature of the industry in question. We therefore allow Marshall's microfoundations to be complementary to other explanations.

As a preliminary step, we look directly at the relationship between industry characteristics and coagglomeration by including the industry variables discussed in Section II.C – i.e. year-of-opening, entry share, technology level, college graduate share and size of entrants/incumbents – as controls in equation (2). The variables are averaged across the industry pair, except for the technology variable, where it is not possible to use an average. In this case, we construct dummy variables indicating whether both sectors are high-technology or whether only one industry is. Table 4 presents the results from these models. The first important finding is that the coefficients of the Marshallian forces are fairly constant as different controls are introduced. The coefficients are also similar in magnitude to the corresponding effects in Table 2 showing that the estimates of the Marshallian forces are robust.

Turning to the non-Marshallian variables, some interesting patterns emerge. We find a positive and significant coefficient on year-of-opening (the inverse of age) in Column (1) and an insignificant coefficient on entry share in Column (2). The former result is consistent with the nursery city/unplanned interactions ideas discussed above. The latter is weakly supportive. The dummies for high- and mixed-technology pairs in Column (3) are instead both negative and significant. Controlling for Marshallian forces, we see more coagglomeration of low-technology industries. This result is the opposite of what one might expect to find based on the predictions of a nursery city model. In Column (4), the coefficient of average college share is significant and negative. Given the strength of human capital effects in other models (e.g., Rauch, 1993, or Rosenthal and Strange, 2008), this is unexpected.[25] Finally, entrant size has a positive and marginally significant coefficient in Column (5), as does incumbent size in Column (6). Controlling for Marshallian forces, we do not find much of a small firm effect. As shown in Column (7), jointly controlling for all the non-Marshallian factors does not affect our conclusions. All coefficients retain their sign, size and significance – with the exception of the proxy for the size of entrants. This is not surprising given its high correlation (0.712) with the proxy for the size of incumbents. In sum, simply including controls for non-Marshallian forces using industry-pair averages in a coagglomeration/microfoundation model, while failing to allow for heterogeneity, generates weak and sometimes puzzling results. We now therefore turn to less restrictive models that let the Marshallian effects differ across industries.[26]

B. New industries and entry

We begin to look at adaptation and nursery city ideas by estimating models based on two partitions of industry pairs. The first focuses on industry age, captured by the age of the oldest active plant. The second focuses on the share of new firm entry in the industry. This approach generates the partitions of pairs detailed in Section II. In the case of age, the new industries group includes pairs where both sectors are younger than the median age across industries. In the case of entry, the dynamic industries group includes pairs where both industries have an entrant share that is above the median. In both cases, we include additional controls for, respectively, age and entry share averaged across the two sectors in the pair.

Results are reported in Table 5. For industry age, we find the largest agglomeration effects for new industry pairs. This is true for all three Marshallian forces. For labor pooling, the effects are smaller for the mixed and the old industry pairs (at 0.153 and 0.081, respectively) than for new pairs (at 0.310). However, all coefficients are highly significant and the extent of variation in these effects is more muted than for the other two Marshallian forces. The knowledge spillover results are very much in the nursery city/unstable

[25]These results hold if we exclude London. Conversely, dropping the three Marshallian proxies from the specifications yields insignificant estimates of the effect of either human capital or technology on coagglomeration.

[26]Our models are at the industry-pair level. Because of this, 'binning' provides a more straight-forward approach to study complementarities between Marshallian and non-Marshallian theories than interacting Marshallian forces with industry characteristics. Consider, for example, firm size. We are interested in how Marshallian links between industry pairs relate to firm size, as in Chinitz. Following the 'binning' approach, we construct three groups of industry pairs. In the first, both industries are characterized by small firms; in the second, the pair is characterized by large firms. Chinitz makes sharp predictions about what we should expect for these groups. The third group has 'mixed' pairs, with one characterized by small firms and the other by large firms. In this case, Chinitz does not make any predictions. While it would be possible to construct an average firm size variable and estimate an interactive specification, the presence of mixed pairs would compromise the interpretation of the results from such an approach.

industry spirit discussed above. They show knowledge effects that are five to ten times stronger for young pairs, at 0.236 (significant), than for mixed and old pairs, at 0.040 (significant) and 0.026 (insignificant), respectively. The same is true for the results on input sharing. The coefficients move from 0.270 (significant) for new industry pairs to 0.049 (insignificant) for the mixed group, and finally to 0.041 (significant) for old pairs. While the pattern of effects on input sharing is not consistent with a nursery city model, the heterogeneity in the coefficients for knowledge spillovers clearly supports this theory.[27]

The results have a relatively similar pattern for industry dynamism. Labor pooling is always significant with the coefficients for the three groups fairly constant and ranging between 0.181 and 0.144. We still find that the largest result for knowledge spillovers occurs for the dynamic industries at 0.181 (significant). This shrinks to 0.033 (significant) and -0.020 (insignificant) for mixed and steady industry pairs. Input-output linkages are closer to a nursery pattern in these dynamic-industry models than in the previous age grouping. Input sharing displays significant coefficients for the mixed and the steady pairs, at 0.103 and 0.052 respectively, and has no significant effect for dynamic industries.[28]

C. High-technology and high-education

We now turn to the related issue of how the relationship between Marshallian forces and coagglomeration depends on the technological status of the industry in question. As noted above, we characterize an industry's technological status (high-technology or not) according to the OECD (1997) classification. This generates three types of industry pairs: both high-technology, both low- technology, or mixed. We estimate equation (2) for each type.

Results are reported in Table 6. We find that labor pooling is significant in all three groups, but its association with coagglomeration is much larger for the low-technology industry group (at 0.332) than for the high-technology group (at 0.046). Input sharing also has the largest coefficient in the low-technology group, at 0.091. While this Marshallian force has positive coefficients for all three groups, the high-technology coefficient is small and insignificant. Knowledge spillovers display the opposite pattern. The largest coefficient is found for high-technology (significant at 0.053), while the effect becomes smaller and insignificant for low-technology (at 0.039).

These results clearly show that agglomeration economies are not simply a high-technology phenomenon. Labor pooling has a stronger effect in the low-technology group, while input sharing is stronger in the mixed- and low-technology groups of industry pairs. Knowledge spillovers, reassuringly, is

[27]We further investigate whether new/old pairs respond differently to Marshallian linkages when they are measured closer/further in time relative to coagglomeration. To do this, we run separate regressions for 1997, 2002 and 2008. The patterns presented in Table 5 are confirmed with no evidence of additional significant heterogeneity.

[28]We checked that our results continue to hold if we focus on single-plant enterprises to identify the cut-offs used to define groups in Table 5 so that new entrants are stand-alone ventures, and not expansions of existing activities.

different – with the largest effect for high-tech industries. This suggests that some of the weaker results for knowledge spillovers reported above, and also presented in Ellison et al. (2010), arise because the sample includes low-technology industries (as well as old and steady industries) where knowledge spillovers are not important.

Table 6 also tabulates results of a similar exercise where industries are partitioned according to the education levels of their workers. In these specifications, we further control for the average share of college graduates across the pair in order to control for direct effects of this variable within groups. The pattern of results is similar to the high-technology vs. low-technology heterogeneity discussed above. Knowledge spillovers have significant effects in high-education (at 0.048) and mixed-education (at 0.050) industry pairs, but not in low-education pairs (insignificant at 0.030). Conversely, input sharing and labor pooling have the largest and most significant effects in low-education pairs, at 0.123 and 0.391 respectively. These shrink to 0.007 (insignificant) and 0.046 (borderline significant) for the high-education pairs.

Taken as a group, these findings are broadly consistent with learning playing an important role in the agglomeration process. Jacobs (1969) calls this phenomenon "the creation of new work". Vernon (1960) instead discusses the process by which new products reach stability. The evidence is also consistent with Duranton and Puga's (2001) nursery city phenomenon, where new products are created in diverse cities and move to specialized cities upon reaching maturity. They provide evidence of firm migration following this pattern in France to support their conclusions. To the best of our knowledge, our paper is the first to examine coagglomeration in this light. The observation that only high-technology/high-education pairs are found to have their coagglomeration associated with stronger knowledge links between the industries is consistent with the nursery city idea. So is the finding that low-technology/low-education pairs have coagglomeration associated with the somewhat more routine labor and input links.

D. Industrial organization

The final set of results deals with industry structure. We consider both a partition based on the size of entrants and another one based on the size of incumbents. Both splits correspond to Chinitz (1961), who argues that the presence of small firms allows entry by other small firms.[29] Naturally, there are other treatments of agglomeration that similarly depend on industry structure, and they will be discussed as well.

Results are presented in Table 7. As above, the table reports results for the three Marshallian forces estimated over three groups of industry pairs. The models include controls for, respectively, entrant size and

[29] Strictly speaking, Chinitz's argument is about heterogeneity in the industrial organization of cities – not of sectors. However, we find that approximately 96% of the variation in the size of entrants and 99% of the variation in the size of incumbents is within-TTWAs across sectors, with the remaining part being within-sectors across TTWAs. This suggests that our analysis that treats industrial composition as fixed across cities and only focuses on sectoral heterogeneity captures the most relevant variation.

incumbent size averaged across the pair, alongside the usual controls for natural advantages. The results are consistent with a small-firm effect, with the largest coefficient found on input sharing for the small entrants' and small incumbents' models (significant at 0.193 and 0.159, respectively). It is worth noting that input sharing is significant for all groups in both models. However, the effects are substantially smaller for industries characterized by mixed- and large-entrants and incumbents, and ranging between 0.068 and 0.082.[30] The labor pooling coefficients are all significant and comparable in magnitude in the three entrant- and the three incumbent-size models. Estimates are above 0.10 for all but one grouping and mainly in the 0.15-0.20 range. Finally, as in all the regressions so far, knowledge is not a statistically significant predictor of coagglomeration in the universe of models. In this case, we find that knowledge flows are only significant and sizeable in the small-entrant sample (at 0.051) and in the mixed-entrant group (at 0.028). They are instead insignificant and small, but always positive and between 0.020 and 0.040, for all three groups of incumbents.

