My precious!

The location and diffusion of scientific research: evidence from the Synchrotron Diamond Light Source

Christian Helmers[§,‡] Henry Overman[‡]

§ Universidad Carlos III, Madrid

‡ SERC LSE

October 19, 2012

Overview

Overview

- Analyze impact of GBP380 million basic scientific research facility on geographic distribution of research

Overview

- Analyze impact of GBP380 million basic scientific research facility on geographic distribution of research
- **Diamond Light Source** (3rd generation synchrotron) single largest investment in research in UK

Overview

- Analyze impact of GBP380 million basic scientific research facility on geographic distribution of research
- **Diamond Light Source** (3rd generation synchrotron) single largest investment in research in UK

- Impact of **location choice** on distribution of research

Overview

- Analyze impact of GBP380 million basic scientific research facility on geographic distribution of research
- **Diamond Light Source** (3rd generation synchrotron) single largest investment in research in UK

- Impact of **location choice** on distribution of research
- Direct impact on clustering of research that uses facility

Overview

- Analyze impact of GBP380 million basic scientific research facility on geographic distribution of research
- **Diamond Light Source** (3rd generation synchrotron) single largest investment in research in UK

- Impact of **location choice** on distribution of research
- Direct impact on clustering of research that uses facility
- Indirect impact on location of related research

Main Questions

Main Questions

- Impact of major 'lumpy' infrastructure investments

Main Questions

- Impact of major 'lumpy' infrastructure investments
- "Sharing of indivisibilities" one of the crucial determinants of agglomeration

Main Questions

- Impact of major 'lumpy' infrastructure investments
- "Sharing of indivisibilities" one of the crucial determinants of agglomeration

- Role of knowledge spillovers

Main Questions

- Impact of major 'lumpy' infrastructure investments
- "Sharing of indivisibilities" one of the crucial determinants of agglomeration

- Role of knowledge spillovers
- As input into research

Main Questions

- Impact of major 'lumpy' infrastructure investments
  - "Sharing of indivisibilities" one of the crucial determinants of agglomeration
- Role of knowledge spillovers
  - As input into research
  - As determinant of location of research

Main Questions

- Impact of major 'lumpy' infrastructure investments
  - "Sharing of indivisibilities" one of the crucial determinants of agglomeration
- Role of knowledge spillovers
  - As input into research
  - As determinant of location of research

$\Rightarrow$ Importance of agglomeration externalities produced by indivisible scientific research facilities for science and innovation

Identification

Identification

- Test whether establishment of Diamond in Didcot has resulted in concentration of research and innovation in geographical proximity beyond what would have happened, had Diamond been sited elsewhere

Identification

- Test whether establishment of Diamond in Didcot has resulted in concentration of research and innovation in geographical proximity beyond what would have happened, had Diamond been sited elsewhere
- Main challenge in identifying **causal** link from Diamond to geographical distribution is potential **endogeneity of Diamond's location**:
  "*[Didcot] is a thriving hub of scientific research [...]. Diamond is surrounded by a number of scientific research facilities making the site a centre of excellence in terms of tools and expertise and therefore the ideal location*."

Identification

- Test whether establishment of Diamond in Didcot has resulted in concentration of research and innovation in geographical proximity beyond what would have happened, had Diamond been sited elsewhere
- Main challenge in identifying **causal** link from Diamond to geographical distribution is potential **endogeneity of Diamond's location**:
  "*[Didcot] is a thriving hub of scientific research [...]. Diamond is surrounded by a number of scientific research facilities making the site a centre of excellence in terms of tools and expertise and therefore the ideal location*."
- If related research increasingly clustered in existing centers regardless of Diamond (government anticipated this) then wrongly attribute affect to Diamond

Identification

Identification

- Use diff-in-diff identification strategy

Identification

- Use diff-in-diff identification strategy
- Exploit **runner up location** at Daresbury (Manchester) to address endogeneity

Identification

- Use diff-in-diff identification strategy
- Exploit **runner up location** at Daresbury (Manchester) to address endogeneity
- *Assumption:* conditional on observable as well as unobservable location-specific characteristics, in the absence of Diamond, **changes** in pattern would have been same around Diamond and Daresbury

Empirical approach

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)
- Research output:

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)
- Research output:
  1. Scientific publications

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)
- Research output:
  1. Scientific publications
  2. Patents

Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)
- Research output:
  1. Scientific publications
  2. Patents
- Focus on UK: *whether one flies from Tokyo to Daresbury or from Tokyo to Oxford is irrelevant.* (Hon. Stunell, March 2000)