Chinitz focused largely on input sharing among small firms as a driver of agglomeration. Our results are consistent with his approach. Conversely, Vernon and Jacobs offer anecdotes of knowledge spillovers generated by large and small firms alike. Our results on coagglomeration seem to suggest, however, that the effect of knowledge flows is more systematic for small entrants. Regarding labor pooling, neither Vernon nor Jacobs directly engages with the implications of firm size for this agglomerative force. While the coefficients on labor pooling are significant for all sizes of entrants and incumbents, we find the smallest coefficients for small entrants and small incumbents. One factor that could potentially come into play is labor poaching, where firms hire away each other's skilled workers (Combes and Duranton, 2006). If small firms are threatened to a greater degree by the possibility of poaching, they might be less likely to co-locate with firms hiring from the same labor pool.

E. Robustness

In this section we discuss a number of issues that could affect the findings reported in Tables 5-7 and report on a series of robustness checks to address them. First, we consider again whether the spatial scale used to construct γ^C affects our results. Following the approach taken in Section IV, we measure coagglomeration using regions instead of TTWAs and re-run our analysis. As shown in Table A4 in the Appendix, the patterns discussed above continue to hold. We also find that our results do not change if we calculate γ^C using both urban and rural TTWAs (results not tabulated). Second, in the regressions presented so far, we controlled for the attribute used to partition the sample averaged across the pair – e.g. the average

[30]Once again, we checked that our results hold if we focus on single-plant enterprises to define groups in Table 7 to limit the possibility that spinoffs and expansions of existing activities affect the observed pattern.

entry share or the average size of incumbents. In some extensions (not tabulated for space reasons), we check whether controlling for the dissimilarity of the sector-pair's characteristics affects our results. We measure dissimilarity as (half of) the absolute value of the difference in the shares of the relevant attribute across the pair. We find that controlling for dissimilarity produces very similar results to controlling for the average. A third issue is that the pattern documented above might be related to the extent of localization of the industry pairs in the different groups. To consider this possibility, we perform all the regressions as in Tables 5 to 7 adding a control for the average localization index (i.e., γ from Ellison and Glaeser, 1997) across the pair. This does not affect the results in any significant way.[31] Fourth, we assess the robustness of our findings relative to the details of the industrial classification we use. In particular, we consider whether our results are driven by the presence of two-digit sectors that are subdivided into many three-digit groupings. To do so, we exclude all two-digit industries partitioned in more than five three-digit sub-groups. Although this leaves us with a quarter of the original sample, this exclusion does not affect our results. Alternatively, we drop all pairs where the two sectors belong to the same two-digit industry (as we did in our IV estimation). This also does not affect our findings.

We also considered potential endogeneity issues. Our approach here is similar to our approach in Section III.B. To check whether any correlation between coagglomeration and agglomeration biases our findings, we estimate the models reported in Tables 5-7 excluding London. We also estimate specifications that control for density measured either as employment or population per land area. These extensions do not affect our conclusions. Furthermore, we estimate IV models which deal with both reverse causation and omitted variables.[32] Our key results continue to hold.

VI. Conclusions

This paper has considered heterogeneity in the microfoundations of agglomeration economies using patterns of coagglomeration of UK industries. The analysis has reached several key conclusions. First, there is notable heterogeneity between industries. Different industries agglomerate for different reasons. Second, this heterogeneity is consistent with Jacobs' (1969) ideas about unplanned interactions being an important aspect of the agglomeration process for labor pooling and knowledge spillovers but not for input sharing. Third, the pattern also provides support for the idea of nursery cities (Duranton and Puga, 2001) in particular, and adaptive agglomeration economies more generally (Vernon, 1960). Last, the results are consistent with Chinitz's (1961) idea that agglomeration effects are stronger in industries dominated by small firms.

[31] Note that the correlation between co-agglomeration and localization of the industry pairs is small and negative at -0.028, so localization features of industries cannot explain the patterns of Figures 1a-1c.

[32] We estimate both univariate and multivariate IV models that yield similar results.

Taken together, this suggests fairly strongly that Marshall's treatment of the microfoundations of agglomeration should be treated as complementary to the analysis of Jacobs and others, rather than as an alternative and competing explanation. Recognizing and interpreting the pattern of heterogeneity through a framework that emphasises the synergies between Marshallian and non-Marshallian approaches has important implications for our understanding of the nature of agglomeration economies.

A second reason why it is essential to recognize this heterogeneity is that there are numerous instances in the agglomeration literature where the circumstances of individual industries and clusters are presented as having broad relevance across industries. Without doubt, the computer industry is the most salient industry in the agglomeration literature. Saxenian (1994) offers an important and often quoted analysis of the Silicon Valley. The car industry is also highly salient in the agglomeration literature. In the US, this industry's declining cluster centered around Detroit is often contrasted to the prosperous computer cluster in Great San Jose. A very informative discussion along these lines can be found in Glaeser (2011).

Our evidence shows clearly that different industries respond differently to agglomerative forces. Therefore, while the detailed analysis of individual cases is often informative – as attested by the influence of this kind of extrapolation – it is important not to simply accept generalizations without further investigation. This point is all too often lost in discussions with policymakers who fail to recognize the uncertainties associated with heterogeneity in agglomeration economies. Given this complexity, cluster policies targeting specific industries are as likely to pick losers as winners. Conversely, policies that are likely to promote growth in general – such as better transport infrastructures, higher levels of education, better amenities and housing to attract workers and entrepreneurs – are more likely to help the 'right' clusters to emerge as a result of the underlying local strength and agglomeration forces, even when these are heterogeneous and therefore difficult to identify *ex ante*.

References

Almazan, A., A. De Motta, and S. Titman (2007), “Firm location and the creation and utilization of human capital,” *Review of Economic Studies* 74, 1305-1327.

Arzaghi, M. and J.V. Henderson (2008), “Networking off Madison Avenue,” *Review of Economic Studies* 75, 1011-1038.

Audretsch, D.B. and M.P. Feldman (1996), “R&D spillovers and the geography of innovation and production”, *American Economic Review* 86, 630-640.

Behrens, K., and F. Robert-Nicoud (2014): “Agglomeration Theory with Heterogeneous Agents”, G. Duranton, J. V. Henderson and W. Strange (eds), *Handbook in Regional and Urban Economics*, Amsterdam (Holland), Elsevier Press.

Berry, C.R. and E.L. Glaeser (2005), “The divergence of human capital levels across cities”, *Papers in Regional Science* 84, 407-444.

Chatterji, A., Glaeser, E. L., & Kerr, W. R. (2013), “Clusters of Entrepreneurship and Innovation”, NBER Working Paper 19013.

Chinitz, B. (1961), “Contrasts in agglomeration: New York and Pittsburgh”, *American Economic Review* 51, 279-289.

Combes, P.-P., and L. Gobillon (2014): “The Empirics of Agglomeration Economies”, G. Duranton, J. V. Henderson and W. Strange (eds), *Handbook in Regional and Urban Economics*, Amsterdam (Holland), Elsevier Press.

Combes, P.-P., and G. Duranton (2006), “Labour Pooling, Labour Poaching and Spatial Clustering”, *Regional Science and Urban Economics* 36, 1-28.

Davis, D. R., and Dingel, J.I. (2013), “The Comparative Advantage of Cities”, Working paper.

Duranton, G. (2011), “California dreamin’: the feeble case for cluster policies”, *Review of Economic Analysis* 3, 3-45.

Duranton, G. and D. Puga (2001), “Nursery cities”, *American Economic Review* 91(5), 1454-1477.

Duranton, G. and D. Puga (2004), "Micro-Foundations of Urban Agglomeration Economies," in Henderson, J.V. and Thisse, J.-F. (Eds.), *Handbook of Urban and Regional Economics, Volume 4*, Elsevier, Amsterdam, 2063-2117.

Ellison, G. and E.L. Glaeser (1997), “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach”, *Journal of Political Economy* 105, 889-927.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010), “What causes industry agglomeration? Evidence from coagglomeration patterns”, *American Economic Review* 100, 1195-1213.

Fallick, B., C. Fleischman and J. B. Rebitzer (2006), “Job hopping in the Silicon Valley: The micro-foundations of an industry cluster”, *Review of Economics and Statistics* 88, 472-481.

Florida, R. (2003), *Rise of the Creative Class: And How It’s Transforming Work, Leisure, Community and Everyday Life*, New York: Basic Books.

Frolich, M. and B. Melly (2010), “Estimation of Quantile Treatment Effects using Stata”, *The Stata Journal*, 10(3), 423-457.

Gabe, T. M., and J. R. Abel (2013): “Shared knowledge and the coagglomeration of occupations”, *Federal Reserve Bank of New York, Staff Report* 612.

Glaeser, E.L. (2011), *The Triumph of the City, How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier and Happier,* New York, NY: Penguin US.

Glaeser, E.L., H. Kallal, J. Scheinkman, and A. Schleifer (1992), “Growth in cities”, *Journal of Political Economy*, 100, 889-927.

Glaeser, E.L. and W.R. Kerr (2009), “Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain?”, *Journal of Economics and Management Strategy* 18, 623-663.

Glaeser, E.L. and A. Saiz (2004), "The rise of the skilled city", *Brookings-Wharton Papers on Urban Affairs*, 47-105.

Hanlon, W. and A. Miscio (2014), "Agglomeration: A Dynamic Approach", mimeo.

Hanson, G.H. (2001), "Scale Economies and the Geographic Concentration of Industry", *Journal of Economic Geography*, 1(3), 255-276.

Helsley, R.W. and W.C. Strange (2001), "Innovation and input sharing", *Journal of Urban Economics* 51, 24-45.

Helsley, R.W. and W.C. Strange (2014), "Coagglomeration, Clusters, and the Scale and Composition of Cities", *Journal of Political Economy*, 122(5), 1064-1093.

Henderson, J.V., A. Kuncoro and M. Turner (1995), "Industrial development in cities", *Journal of Political Economy*, 103, 1067-1085.

Holmes, T.J. (1999), "Localization of industry and vertical disintegration", *Review of Economics and Statistics* 81, 314-25.

Jacobs, J. (1969), *The Economy of Cities*. New York, NY: Vintage.

Jacobs, W., H. R. A. Koster and F. Van Oort (2013): "Co-agglomeration of knowledge-intensive business services and multinational enterprises", *Journal of Economic Geography*, 14, 443-475.

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quarterly Journal of Economics*, 108(3), 577-598.

Jofre-Monseny, J., R. Marín-López and E. Viladecans-Marsal (2011), "The mechanisms of agglomeration: Evidence from the effect of inter-industry relations on the location of new firms", *Journal of Urban Economics* 70, 61-74.

Kolko, J. (2010): "Urbanization, agglomeration, and coagglomeration of service industries", in Glaeser, E. L. (ed), *The Economics of Agglomeration*, 151–180, Cambridge, Mass., National Bureau of Economic Research Press.

Krugman, P. (1991), "Increasing returns and economic geography", *Journal of Political Economy* 99, 483-499.

Lin, J. (2012), "Technological adaptation, cities, and new work," *Review of Economics and Statistics* 93, 554-574.

Marshall, A. (1890), *Principles of Economics*, MacMillan: London.

Moretti, E. (2004): "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions", *American Economic Review*, 94, 656-690.

OECD (1997), "Revision of High-Technology Sector and Product Classification", STI Working Paper 1997/2.

Office for National Statistics, *Business Structure Database, 1997-2011: Secure Access* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], October 2012. SN: 6697.

Office for National Statistics. Social Survey Division, and Northern Ireland Statistics and Research Agency. Central Survey Unit, *Quarterly Labour Force Survey, 1992-2011: Secure Access* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], May 2013. SN: 6727.

Overman, H.G. and D. Puga (2010), "Labor pooling as a source of agglomeration: An empirical investigation", in E. L. Glaeser (Ed.), *Agglomeration Economics - NBER Collections*, Chicago, IL: The University of Chicago Press, 133-150.

O'Sullivan, A. and W. Strange (2015), "The emergence of coagglomeration," Working paper.

Porter, M. (1990), *The Competitive Advantage of Nations*, New York, NY: The Free Press.

Rauch, J. (1993), "Productivity gains from geographic concentration of human capital: Evidence from the cities", *Journal of Urban Economics* 34, 380-400.

Rosenthal, S. and W.C. Strange (2001), “The determinants of agglomeration”, *Journal of Urban Economics* 50, 191-229.

Rosenthal, S. and W.C. Strange (2003), “Geography, industrial organization, and agglomeration”, *Review of Economics and Statistics* 85, 377-393.

Rosenthal, S. and W.C. Strange (2004), “Evidence on the Nature and Sources of Agglomeration Economies”, in Henderson, J.V. and Thisse, J.-F. (Eds.), *Handbook of Urban and Regional Economics, Volume 4*, Amsterdam: Elsevier, 2119-2172.

Rosenthal, S., and W.C. Strange (2008), “The attenuation of human capital spillovers”, *Journal of Urban Economics* 64, 373-389.

Rosenthal, S. and W.C. Strange (2010), “Big Establishment/Small Effects: Agglomeration, Industrial Organization, and Entrepreneurship”, Chapter 10 in E. L. Glaeser (Ed.), *Agglomeration Economics - NBER Collections*, Chicago, IL: The University of Chicago Press, 277-302.

Saxenian, A.-L. (1994), *Regional Advantage: Cultural Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Serafinelli, M. (2015): “Good Firms, Worker Flows and Local Productivity”, mimeo, University of Toronto.

Silverman, B.S. (2002), *Technological Resources and the Logic of Corporate Diversification*, London, UK: Routledge.

Strange, W.C., W. Hejazi and J. Tang (2006), “The Uncertain City: Competitive Instability, Skills, Innovation, and the Strategy of Agglomeration”, *Journal of Urban Economics*, 59(3), 331-351.

Vernon, R. (1960), *Metropolis 1985*, Cambridge, MA: Harvard University Press.

Table 1: Descriptive statistics

	Mean	*Std. Dev.*	*Min*	*Max*
<u>*Coagglomeration measures and Marshallian forces*</u>				
TTWA total employment coagglomeration (γ^C)	0.000	0.005	-0.028	0.107
Labor pooling (correlation)	0.237	0.188	-0.022	0.968
Input-output sharing (maximum)	0.009	0.033	0.000	0.547
Input sharing (maximum)	0.007	0.029	0.000	0.547
Output sharing (maximum)	0.005	0.021	0.000	0.546
Knowledge spillovers – prob. mapping, industry of manufacture (IOM, maximum of inward/outward citation)	0.016	0.037	0.000	0.413
Knowledge spillovers – prob. mapping, sector of use (SOU, maximum of inward/outward citation)	0.012	0.026	0.000	0.540
<u>*Additional Controls*</u>				
Energy dissimilarity index	0.013	0.016	0.000	0.097
Water dissimilarity index	0.001	0.001	0.000	0.006
Transport dissimilarity index	0.014	0.018	0.000	0.084
Natural Resources dissimilarity index	0.041	0.076	0.000	0.369
Services dissimilarity index	0.018	0.016	0.000	0.082

Note: All pairwise combinations of manufacturing SIC1992 3-digit industries are included except Manufacture of tobacco (SIC160). In addition, we combined Manufacture of leather clothes (SIC181) and Dressing and dyeing of fur (SIC183) with Manufacture of wearing apparel (SIC182); Manufacture of coke oven products (SIC231) and Processing of nuclear fuel (SIC233) with Refined petroleum products (SIC232). We also combined the following sectors: Manufacture of vegetable and animal oils and fats (SIC154) with Manufacture of other food products (SIC158); Manufacture of man-made fibers (SIC247) with Manufacture of other chemical products (SIC246); Manufacture of cement, lime and plaster (SIC265) with Manufacture of articles of concrete, plaster and cement (SIC266); Reproduction of recorded media (SIC223) with Printing (SIC222). Our final sample consists of 94 manufacturing 3-digit sectors for a total of 4,371 unique pairwise correlations a year for twelve years (1997-2008). The complete dataset contains 52,452 observations. Labor correlation indices are computed from the UK Labour Force Survey 1995-1999. Input-Output measures are calculated ONS UK Input-Output Tables for 1995-1999. Knowledge spillover measures are calculated using the UK data retrieved from the EPO-PATSTAT dataset made available to us by Bocconi University. Cited patents sampled for the years 1978 to 1997. Citing patents sampled for the years 1981 to 2000. Additional control measures are calculated using the UK Input-Output tables for 1995-1999.

Table 2: The relationship between coagglomeration γ^C and Marshallian forces

	(1)	(2)	(3)	(4)
Specification details:	OLS – Univar.	OLS – Univar.	OLS – Multivar.	OLS – Multivar.
Labor pooling (LP)	0.191 (0.018)***	0.198 (0.018)***	0.156 (0.019)***	0.165 (0.020)***
Input-output sharing (IO)	0.138 (0.026)***	0.137 (0.027)***	0.083 (0.025)***	0.082 (0.025)***
Knowledge spill. – IOM (KS)	0.106 (0.015)***	0.099 (0.014)***	0.031 (0.013)**	0.024 (0.013)*
Resource use diss. Controls	*No*	*Yes*	*No*	*Yes*

Note: See note to Table 1 for details on variable definitions. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses. All regressions consider the period 1997-2008.