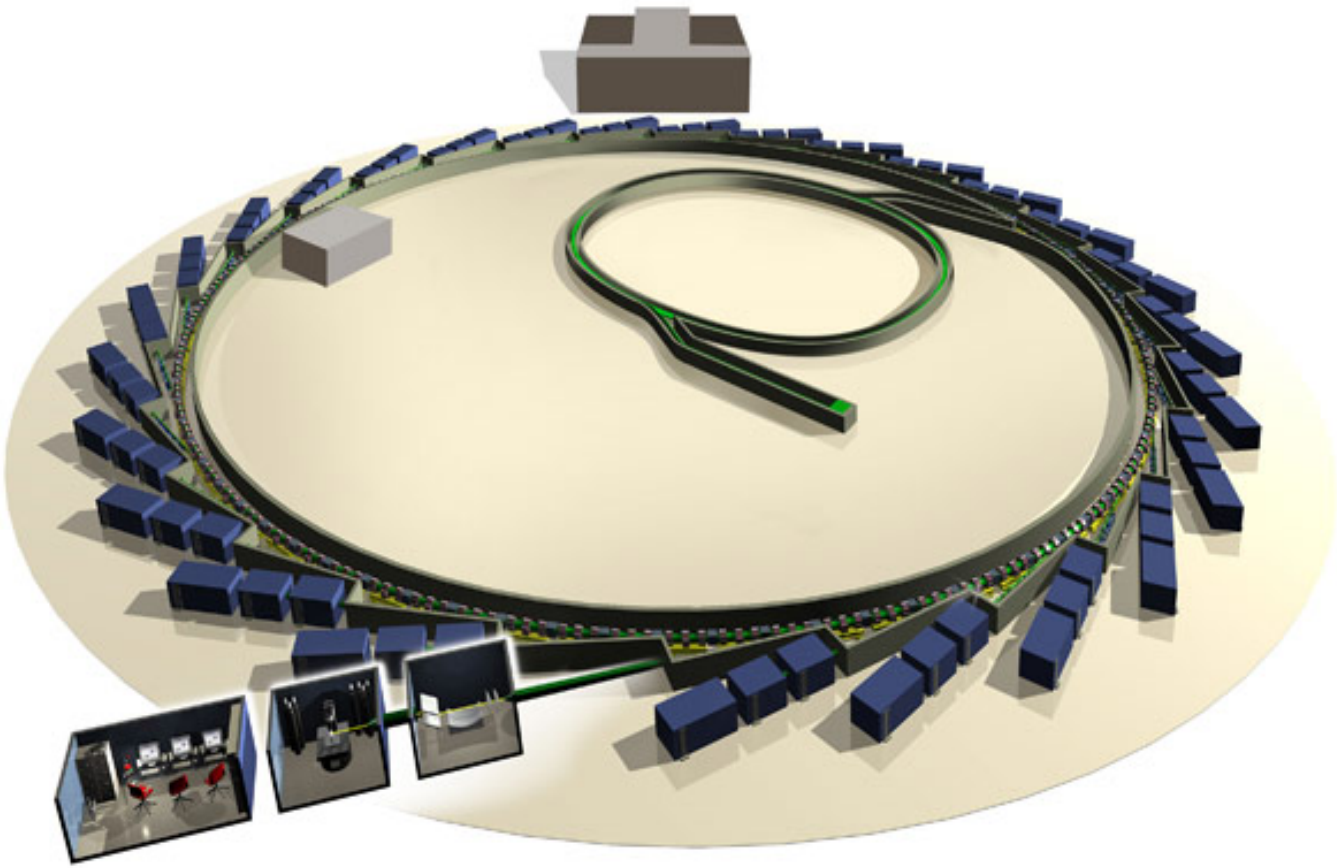
Empirical approach

- Distance to Diamond of scientists/inventors (in geographical area)
- Hard sciences produce **codifiable** knowledge (e.g. structural biology, physics, chemistry, materials science etc.)
- Research likely 'cutting edge' (i.e., publishable and patentable)
- Research output:
  1. Scientific publications
  2. Patents
- Focus on UK: *whether one flies from Tokyo to Daresbury or from Tokyo to Oxford is irrelevant.* (Hon. Stunell, March 2000)
- Does it matter where infrastructure is sited **within** the UK?

The Diamond Synchrotron

The Diamond Synchrotron

- What is Diamond used for?



The Diamond Synchrotron

The Diamond Synchrotron

- What is Diamond used for?

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)
  - Expensive GBP380 million (86% Government, 14% Wellcome trust)

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)
  - Expensive GBP380 million (86% Government, 14% Wellcome trust)
- Phase 1 began operation January 2007

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)
  - Expensive GBP380 million (86% Government, 14% Wellcome trust)
- Phase 1 began operation January 2007
- "Beam time" free to academic users

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)
  - Expensive GBP380 million (86% Government, 14% Wellcome trust)
- Phase 1 began operation January 2007
- "Beam time" free to academic users
- 3rd generation produce more intense and brighter synchrotron light than 2nd generation

The Diamond Synchrotron

- What is Diamond used for?
- Synchrotron facilities are circular particle accelerators that produce beams of x-rays, infrared and ultraviolet light
  - Used for looking at very small things
  - Big (561m storage ring, floor area 45,500m^2)
  - Expensive GBP380 million (86% Government, 14% Wellcome trust)
- Phase 1 began operation January 2007
- "Beam time" free to academic users
- 3rd generation produce more intense and brighter synchrotron light than 2nd generation
- Novel research in **hard sciences**

The Diamond Synchrotron

The Controversy

The Controversy

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)
- Wellcome pushed for Harwell (Didcot/Oxford)

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)
- Wellcome pushed for Harwell (Didcot/Oxford)
- "Greater scientific benefits would result from a location close to the existing neutron source and to MRC units and the University of Oxford"

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)
- Wellcome pushed for Harwell (Didcot/Oxford)
- "Greater scientific benefits would result from a location close to the existing neutron source and to MRC units and the University of Oxford"
- Daresbury supporters claimed Wellcome overstating the benefits of proximity

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)
- Wellcome pushed for Harwell (Didcot/Oxford)
- "Greater scientific benefits would result from a location close to the existing neutron source and to MRC units and the University of Oxford"
- Daresbury supporters claimed Wellcome overstating the benefits of proximity
- High-level dispute (e.g. parliamentary debate), broad public attention (e.g. BBC, Financial Times, the Times Higher Education, The Guardian, Nature)