Table 3: Instrumental variable regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS – Univar.	IV – Univar.	OLS – Multivar.	IV – Multivar.	IV – Multivar. LP & IO	IV – Multivar. IO & KS	IV – Multivar. LP & KS
Labor pooling (LP)	0.161 (0.017)***	0.113 (0.020)***	0.133 (0.018)***	0.116 (0.032)***	0.100 (0.024)***	--	0.116 (0.030)***
Output sharing (IO)	0.105 (0.017)***	0.127 (0.026)***	0.061 (0.016)***	0.083 (0.024)***	0.082 (0.024)***	0.121 (0.028)***	--
Knowledge spillovers – IOM (KS)	0.078 (0.013)***	0.088 (0.016)***	0.033 (0.013)**	-0.021 (0.023)	--	0.031 (0.019)*	0.017 (0.022)
First-stage statistics							
t-stat on LP	--	*19.93*	--	*16.91*	*18.78*	--	*17.04*
t-stat on IO	--	*6.39*	--	*6.22*	*6.32*	*5.95*	--
t-stat on KS	--	*6.83*	--	*5.94*	--	*6.45*	*6.56*
Kleinbergen-Paap F-Stat	--	--	--	*18.17*	*33.10*	*23.26*	*23.94*

Note: See note to Table 1 for details on variable definitions. All regressions control for dissimilarity in use of resources. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses. Sample excludes 3-digit SIC sectors within the same 2-digit SIC sector and the following sectors which were aggregated in the data construction: Manufacture of wearing apparel and accessories (SIC182); Manufacture of refined petroleum products (SIC232); Manufacture of other chemical products (SIC246); Manufacture of other food products (SIC158); Manufacture of articles of concrete, plaster and cement (SIC266); and Printing and service activities related to printing (SIC222). Number of observations 43,644 (3,637 industry pairs). Instrumental variable regressions use labor correlation, input-output linkages and patent citations flows calculated using US data as instruments. See Web Appendix for more details. Cells in Column (1) and (2) come from separate regressions. Cells in Columns (3)-(7) come from regressions that simultaneously enter Marshallian forces as detailed in the headings. Knowledge spillovers measure is based on probabilistic mapping – Industry of manufacturing (Knowledge Spillovers – IOM).

Table 4: The relationship between coagglomeration γ^C, Marshallian forces and non-Marshallian mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Additional Control details:	Year of Opening	Entry Share	Tech.	Education	Size of Entrants	Size of Incumbents	Joint Controls
Labor pooling (LP)	0.166	0.165	0.188	0.176	0.164	0.161	0.189
	(0.020)***	(0.020)***	(0.022)***	(0.021)***	(0.020)***	(0.019)***	(0.023)***
Input-output sharing (IO)	0.085	0.082	0.075	0.079	0.080	0.081	0.073
	(0.026)***	(0.025)***	(0.025)***	(0.025)***	(0.025)***	(0.025)***	(0.025)**
Knowledge spill. – IOM (KS)	0.028	0.024	0.027	0.029	0.023	0.024	0.036
	(0.013)**	(0.014)*	(0.014)**	(0.014)**	(0.014)*	(0.014)*	(0.014)***
Year of opening	0.046						0.048
	(0.014)***						(0.014)***
Entry share		0.003					0.003
		(0.012)					(0.014)
High tech			-0.227				-0.223
			(0.055)***				(0.065)***
Mix tech			-0.087				-0.083
			(0.035)***				(0.039)**
Share college graduates				-0.059			-0.045
				(0.018)***			(0.021)**
Size of entrants					0.028		0.009
					(0.014)*		(0.020)
Size of incumbents						0.032	0.058
						(0.016)*	(0.025)**

Note: See note to Table 1 for details on variable definitions. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses. All regressions consider the period 1997-2008, control for dissimilarity in use of resources and include all Marshallian forces simultaneously. Estimates obtained using OLS. Year of opening refers to the first year of openings across all plants in a given industry averaged across the sector pairs. Share of entry refers to the fraction of new firms averaged across sector pairs. High-tech and low-tech industries are categorized according to the OECD classification (1997). The mixed-tech industry pairs consist of one high-tech industry and one low-tech industry. Share of college graduates refers to the average share across the sector pairs. Education level calculated using the UK LFS 1995-1999 data. Size of entrants and size of incumbents refer to the number of employees of new and existing plants averaged across the sector pairs.

Table 5: Heterogeneous agglomeration – Adaptation

	(1)	(2)	(3)
	New	Mixed	Old
Labor pooling (LP)	0.310	0.153	0.081
	(0.058)***	(0.026)***	(0.021)***
Input-output sharing (IO)	0.270	0.049	0.041
	(0.083)***	(0.030)	(0.018)**
Knowledge spillovers – IOM (KS)	0.236	0.040	0.026
	(0.121)**	(0.017)**	(0.019)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081
	Dynamic	Mixed	Steady
Labor pooling (LP)	0.181	0.144	0.180
	(0.059)***	(0.026)***	(0.028)***
Input-output sharing (IO)	0.113	0.103	0.052
	(0.095)	(0.031)***	(0.021)**
Knowledge spillovers – IOM (KS)	0.181	0.033	-0.020
	(0.074)**	(0.018)*	(0.015)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081

Note: See note to Table 1 for details on variable definitions. Number of pairs refers to unique (non-repeated) sector combinations. All regressions control for dissimilarity in use of resources. Regressions further control for the following variables averaged across the sector pairs: first year of opening (top panel); entry share (bottom panel). Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses.

Table 6: Heterogeneous agglomeration – Technology and education

	(1)	(2)	(3)
	High-tech	Mixed-tech	Low-tech
Labor pooling (LP)	0.046	0.110	0.332
	(0.017)***	(0.019)***	(0.049)***
Input-output sharing (IO)	0.020	0.064	0.091
	(0.012)	(0.020)***	(0.045)**
Knowledge spillovers – IOM (KS)	0.053	0.031	0.039
	(0.024)**	(0.016)*	(0.041)
N of. Observations/Pairs	7140/595	24780/2065	20532/1711
	High-educ.	Mixed-educ.	Low-educ.
Labor pooling (LP)	0.046	0.167	0.391
	(0.023)*	(0.031)***	(0.061)***
Input-output sharing (IO)	0.007	0.066	0.123
	(0.013)	(0.029)**	(0.057)**
Knowledge spillovers – IOM (KS)	0.048	0.050	0.030
	(0.020)**	(0.021)**	(0.040)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081

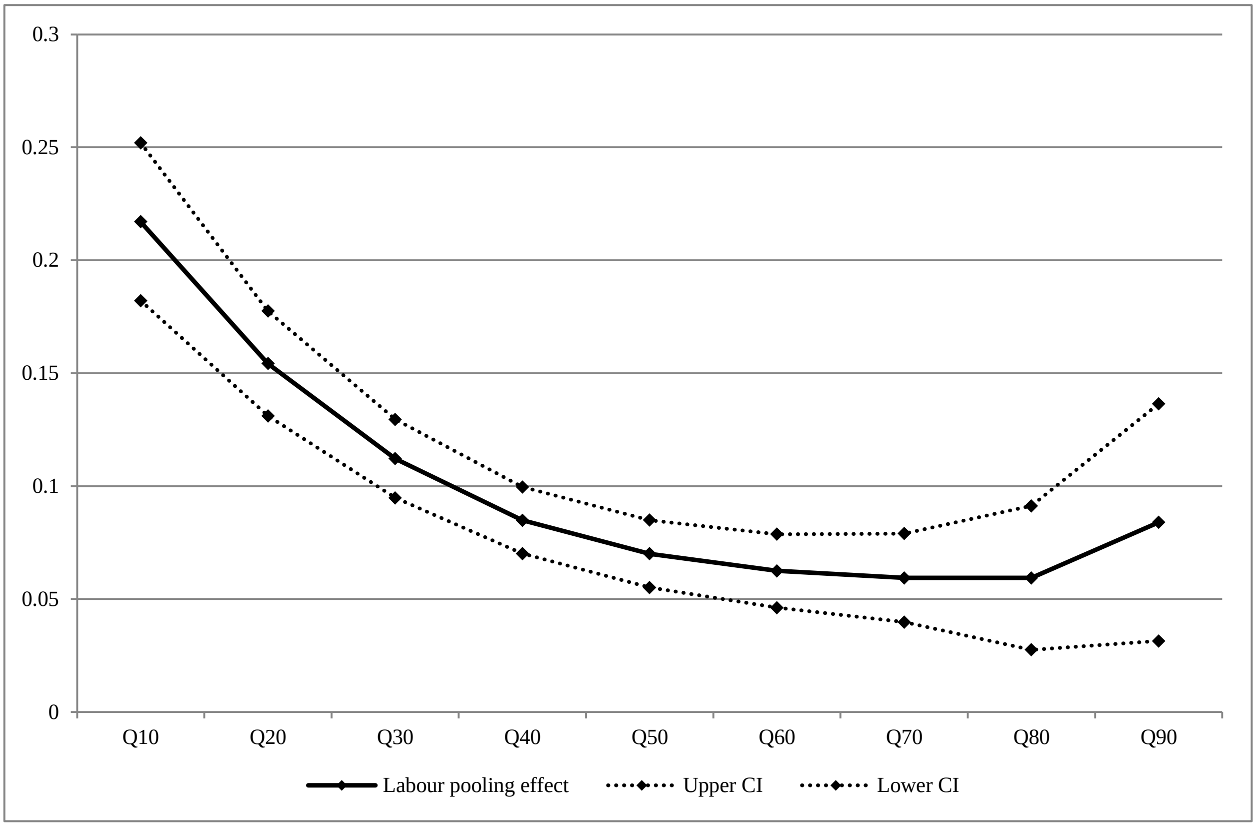
Note: See note to Table 1 for details on variable definitions. Number of pairs refers to unique (non-repeated) sector combinations. All regressions control for dissimilarity in use of resources. Regressions in the bottom panel further control for the share of college graduates averaged across the sector pairs. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses.

Table 7: Heterogeneous agglomeration– Organization

	(1)	(2)	(3)
	Small entrants	Mixed entrants	Large entrants
Labor pooling	0.113	0.181	0.223
	(0.027)***	(0.027)***	(0.056)***
Input-output sharing	0.193	0.068	0.082
	(0.112)*	(0.034)**	(0.039)**
Knowledge spillovers – IOM	0.051	0.028	-0.022
	(0.030)*	(0.016)*	(0.030)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081
	Small incumbents	Mixed incumbents	Large incumbents
Labor pooling	0.065	0.223	0.149
	(0.026)**	(0.031)***	(0.047)***
Input-output sharing	0.159	0.071	0.068
	(0.076)**	(0.035)**	(0.039)*
Knowledge spillovers – IOM	0.034	0.020	0.040
	(0.024)	(0.018)	(0.031)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081

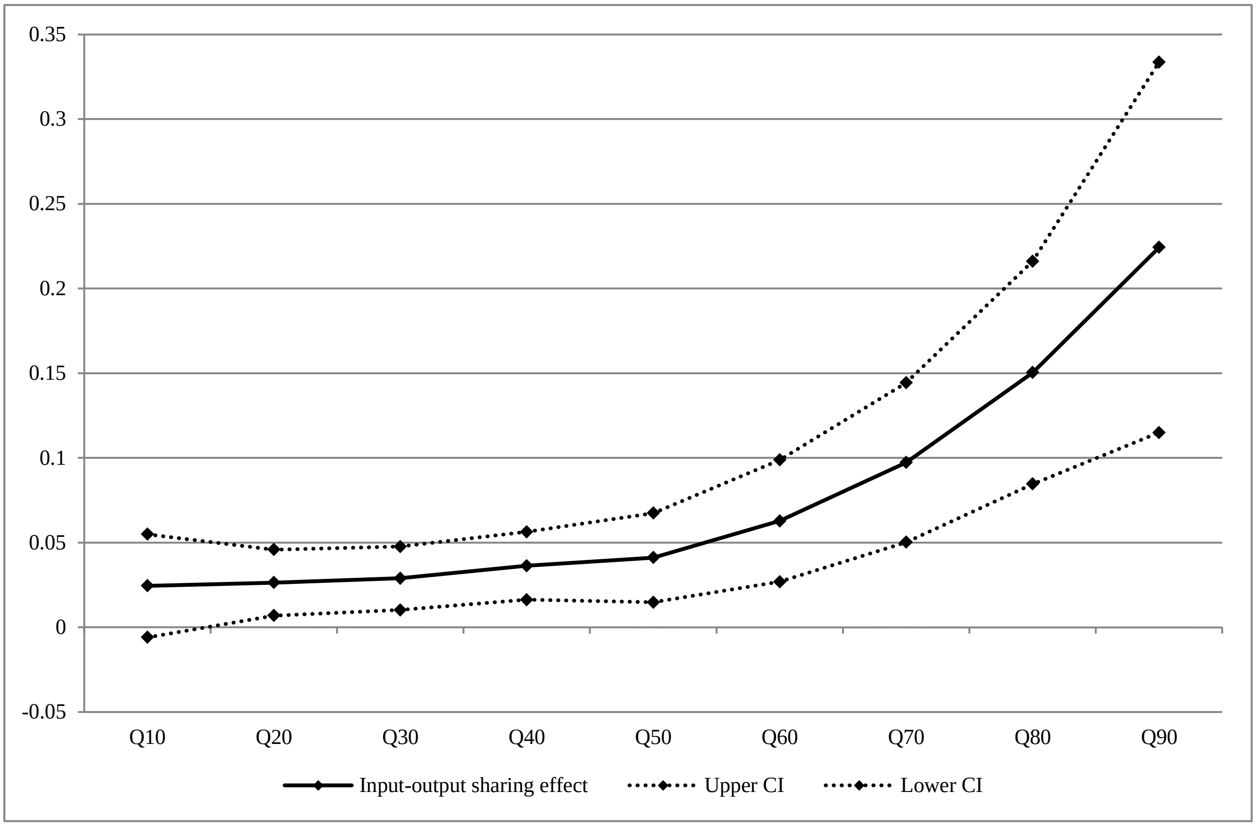
Note: See note to Table 1 for details on variable definitions. Number of pairs refers to unique (non-repeated) sector combinations. All regressions control for dissimilarity in use of resources. Regressions further control for the following variables averaged across the sector pairs: size of entrants (top panel); size of incumbents (bottom panel). Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses.

Figure 1a: The effect of Marshallian forces at different quantiles of γ^C: Labor pooling



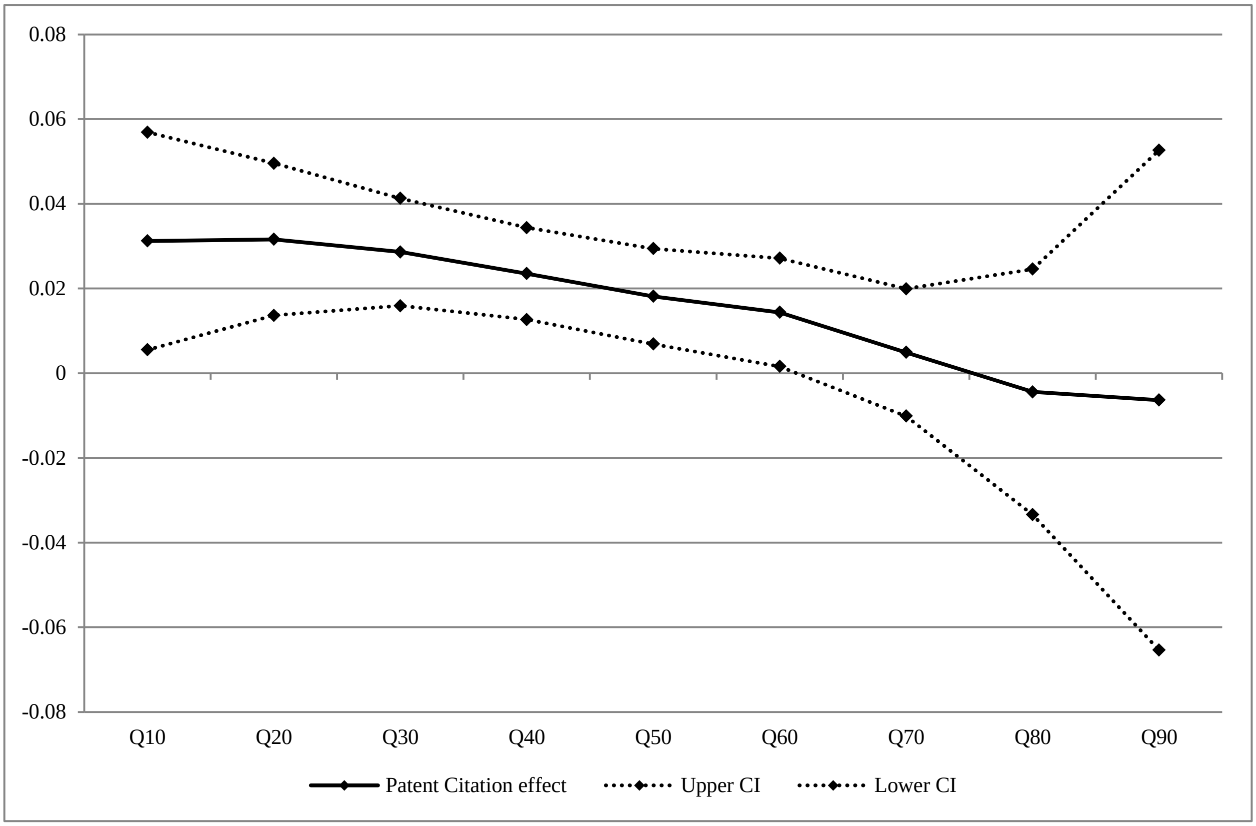
Note: See note to Table 1 for details on variable definitions. Variables are transformed to have unit standard deviation for interpretation. The figure plots regression coefficients and 95% confidence intervals from quantile regressions that simultaneously include all three Marshallian forces. Confidence intervals from bootstrapped standard errors clustered on industry pairs. All regressions control for dissimilarity in use of resources.

Figure 1b: The effect of Marshallian forces at different quantiles of γ^C: Input-output sharing



Note: See note to Table 1 for details on variable definitions. Variables are transformed to have unit standard deviation for interpretation. The figure plots regression coefficients and 95% confidence intervals from quantile regressions that simultaneously include all three Marshallian forces. Confidence intervals from bootstrapped standard errors clustered on industry pairs. All regressions control for dissimilarity in use of resources.

Figure 1c: The effect of Marshallian forces at different quantiles of γ^C: Knowledge Spillovers – IOM



Note: See note to Table 1 for details on variable definitions. Variables are transformed to have unit standard deviation for interpretation. The figure plots regression coefficients and 95% confidence intervals from quantile regressions that simultaneously include all three Marshallian forces. Confidence intervals from bootstrapped standard errors clustered on industry pairs. All regressions control for dissimilarity in use of resources. Knowledge spillovers measure is based on probabilistic mapping – Industry of manufacturing (Knowledge Spillovers – IOM).