The Controversy

- Initially government intended to site in Daresbury (Manchester) next to existing 2nd generation synchrotron (distance to Diamond approx. 215km – opened 1981)
- Wellcome pushed for Harwell (Didcot/Oxford)
- "Greater scientific benefits would result from a location close to the existing neutron source and to MRC units and the University of Oxford"
- Daresbury supporters claimed Wellcome overstating the benefits of proximity
- High-level dispute (e.g. parliamentary debate), broad public attention (e.g. BBC, Financial Times, the Times Higher Education, The Guardian, Nature)
- Government went with Harwell in March 2000

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise
  - 3 ring dummies: 0-25km, 25km-125km, 125km-175km

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise
  - 3 ring dummies: 0-25km, 25km-125km, 125km-175km
  - >175km from Diamond omitted category

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise
  - 3 ring dummies: 0-25km, 25km-125km, 125km-175km
  - >175km from Diamond omitted category
- $I(t \geq 2007)$: indicator variable $= 1$ from 2007 onwards

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise
  - 3 ring dummies: 0-25km, 25km-125km, 125km-175km
  - >175km from Diamond omitted category
- $I(t \geq 2007)$: indicator variable $= 1$ from 2007 onwards
- D_t a dummy variable taking value one if year is equal to t

Estimate following model:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at} \quad (1)$$

- c_{at}: count of articles/patents in area a at time t
- D^r_{DI}: set of R 'ring' dummies which $= 1$ if area within given distance of **Diamond** & $= 0$ otherwise
  - 3 ring dummies: 0-25km, 25km-125km, 125km-175km
  - >175km from Diamond omitted category
- $I(t \geq 2007)$: indicator variable $= 1$ from 2007 onwards
- D_t a dummy variable taking value one if year is equal to t
- α_a: area FE (if included, $\sum_R D^r_{DI}$ drops out)

Use 'runner-up' location:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \sum_R D^r_{DA} + \sum_R D^r_{DA} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_a t \times I(t \geq 2007) + \epsilon_{at} \quad (2)$$

Use 'runner-up' location:

$$c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} + \sum_R D^r_{DI} \times I(t \geq 2007) + \\ + \sum_R D^r_{DA} + \sum_R D^r_{DA} \times I(t \geq 2007) + \\ + \beta_1 X_{at} + \beta_2 X_a t \times I(t \geq 2007) + \epsilon_{at} \quad (2)$$

- D^r_{DA}: set of R 'ring' dummies which $= 1$ if area within given distance of **Daresbury** & $= 0$ otherwise

Use third location (Newcastle upon Tyne – Institute for Cell and Molecular Biosciences):

$$
\begin{aligned}
c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} \times I(t \geq 2007) + \\
+ \sum_R D^r_{DA} \times I(t \geq 2007) + \sum_R D^r_{NT} \times I(t \geq 2007) + \\
+ \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at}
\end{aligned} \tag{3}
$$

Use third location (Newcastle upon Tyne – Institute for Cell and Molecular Biosciences):

$$
\begin{aligned}
c_{at} = \alpha_a + \sum_t D_t + \sum_R D^r_{DI} \times I(t \geq 2007) + \\
+ \sum_R D^r_{DA} \times I(t \geq 2007) + \sum_R D^r_{NT} \times I(t \geq 2007) + \\
+ \beta_1 X_{at} + \beta_2 X_{at} \times I(t \geq 2007) + \epsilon_{at}
\end{aligned} \tag{3}
$$

- D^r_{NT}: set of R 'ring' dummies which $= 1$ if area within given distance of **Newcastle-upon-Tyne** & $= 0$ otherwise

Data: Overview

Data: Overview

- Identify *relevant* research & researchers from Diamond user output:

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
- **Academic publications**

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

- ▷ Identify *related* research:

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

- ▷ Identify *related* research:
  - **Similar academic publications**

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

- ▷ Identify *related* research:
  - **Similar academic publications**
  - Similar patent applications

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

- ▷ Identify *related* research:
  - **Similar academic publications**
  - Similar patent applications
- ▶ Indirect Effect

Data: Overview

- ▷ Identify *relevant* research & researchers from Diamond user output:
  - **Academic publications**
  - Patent applications
- ▶ Direct Effect

- ▷ Identify *related* research:
  - **Similar academic publications**
  - Similar patent applications
- ▶ Indirect Effect

- ▷ Analysis at Local Administrative District level (379 LADs)

Data: Diamond User Output

Data: Diamond User Output

- Start with **complete** list of scientific publications from work at Diamond

Data: Diamond User Output

- ▷ Start with **complete** list of scientific publications from work at Diamond
- Standardize/clean names & affiliations

Data: Diamond User Output

- ▷ Start with **complete** list of scientific publications from work at Diamond
  - Standardize/clean names & affiliations
  - 347 publications in 121 journals (December 2010)

Data: Diamond User Output

- ▷ Start with **complete** list of scientific publications from work at Diamond
  - Standardize/clean names & affiliations
  - 347 publications in 121 journals (December 2010)
  - 1,760 authors affiliated to 441 institutions worldwide

Data: Diamond User Output

- ▷ Start with **complete** list of scientific publications from work at Diamond
  - Standardize/clean names & affiliations
  - 347 publications in 121 journals (December 2010)
  - 1,760 authors affiliated to 441 institutions worldwide
  - Location determined by affiliation postcode

Data: Diamond User Output

- ▷ Start with **complete** list of scientific publications from work at Diamond
  - Standardize/clean names & affiliations
  - 347 publications in 121 journals (December 2010)
  - 1,760 authors affiliated to 441 institutions worldwide
  - Location determined by affiliation postcode
  - Focus on GB affiliations: 1,282 authors affiliated to 194 institutions

Data: Academic publications

Data: Academic publications

- How to identify similar/relevant scientific publications?

Data: Academic publications

- How to identify similar/relevant scientific publications?
- Find 5 most similar articles to "Diamond Articles" according to adjusted **references overlap**

Data: Academic publications

- How to identify similar/relevant scientific publications?
- Find 5 most similar articles to "Diamond Articles" according to adjusted **references overlap**
- Find 1,528 related articles

Data: Academic publications

- How to identify similar/relevant scientific publications?
- Find 5 most similar articles to "Diamond Articles" according to adjusted **references overlap**
- Find 1,528 related articles
- Focus on authors with at least 1 UK affiliation

Summary statistics for Academic publications
Diamond academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	5.69	5	2.99	1	20
# affiliations per article	2.19	2	1.23	1	7
# affiliations per author	1.13	1	0.38	1	3
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	180.1	120.6	148.4	0	539.3
Distance (km) to Daresbury	206.2	219.4	76.7	0	340.0
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	152.4	116.5	153.1	0	623.4
Distance (km) to Daresbury	196.9	216.5	81.6	0	425.6

Summary statistics for Academic publications
Diamond academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	5.69	5	2.99	1	20
# affiliations per article	2.19	2	1.23	1	7
# affiliations per author	1.13	1	0.38	1	3
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	180.1	120.6	148.4	0	539.3
Distance (km) to Daresbury	206.2	219.4	76.7	0	340.0
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	152.4	116.5	153.1	0	623.4
Distance (km) to Daresbury	196.9	216.5	81.6	0	425.6

Summary statistics for Academic publications
Diamond academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	5.69	5	2.99	1	20
# affiliations per article	2.19	2	1.23	1	7
# affiliations per author	1.13	1	0.38	1	3
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	180.1	120.6	148.4	0	539.3
Distance (km) to Daresbury	206.2	219.4	76.7	0	340.0
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	152.4	116.5	153.1	0	623.4
Distance (km) to Daresbury	196.9	216.5	81.6	0	425.6

Summary statistics for Academic publications
Related academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	3.85	3	2.47	1	17
# affiliations per article	1.66	1	0.92	1	8
# affiliations per author	1.26	1	0.50	1	4
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	170.5	120.6	136.9	0	554.3
Distance (km) to Daresbury	192.0	209.7	85.7	0	347.3
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	164.4	118.4	155.1	0	624.3
Distance (km) to Daresbury	199.2	212.0	84.7	0	426.7

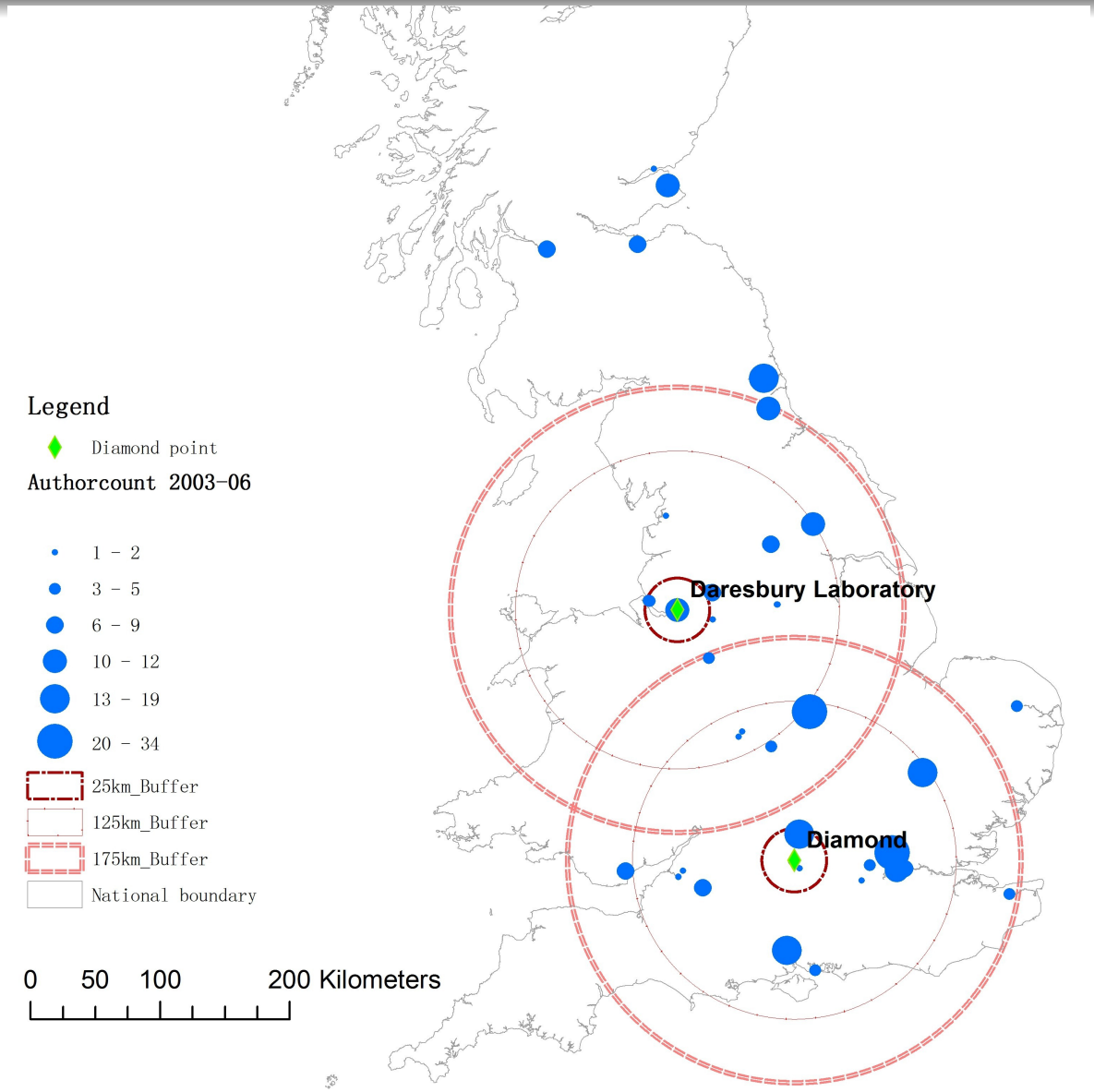
Summary statistics for Academic publications
Related academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	3.85	3	2.47	1	17
# affiliations per article	1.66	1	0.92	1	8
# affiliations per author	1.26	1	0.50	1	4
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	170.5	120.6	136.9	0	554.3
Distance (km) to Daresbury	192.0	209.7	85.7	0	347.3
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	164.4	118.4	155.1	0	624.3
Distance (km) to Daresbury	199.2	212.0	84.7	0	426.7