Appendix tables

Table A1: Fifteen most co-agglomerated industry pairs – based on coagglomeration measure γ^C in 1997

Rank	Industry 1	Industry 2	Coagglomeration	1st TTWA	2nd TTWA	3rd TTWA
1	Ceramic goods other than construction (262)	Ceramic tiles and flags (263)	0.105	Stoke-on-Trent	Exeter	London
2	Knitted and crocheted fabrics (176)	Knitted and crocheted articles (177)	0.086	Leicester	Nottingham	Derby
3	Publishing (221)	Jewellery and related articles (362)	0.054	London	Birmingham	Sheffield
4	Spinning of textiles (171)	Textile weaving (172)	0.054	Bradford	Huddersfield	Leeds
5	Publishing (221)	Printing and reproduction of recorded media (222)	0.037	London	Manchester	Birmingham
6	Finishing of textiles (173)	Knitted and crocheted articles (177)	0.037	Leicester	Manchester	Nottingham
7	Finishing of textiles (173)	Knitted and crocheted fabrics (176)	0.035	Leicester	Nottingham	Manchester
8	Ceramic goods other than construction (262)	Construction products in baked clay (264)	0.033	Stoke-on-Trent	Crawley	Peterborough
9	Basic iron and steel and ferro-alloys (271)	Cutlery, tools and general hardware (286)	0.033	Sheffield	Birmingham	Wolverhampton
10	Basic iron and steel and ferro-alloys (271)	Other first processing of iron and steel (273)	0.033	Sheffield	Dudley	Swansea
11	Other first processing of iron and steel (273)	Forging, pressing, stamping and roll forming of metal (284)	0.031	Dudley	Birmingham	Sheffield
12	Tanning and dressing of leather (191)	Footwear (193)	0.031	Northampton	Hull	Glasgow
13	Knitted and crocheted articles (177)	Footwear (193)	0.030	Leicester	Northampton	Blackburn
14	Iron and Steel tubes (272)	Other first processing of iron and steel (273)	0.029	Dudley	Birmingham	Sheffield
15	Spinning of textiles (171)	Finishing of textiles (173)	0.028	Bradford	Manchester	Huddersfield

Note: See note to Table 1 for details on variable definitions.

Table A2: Characterizing the sectoral breakdown

Sector pair is:	N. of Obs./Pairs	Mean γ^C	Mean LP	Mean IO	Mean KS (IOM)	Mean of Cut-off Variable	Mean E-G Localization Index
New	12972/1081	0.000	0.203	0.006	0.007	1972	0.039
Mixed	26508/2209	0.000	0.234	0.008	0.016	1958	0.027
Old	12972/1081	0.000	0.277	0.014	0.022	1945	0.014
Dynamic	12972/1081	0.000	0.219	0.007	0.012	0.131	0.032
Mixed	26508/2209	0.000	0.234	0.008	0.015	0.106	0.027
Steady	12972/1081	0.000	0.262	0.012	0.020	0.082	0.021
High tech.	7140/595	0.000	0.412	0.014	0.029	--	0.009
Mix tech.	24780/2065	0.000	0.221	0.007	0.016	--	0.023
Low tech.	20532/1711	0.000	0.195	0.010	0.011	--	0.037
High education	12972/1081	0.000	0.328	0.009	0.022	0.148	0.015
Mix education	26508/2209	0.000	0.219	0.008	0.014	0.097	0.027
Low education	12972/1081	0.000	0.184	0.011	0.012	0.047	0.038
Small entrants	12972/1081	0.000	0.248	0.005	0.016	5.676	0.018
Mix entrants	26508/2209	0.000	0.233	0.010	0.014	10.08	0.027
Large entrants	12972/1081	0.000	0.234	0.012	0.018	14.48	0.035
Small incumbents	12972/1081	0.000	0.240	0.005	0.019	10.54	0.017
Mix incumbents	26508/2209	0.000	0.233	0.009	0.014	22.11	0.027
Large entrants	12972/1081	0.000	0.243	0.012	0.016	33.69	0.036

Note: Number of pairs refers to unique (non-repeated) sector combinations. High-tech and low-tech industries are categorized according to the OECD classification (1997). High-education and low-education industries are classified according to the share of college graduates above and below the median across all years (the median is 0.0783). Education level calculated using the UK LFS 1995-1999 data. New/old industry pairs consist of industries where the first year of opening is above/below the median across all years (the median is 1967). Dynamic/steady industry pairs consist of industries where the share of entrants is above/below the median across all years (the median is 0.100). Small/large entrants refer to industry pairs where the average size of entrants is below/above the median across all years (the median is 8.59). Small/large incumbents refer to industry pairs where the average size of incumbents is below/above the median across all years (the median is 18.95). Mixed pairs consist of one old/big entrants/big incumbents/steady industry and one new/small entrants/small incumbents/dynamic industry. Variables in the penultimate column refer to the cut-off variables (first year of opening, entry share, size of entrants and size of incumbents) averaged across the sector pairs. Mean Ellison-Glaeser localization index across all sectors: 0.027 (std. dev.: 0.048).

Table A3: Further robustness checks and extensions

	(1)	(2)	(3)	(4)	(5)
	Staggered Marshallian forces	γ^C excluding London	Control for popul. density	Control for empl. density	Control for Herfind. Index
Labor pooling	0.160	0.139	0.142	0.158	0.165
	(0.019)***	(0.022)***	(0.018)***	(0.020)***	(0.020)***
Output sharing	0.081	0.119	0.091	0.088	0.083
	(0.026)***	(0.033)***	(0.026)***	(0.026)***	(0.026)***
Knowledge spillovers – IOM	0.023	0.017	0.026	0.023	0.026
	(0.014)*	(0.014)	(0.014)*	(0.014)*	(0.014)*

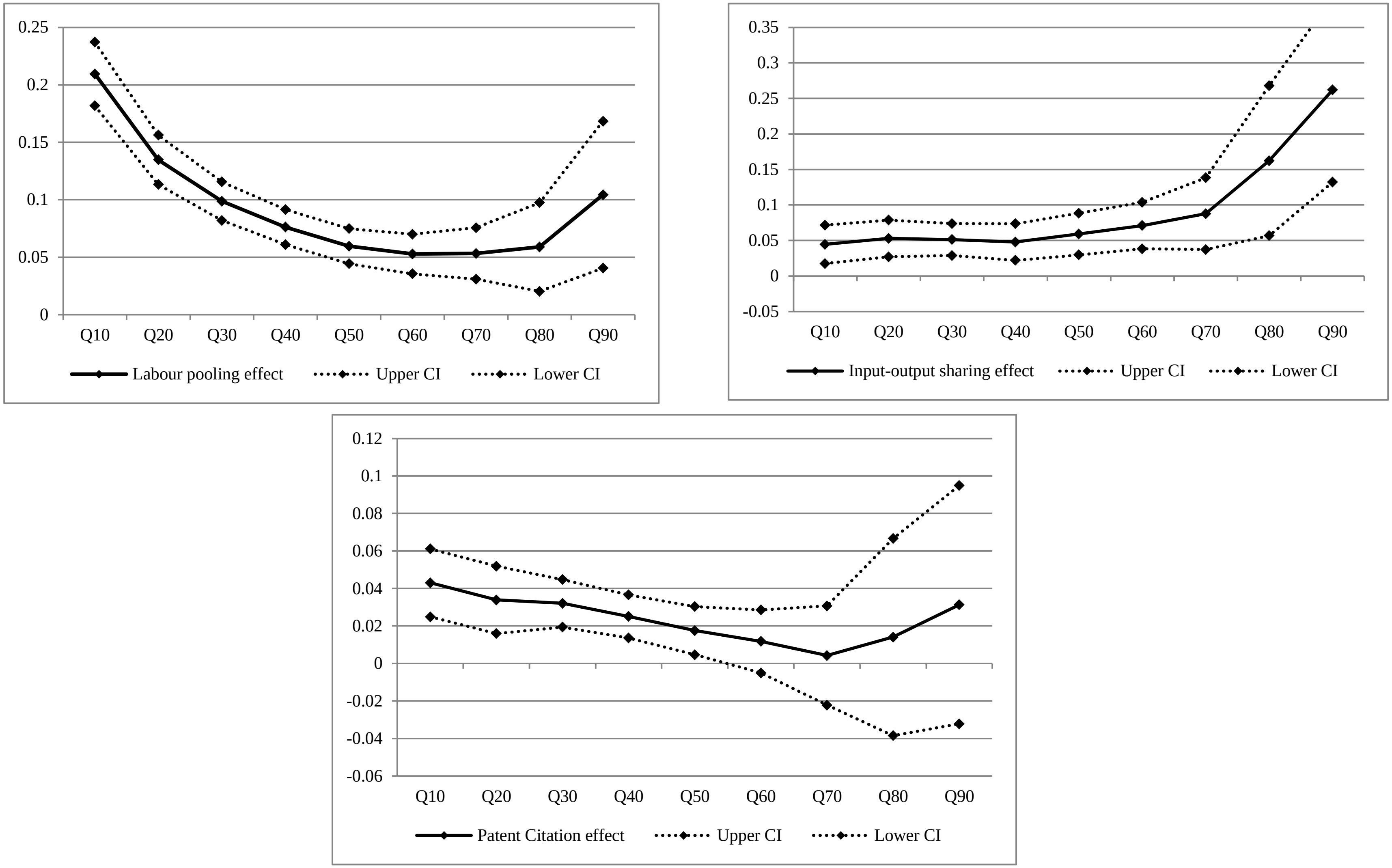
Note: See note to Table 1 for details on variable definitions. All regressions control for dissimilarity in use of resources. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses. Knowledge spillovers measure is based on probabilistic mapping – Industry of manufacturing (knowledge spillovers – IOM). Column (1) considers γ^C for years from 2000 and Marshallian forces calculated up to 1999. Column (2) excludes London from the calculations of γ^C. Column (3) controls for the average (un-weighted) population density of the TTWAs in which the two sectors are operating, averaged within the pair. Column (4) controls for the average (un-weighted) employment density of the TTWAs in which the two sectors are operating, average within the pair. Column (5) controls for the Herfindahl index of the two sectors, averaged within the pair.