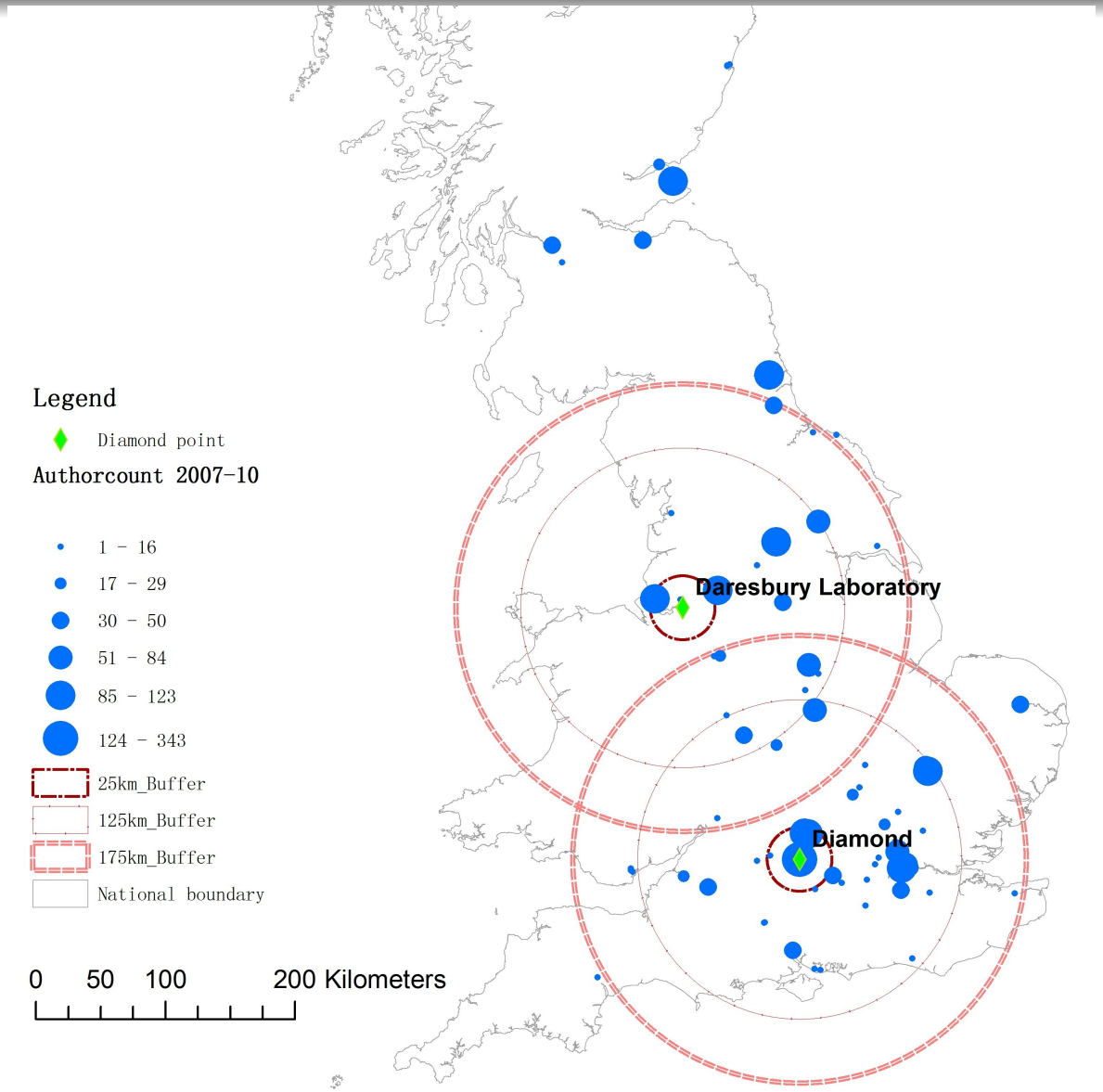
Summary statistics for Academic publications
Related academic journal articles