Table A4: The heterogeneous relationship between coagglomeration γ^C and Marshallian forces – Regional level of aggregation

Panel A: Adaptation	*Dynamic*	*Mixed*	*Steady*
Labor pooling (LP)	0.134	0.122	0.161
	(0.048)***	(0.025)***	(0.028)***
Input-output sharing (IO)	0.184	0.127	0.059
	(0.106)*	(0.038)***	(0.023)**
Knowledge spillovers – IOM (KS)	0.209	0.030	-0.018
	(0.071)***	(0.016)*	(0.015)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081
Panel B: Technology	*High-tech*	*Mixed-tech*	*Low-tech*
Labor pooling (LP)	0.096	0.127	0.197
	(0.024)***	(0.021)***	(0.040)***
Input-output sharing (IO)	0.009	0.096	0.135
	(0.016)	(0.028)***	(0.051)**
Knowledge spillovers – IOM (KS)	0.074	0.022	0.039
	(0.023)***	(0.016)	(0.037)
N of. Observations/Pairs	7140/595	24780/2065	20532/1711
Panel C: Organization	*Small incumbents*	*Mixed incumbents*	*Large incumbents*
Labor pooling	0.102	0.161	0.162
	(0.021)**	(0.025)***	(0.049)***
Input-output sharing	0.138	0.127	0.069
	(0.042)***	(0.033)***	(0.054)
Knowledge spillovers – IOM	0.062	0.007	0.035
	(0.022)***	(0.018)	(0.028)
N of. Observations/Pairs	12972/1081	26508/2209	12972/1081

Note: See note to Table 1 for details on variable definitions. Number of pairs refers to unique (non-repeated) sector combinations. All regressions control for dissimilarity in use of resources. Regressions further control for the following variables averaged across the sector pairs: entry share (top panel); size of entrants (top panel). Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses.

Figure A1: The effect of Marshallian forces at difference quantiles of γ^{C} – Regional level of aggregation



Note: See note to Table 1 for details on variable definitions. Variables are transformed to have unit standard deviation for interpretation. The figure plots regression coefficients and 95% confidence intervals from quantile regressions that simultaneously include all three Marshallian forces. Confidence intervals from bootstrapped standard errors clustered on industry pairs. All regressions control for dissimilarity in use of resources.

Web Appendix: Data Construction –Not for Publication – This will posted on the authors' webpage

The Business Structure Database (BSD)

Our measures of coagglomeration of UK manufacturing sectors are constructed aggregating micro-level data from the Business Structure Database (BSD) covering the period 1997 to 2008. The data is an annual snapshot (taken in April at the closing of the fiscal year) of the Inter-Departmental Business Register (IDBR), which consists of constantly-updated administrative data collected for taxation purposes. Any business liable for value-added taxation (VAT) and/or with at least one employee registered for tax collection appears on the IDBR. Estimates produced by the Office for National Statistics (ONS) in 2004 show that the businesses listed on the IDBR account for almost 99 per cent of economic activity in the UK.

The data are structured into enterprises and local units. An enterprise is the overall business organization. The local unit can be thought of as a plant or establishment. In the majority of cases (70 per cent), enterprises only have one local unit, while the remaining 30 per cent of the cases represent enterprises with multiple local units. In our work, we make use of data at the local unit level including plants belonging both to single- and multi-plant enterprises and located in England, Wales and Scotland. We neglect Northern Ireland because of poor data coverage.

The initial raw data includes approximately three million local units every year. However, we carry out a series of checks and drop a number of units. First, we investigate the consistency of opening and closing dates of BSD units with their actual existence in the dataset and drop a number of anomalous cases where we identify establishments opening/closing in a specific year, disappearing/reappearing in a subsequent year only to open/close again in a subsequent wave. Stated differently, we only count firms' birth and death once. Second, we check the consistency of units' postcodes and sectors of activity over the years, and drop cases with missing or anomalous information. [33] For example, when we observe two or more plants operating in the same 3-digit industry, sharing the same postcode and being part of the same enterprise, we believe this being a reporting error and drop them. Similarly, we observe a number of same-postcode same-three-digit industry combinations representing anomalous concentration of identical activities at a single address. We believe this is another coding error and drop the plants that belong to the top 5% of the distribution of the number of plants sharing same three-digit industry and the same postcode. Finally, we drop active units with zero employment since this figure includes the owners/managers of the establishment, so it cannot be zero for an active unit, as well as units with an unusually large size (i.e., total employment above the 99th percentile of the distribution for each three-digit industry sector). After applying these restrictions, our dataset still comprises of more than two million plants annually over 12 years (1997-2008).

[33] A UK postcode usually corresponds to a very limited number of addresses or a single large delivery point. While it might not always be a geographically accurate description of where a property is located, it is generally a good approximation. For instance, a building which contains several businesses, but only one external door will only have the external door listed as a delivery point. This example shows that UK postcodes are geographically accurate up to the level of a front door in a particular street.

In terms of industrial classification, we focus on manufacturing and adopt the three-digit Standard Industry Classification (SIC) 1992. We also apply a number of restrictions and re-combine a number of sectors to avoid having a limited or erratic evolution in the number of plants and employment during the sample period. Specifically, we exclude Tobacco (SIC160) because the number of plants in this sector tends to be small throughout the sample period (e.g. 43 in 1997). In addition, we combine Leather (SIC181) and Fur Clothes (SIC183) with Other Wearing Apparel (SIC182) to avoid small sample size problems in SIC181 and SIC183. For similar reasons, we also combine the following industries: Manufacture of Vegetal and Animal Oils and Fats (SIC154) with Other Food Products (SIC158); Reproduction of Recorded Media (SIC223) with Printing (SIC222); Coke Oven Products (SIC231) and Processing of Nuclear Fuel (SIC233) with Refined Petroleum Products (SIC232); Man-made Fabrics (SIC246) with Other Chemical Products (SIC247); Articles of Concrete, Plaster and Cement (SIC266) with Manufacture of Cement, Lime and Plaster (SIC265). Our final sample consists of 94 manufacturing 3-digit sectors for a total of 4,371 unique pairwise correlations a year for twelve years (1997-2008).

In terms of geography, our unit of aggregation is the Travel-to-Work Area (TTWA). These are entities constructed to guarantee that at least 75% of the resident population works in the area and that 75% of the people working in the area are resident there. TTWAs were devised to delineate areas that can be considered as self-contained labor markets and economically relevant aggregates. In 2007, there were 243 TTWAs within the United Kingdom. In most of our work, we mainly focus on 84 urban TTWAs with population in excess of 100,000. The reason why we focus on TTWAs with more than 100,000 inhabitants is that they cover urban areas where most of the productive activities take place. Rural areas in the UK refer to fairly sparsely populated areas such as the Scottish Highlands, the Welsh Mountains and the Peak and Lake Districts. Very few productive activities are concentrated in these areas – with the exception of tourism and related services. These industries are not covered by our analysis which focuses on solely manufacturing. However, in some robustness checks, we extend our analysis to include all TTWAs – urban and rural, irrespective of their population. The correlation between the coagglomeration metric γ^C that we use in the paper and the γ^C measure for all TTWAs (urban and rural) is very high, at 0.993. This backs our intuition that urban areas drive manufacturing coagglomeration patterns. Note that we aggregate the individual areas of Clacton, Colchester, Lincoln, Grantham, Torquay, and Paignton-Totes into the following urban TTWAs: *(1)* Clacton & Colchester; *(2)* Lincoln & Grantham; and *(3)* Torquay & Paignton-Totes. Even before the aggregation, the areas of Colchester, Lincoln and Torquay each had a population above the 100,000 threshold.

UK Labour Force Survey (LFS)

The UK Labour Force Survey (LFS) is a quarterly representative survey of households living at private addresses in the UK and is conducted by the Office for National Statistics (ONS) to collect information about individuals' labor market experiences. In our analysis, we use the years between 1995 and 1999 which allow for a consistent coding of the industrial and occupational classification of workers' jobs.

Each quarter of the LFS contains between 64,000 (earlier years) and 52,000 (later years) households, equivalent to about 120,000-150,000 individuals. In our analysis, we focus on 16-59 aged women and 16-64 aged men, and on individuals either working as employees or as self-employed. Excluding self-employed individuals does not affect our analysis.

In order to assign each individual to a TTWA, we retain workers living in England, Scotland and Wales (LFS data for Northern Ireland has poor coverage), and with a valid geographical identifier, namely the ward of residence (roughly equivalent to a US census tract). Additionally, we select individuals with non-missing information on: *(i)* gender and age; *(ii)* educational qualifications; *(iii)* industry and occupation. We exclude people working for the armed forces.