	Mean	Median	Std. Dev.	Min.	Max.
Descriptive Statistics of Authors & Affiliations (UK only)					
# authors per article	3.85	3	2.47	1	17
# affiliations per article	1.66	1	0.92	1	8
# affiliations per author	1.26	1	0.50	1	4
Geographical Distribution of Authors' Affiliations					
$<$ 2007 (*Before* Establishment of Diamond)					
Distance (km) to Diamond	170.5	120.6	136.9	0	554.3
Distance (km) to Daresbury	192.0	209.7	85.7	0	347.3
$\geq$ 2007 (*After* Establishment of Diamond)					
Distance (km) to Diamond	164.4	118.4	155.1	0	624.3
Distance (km) to Daresbury	199.2	212.0	84.7	0	426.7

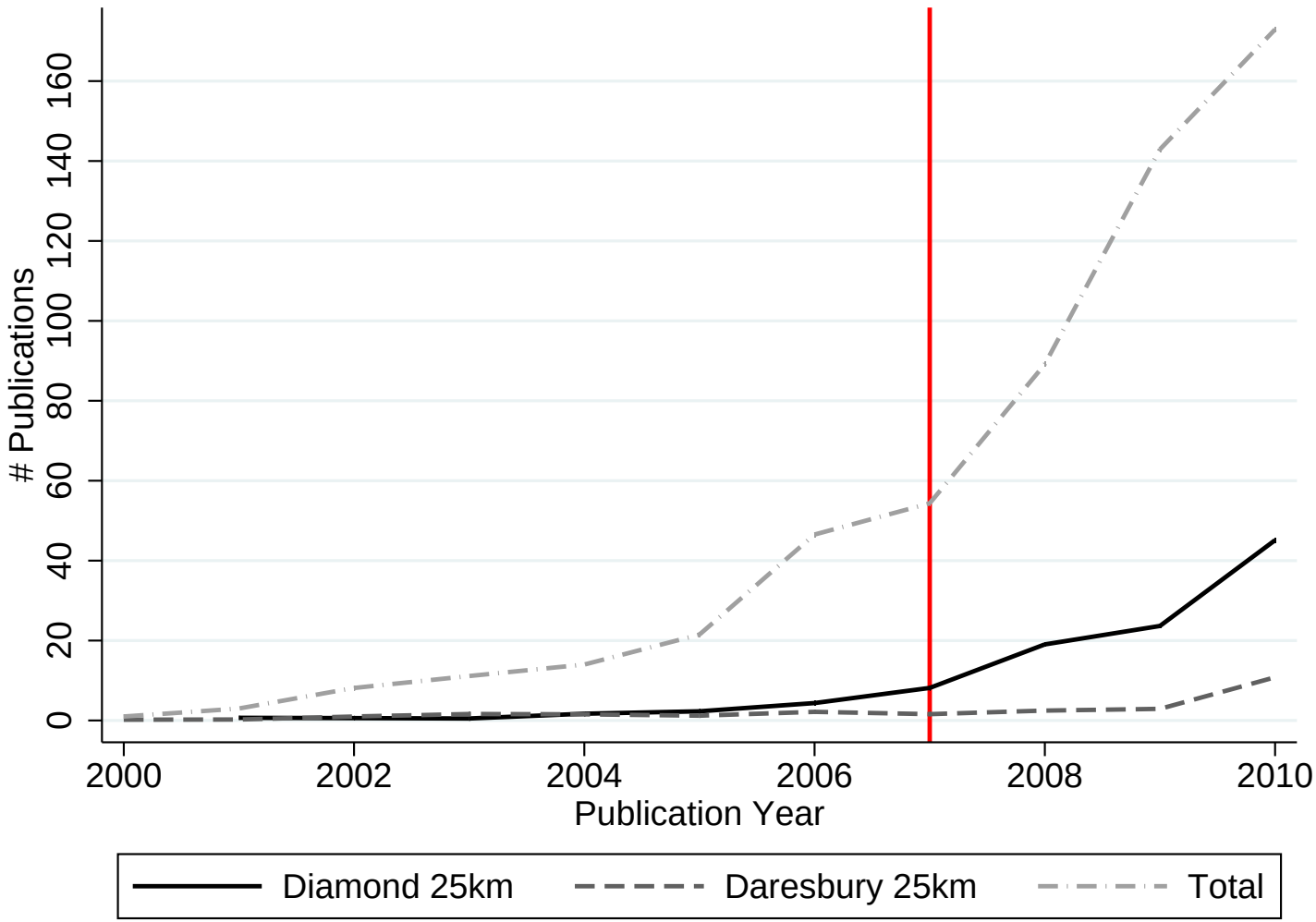
Academic Publications: **Pre**-Diamond 2003-2006