These restrictions leave us with a set of approximately 200,000 individuals each year for a total of 1.03m, of which 820,000 and 210,000 live in urban and rural areas, respectively. Next, we select individuals living in urban areas and working in manufacturing only. The final sample consists of about 35,000 workers a year for a total of 166,000 individuals. We use the UK Standard Industrial Classification (SIC) 1992 and the UK Standard Occupational Classification (SOC) 1990 at the three-digit level for these individuals' jobs to construct a proxy for the extent of labor pooling occurring between manufacturing sectors.

UK Input-Output Tables

To capture the flow of goods between industry pairs, we use the ONS Input-Output Analytical Tables for 1995 to 1999. For each industry, we calculate the shares of inputs bought from/sold to other industries as a fraction of the total intermediate inputs/outputs. Note that we exclude direct sales to consumers.

The sector classification in the I-O Tables is more aggregated than the three-digit SIC classification we use and only includes 77 manufacturing industries. In order to assign I-O shares to a SIC three-digit sector belonging to the same I-O sector, we use an apportioning procedure based on their employment share within the group averaged over 1995-1999. These shares are obtained using the relevant years of the BSD (our main dataset).

The EPO-CESPRI Dataset

The main data source for our analysis of patent citation flows is the EPO-CESPRI data provided by Bocconi University. This database provides cleaned and consistently coded information extracted from the European Patent Office (EPO) data for the period 1977 and 2009. Approximately 144,000 patents were filed by 160,000 UK inventors (multiple-investors can be recorded for each patent). These generate a stream of more than 77,000 citations of UK patents over the observed time-window.

In order to construct knowledge spillover measures we impose a number of restrictions. First, we exclude self-citations from the same inventor or the applying company at which he/she is based. Second, we exclude citing patents filed after 2000 and before 1981, and cited patents filed after 1997. The aim of these restrictions is twofold: *(a)* we want to guarantee that on average cited patents are at least three years older than citing ones; *(b)* we want to guarantee that our knowledge-spillover measures are constructed for the initial years of our sample (i.e. up to 2000) so that they are measured at a similar time as the labor-pooling and input-output sharing metrics. Expanding the sample to include all years does not affect the results.

It should be noted that while the US Patent and Trademark Office (USPTO) requires patent applicants to declare all relevant references and citations, the EPO does not apply this rule and all citations come directly from the patent examiners. As a result, the average number of EPO patent citations is much smaller than the corresponding figure for USPTO patents, and EPO numbers do not suffer from UPSTO-type "citation inflation" (see Hall et al. 2000). According to Breschi and Lissoni (2004), USPTO patents cited approximately 13 other patents and received on average 10.2 citations. The corresponding numbers for EPO patents are much lower at 4 and 2.8, respectively.

Patents in the EPO dataset (as in any other patent database) are categorized using technological classes rather than a standard industrial classification. To create a mapping industrial sectors and technological classification, we follow the literature and adopt two approaches: *(1)* a probabilistic mapping based on the Industry of Manufacture (IOM); and *(2)* an alternative probabilistic mapping based on the Sector of Use (SOU). These are based on correspondences developed by Silverman (2002) who studied approximately 150,000 patents filed at the Canadian Patent Office between 1990 and 1993. More information is available from Silverman's website: http://www-2.rotman.utoronto.ca/~silverman/ipcsic/documentation_IPC-SIC_concordance.htm

The OECD Technology Classification

Based on the intensity of both direct R&D (i.e. R&D expenditure) and indirect R&D (i.e. embodied technology flows) in the output of manufacturing sectors across 10 OECD countries over the period 1980 to 1996, the OECD classifies as high tech or medium-high tech the following manufacturing industries: SIC241 "Manufacturing of basic chemicals" to SIC246 "Manufacturing of other chemicals & man-made fibres"; SIC291 "Manufacturing of other machinery for production/use of mechanical power N.E.C." to SIC297 "Manufacturing of domestic appliances"; SIC300 "Manufacturing of office machinery & computers"; SIC311 "Manufacturing of electric motors, generators & transformers" to SIC316 "Manufacturing of electrical equipment N.E.C."; SIC231 "Manufacturing of electronic valves, tubes & electronic components" to SIC323 "Manufacturing of TV/radio receivers & sound/video recording/reproducing"; SIC331 "Manufacturing of medical, surgical & orthopedic equipment" to SIC335 "Manufacture of watches & clocks; SIC341 "Manufacturing of motor vehicles" to SIC343 "Manufacturing of parts & accessories for vehicles/engines"; and SIC352 "Manufacturing of railway/tramway locomotives/rolling stock" to SIC355 "Manufacturing of other transport equipment N.E.C.". See OECD (1997) for more details.

US Data for Instrumental Variables

In order to address potential endogeneity issues, we follow Ellison et al. (2010) and instrument our UK-based proxies for the three Marshallian forces using almost identical measures obtained from US data.

Starting with labor pooling, we create a measure of the similarity in the occupational inputs of two industries using the National Industrial-Occupation Employment Matrix (NIOEM) published by the Bureau of Labor Statistics. Following the approach we have taken for the UK data, we construct the shares of different types of workers used in each manufacturing sector and then correlate the percentage of different types of occupations across industry pairs to obtain a proxy for labor sharing. In order to link this proxy to our data, we map US NIOEM industry codes to UK SIC codes. Since the US manufacturing classification is less detailed than the one that we adopt (79 vs. 94 sectors), we attribute the same US industry-occupation shares to multiple UK sectors. Note also that we construct the US labor correlation measure using all available data spanning the period 1983 to 1998. Restricting the calculations of this instrument to the period 1995-1998 does not affect our IV results.

We construct an instrument for input-output sharing following a similar approach. To begin with, we use the 1987 Benchmark Input-Output Accounts published by the Bureau of Economic Analysis (BEA) to measure the flows of intermediate inputs exchanged between US industries at the same level of aggregation as used in Ellison et al. (2010) and map these values from the 140 US manufacturing sectors to the 94 UK industries. In our regression analysis we focus on the maximum between the inputs and outputs that two industries are sharing irrespective of the direction of the flow (given that our data treats industry pairs symmetrically). Consistently, we use US data to construct a proxy for the maximum of the input-output linkages between industries and use this as an instrument.

Finally, we construct our instrument for knowledge spillovers using the NBER Patent Data initially assembled by Hall et al. (2001). The data cover patents granted by the US Patent and Trademark Office (USPTO) between 1975 and 1999 and record citation flows across patents. Following our main approach, we use a probabilistic mapping based on the industry of manufacture (IOM) to map technology to industrial classes and convert citations across US sectors to our UK classification based on 94 industries. Note that this instrument is different from the one adopted by Ellison et al. (2010) who used UK patents registered at the USPTO to instrument for US patents registered at the same office. Conversely, we use information coming from the USPTO about flows of citations among US patents to instrument for citations among UK patents registered at the European Patent Office (EPO). Part of the mechanical problems discussed by Ellison et al. (2010) in relation to this instrument is thus by-passed.

References

Breschi, S. and F. Lissoni (2004), "Knowledge networks from patent data: methodological issues and research targets", CESPRI Working Paper 150.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010), "What causes industry agglomeration? Evidence from coagglomeration patterns", *American Economic Review* 100, 1195-1213.

Hall, B., A. Jaffe and M. Trajtenberg (2000), "Market Value and Patent Citations: A First Look," NBER Working Paper 7441.

Hall, B., A. Jaffe and M. Trajtenberg (2001), "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tool", NBER Working Paper 8498.

OECD (1997), "Revision of High-Technology Sector and Product Classification", STI Working Paper 1997/2.

Web Appendix: Additional Regressions – Not for Publication – These will posted on the authors' webpage

Table W1: Additional regressions of coagglomeration measure γ^C on Marshallian forces

	KS measured as Industry of Manufacture (IOM)				*KS measured as Industry of Manufacture (SOU)*				
	(1)	(2)	(3)	(4)					
Dependent variable/	γ^C	γ^C	γ^C	γ^C	γ^C	γ^C	γ^C	γ^C	γ^C
Timing is:	1997-2008	1997	2002	2008	1997-2008.	1997-2008	1997	2002	2008
Labor pooling (LP)	0.166	0.188	0.166	0.146	0.152	0.154	0.170	0.153	0.138
	(0.020)***	(0.024)***	(0.022)***	(0.020)***	(0.019)***	(0.019)***	(0.023)***	(0.021)***	(0.019)***
Input-output sharing (IO)	--	0.083	0.099	0.071	0.071	--	0.067	0.088	0.063
		(0.026)***	(0.028)***	(0.025)***	(0.024)***		(0.024)***	(0.027)***	(0.024)***
Knowledge spill. – IOM (KS)	0.026	0.028	0.029	0.022					
	(0.013)*	(0.014)**	(0.015)*	(0.013)					
Knowledge spill. – SOU (KS)					0.075	0.077	0.101	0.080	0.057
					(0.023)***	(0.022)***	(0.026)***	(0.025)***	(0.022)***
Input sharing	0.057	--	--	--		0.049			
	(0.028)**					(0.027)*			
Output sharing	0.025	--	--	--		0.021			
	(0.031)					(0.030)			

Note: See note to Table 1 for details on variable definitions and samples. All regressions include all Marshallian forces at the same time and control for dissimilarity in use of resources. Variables are standardized to have zero mean and unit standard deviation. Robust standard errors clustered on industry pairs are reported in parentheses.