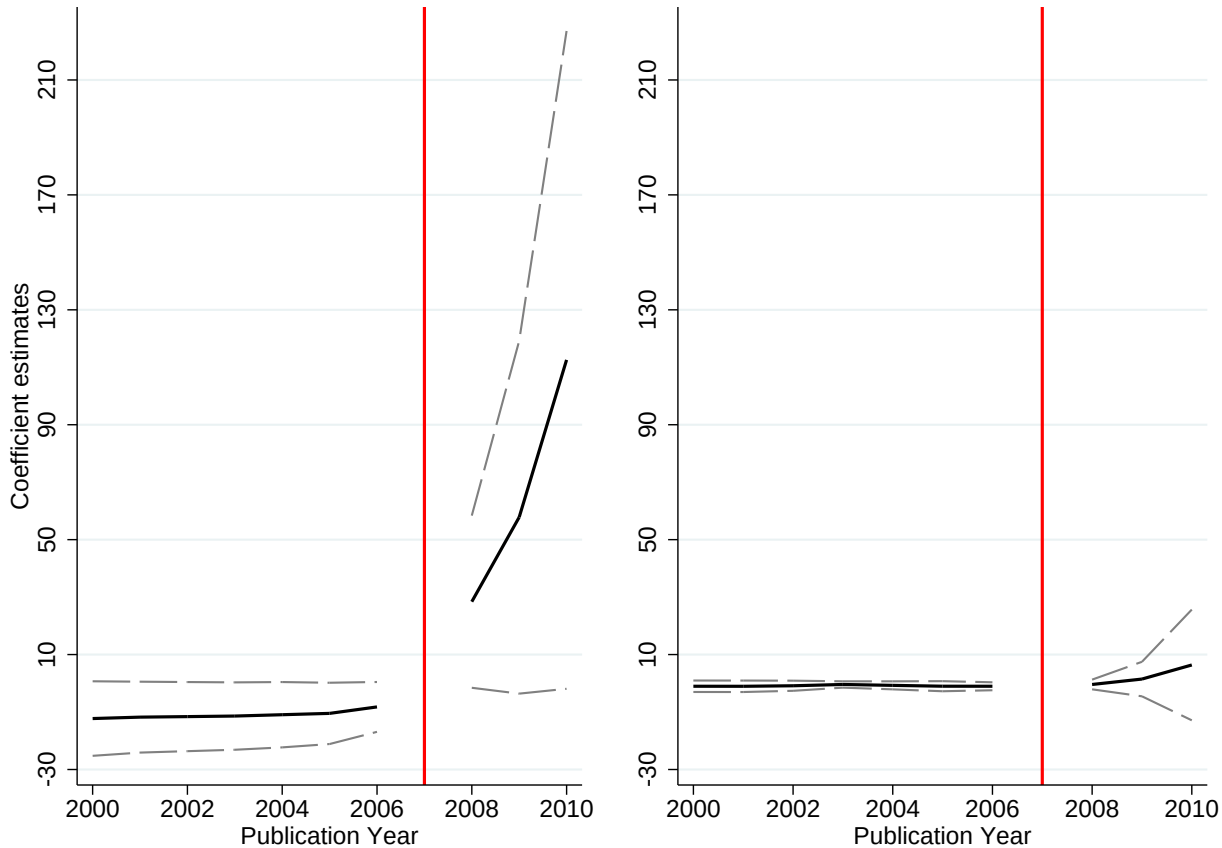
Academic Publications: **Post**-Diamond 2007-2010



Academic Publications: # Articles – Distance to **Diamond** vs **Daresbury**



Academic Publications: annual estimates for **Diamond** vs **Daresbury**



Notes: Annual coefficient estimates β_{Ct} from the regressions $p_{at} = \alpha_a + \sum_t D_t + \beta_{Dt} D_C^{25} \times D_t + \epsilon_{at}$ (with $C = [DI, DA]$) for Diamond and Daresbury (within 25km) where $t = 2000, 2001, ..., 2010$, 2007 is the omitted category.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond	25km	6.649*	0.529*	0.493	
		(3.474)	(0.308)	(0.305)	
	125km	0.047	0.015	-0.001	
		(0.109)	(0.026)	(0.021)	
	175km	-0.228***	-0.031*	-0.003	
		(0.069)	(0.049)	(0.016)	
Diamond $\times$ $I(t \geq 2007)$	25km		16.834*	16.135*	16.250*
			(8.983)	(8.883)	(8.881)
	125km		0.087	-0.317*	-0.340**
			(0.238)	(0.166)	(0.159)
	175km		-0.528***	-0.045	-0.116
			(0.150)	(0.131)	(0.126)
Time dummies		YES	YES	YES	YES
Controls		NO	NO	YES	YES
Fixed Effects		NO	NO	NO	YES
Daresbury		NO	NO	NO	NO
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above $\times$ $I(t \geq 2007)$; Labor force; Labor force $\times$ $I(t \geq 2007)$.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond	25km	6.649*	0.529*	0.493	
		(3.474)	(0.308)	(0.305)	
	125km	0.047	0.015	-0.001	
		(0.109)	(0.026)	(0.021)	
	175km	-0.228***	-0.031*	-0.003	
		(0.069)	(0.049)	(0.016)	
Diamond $\times$ $I(t \geq 2007)$	25km		16.834*	16.135*	16.250*
			(8.983)	(8.883)	(8.881)
	125km		0.087	-0.317*	-0.340**
			(0.238)	(0.166)	(0.159)
	175km		-0.528***	-0.045	-0.116
			(0.150)	(0.131)	(0.126)
Time dummies		YES	YES	YES	YES
Controls		NO	NO	YES	YES
Fixed Effects		NO	NO	NO	YES
Daresbury		NO	NO	NO	NO
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above $\times$ $I(t \geq 2007)$; Labor force; Labor force $\times$ $I(t \geq 2007)$.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond	25km	6.649*	0.529*	0.493	
		(3.474)	(0.308)	(0.305)	
	125km	0.047	0.015	-0.001	
		(0.109)	(0.026)	(0.021)	
	175km	-0.228***	-0.031*	-0.003	
		(0.069)	(0.049)	(0.016)	
Diamond $\times$ $I(t \geq 2007)$	25km		16.834*	16.135*	16.250*
			(8.983)	(8.883)	(8.881)
	125km		0.087	-0.317*	-0.340**
			(0.238)	(0.166)	(0.159)
	175km		-0.528***	-0.045	-0.116
			(0.150)	(0.131)	(0.126)
Time dummies		YES	YES	YES	YES
Controls		NO	NO	YES	YES
Fixed Effects		NO	NO	NO	YES
Daresbury		NO	NO	NO	NO
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above $\times$ $I(t \geq 2007)$; Labor force; Labor force $\times$ $I(t \geq 2007)$.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond × $I(t \geq 2007)$	25km	16.788*	16.277*	17.071*	16.482*
		(8.985)	(8.886)	(8.987)	(8.887)
	125km	0.044	-0.309	0.327	-0.093
		(0.307)	(0.233)	(0.245)	(0.179)
	175km	-0.572**	-0.087	-0.289*	0.130
		(0.245)	(0.219)	(0.159)	(0.176)
Daresbury × $I(t \geq 2007)$	25km	1.060	0.811	1.343	1.052
		(1.113)	(1.029)	(1.098)	(1.018)
	125km	-0.107	-0.002	0.117	0.187
		(0.316)	(0.248)	(0.250)	(0.196)
	175km	-0.347*	-0.016	-0.357**	0.140
		(0.329)	(0.254)	(0.151)	(0.178)
Time dummies		YES	YES	YES	YES
Fixed Effects		YES	YES	YES	YES
Controls		NO	YES	NO	YES
Newcastle		NO	NO	YES	YES
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above × $I(t \geq 2007)$; Labor force; Labor force × $I(t \geq 2007)$.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond $\times$ $I(t \geq 2007)$	25km	16.788*	16.277*	17.071*	16.482*
		(8.985)	(8.886)	(8.987)	(8.887)
	125km	0.044	-0.309	0.327	-0.093
		(0.307)	(0.233)	(0.245)	(0.179)
	175km	-0.572**	-0.087	-0.289*	0.130
		(0.245)	(0.219)	(0.159)	(0.176)
Daresbury $\times$ $I(t \geq 2007)$	25km	1.060	0.811	1.343	1.052
		(1.113)	(1.029)	(1.098)	(1.018)
	125km	-0.107	-0.002	0.117	0.187
		(0.316)	(0.248)	(0.250)	(0.196)
	175km	-0.347*	-0.016	-0.357**	0.140
		(0.329)	(0.254)	(0.151)	(0.178)
Time dummies		YES	YES	YES	YES
Fixed Effects		YES	YES	YES	YES
Controls		NO	YES	NO	YES
Newcastle		NO	NO	YES	YES
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above $\times$ $I(t \geq 2007)$; Labor force; Labor force $\times$ $I(t \geq 2007)$.

Academic Articles: OLS (379 LAD – 2000-2010)

		[I]	[II]	[III]	[IV]
Diamond $\times$ $I(t \geq 2007)$	25km	16.788*	16.277*	17.071*	16.482*
		(8.985)	(8.886)	(8.987)	(8.887)
	125km	0.044	-0.309	0.327	-0.093
		(0.307)	(0.233)	(0.245)	(0.179)
	175km	-0.572**	-0.087	-0.289*	0.130
		(0.245)	(0.219)	(0.159)	(0.176)
Daresbury $\times$ $I(t \geq 2007)$	25km	1.060	0.811	1.343	1.052
		(1.113)	(1.029)	(1.098)	(1.018)
	125km	-0.107	-0.002	0.117	0.187
		(0.316)	(0.248)	(0.250)	(0.196)
	175km	-0.347*	-0.016	-0.357**	0.140
		(0.329)	(0.254)	(0.151)	(0.178)
Time dummies		YES	YES	YES	YES
Fixed Effects		YES	YES	YES	YES
Controls		NO	YES	NO	YES
Newcastle		NO	NO	YES	YES
Obs		4,121	4,121	4,121	4,121

Controls: % NVQ4 and above; % NVQ4 and above $\times$ $I(t \geq 2007)$; Labor force; Labor force $\times$ $I(t \geq 2007)$.

Robustness – variations of the basic model

- Channels:
  - Research input: author counts by LA and year
  - Research input: 'unique' institution counts by LA and year
- Vary size of distance rings (30km, 100km, 150km)
- Vary number of distance rings (25km, 125km; 25km, 75km, 125km, 175km)
- Daresbury shut-down effect
- Diamond construction effect
- Limit sample to LAs that report a positive author/article count in at least 1 sample year
- Alternative ways of constructing related articles sample (field/journal restrictions)
- Count data model

Conclusion

Conclusion

- Does the location of basic scientific research infrastructure affect its use and impact?

Conclusion

- Does the location of basic scientific research infrastructure affect its use and impact?
- Analyze the impact of a GBP380 million scientific facility on geographic distribution of research

Conclusion

- Does the location of basic scientific research infrastructure affect its use and impact?
- Analyze the impact of a GBP380 million scientific facility on geographic distribution of research
- Find evidence of impact on clustering of research that uses facility & related research (moved closer to Diamond)

Conclusion

- Does the location of basic scientific research infrastructure affect its use and impact?
- Analyze the impact of a GBP380 million scientific facility on geographic distribution of research
- Find evidence of impact on clustering of research that uses facility & related research (moved closer to Diamond)
- Not driven by increased proximity to existing clusters (unless unique to Diamond)

Conclusion

- Does the location of basic scientific research infrastructure affect its use and impact?
- Analyze the impact of a GBP380 million scientific facility on geographic distribution of research
- Find evidence of impact on clustering of research that uses facility & related research (moved closer to Diamond)
- Not driven by increased proximity to existing clusters (unless unique to Diamond)
- Work in progress...