# Case Studies in Microeconomic Evaluation Data and methods for learning what works

## Case study 3

*Competence Centre on Microeconomic Evaluation (CC-ME)*
*Joint Research Centre 30 Nov 2022*

# Outline

- Introduction to **Regression Discontinuity Design**
- **Case study**: The impact of the European Research Council grants on scientific productivity

# Regression Discontinuity Design for causal inference

- Main goal is to learn about effect of a policy or intervention (**treatment**)
  - Did units who participated to the intervention (exposed to the policy) benefit from it?
- Compare the outcome of a unit who participated to the intervention, to the outcome of **the same unit had it not participated to the intervention**
  - Fundamental problem of causal inference: it is not observable.
  - Need of a **counterfactual**
- If treatment is random $\rightarrow$ easy (**Randomized control trial**)
- If treatment is not random $\rightarrow$ **observational studies**: find a proper control group
  - Regression Discontinuity; Difference-in-differences; Matching methods; Instrumental variable

**How to chose which CIE method to use?**

1. **Design**: how are treated units selected?
2. **Data**

# RDD: main ingredients

1. Units have a **score** (*running variable, forcing variable*)
2. Treatment is assigned based on the score and a known **threshold**(*cutoff*)
3. The **treatment** is:
   - given to units whose score is greater than the threshold
   - not given to units whose score is lower then the threshold

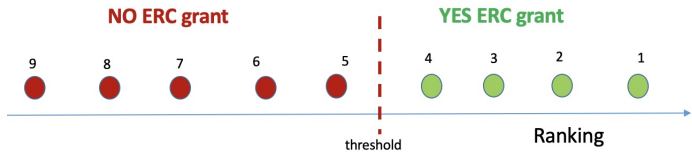
   (or vice-versa)

# Examples

| Units | Treatment | Score | Outcome |
|---|---|---|---|
| Students | scholarship | income | academic progression |
| Students | college admission | high school gpa | future earnings |
| Regions | EU Funds | GDP | grow |
| Municipalities | political party | margin of victory | spending type, women edu |
| Households | poverty program | poverty index | child health |

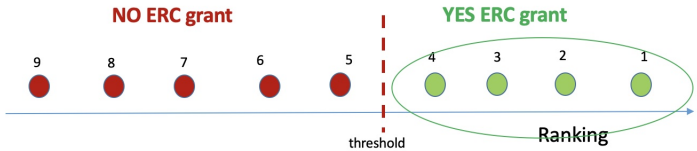
# Example: the European Research Council grant

| Units | Treatment | Score | Outcome |
|---|---|---|---|
| ERC Applicants | ERC grant | Evaluation score or rank | scientific productivity |



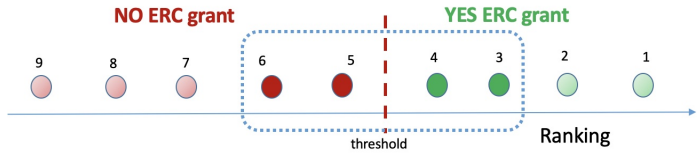
# THE counterfactual question

What would have happened to the ERC grant recipients had they not won the grant?
Need a control group!




- Extreme case of lack of common support!

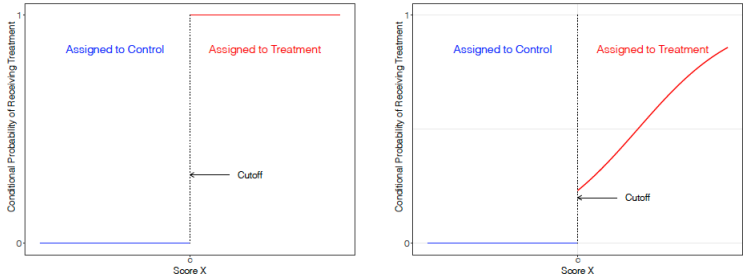
# RDD intuition




- Around the threshold, ERC applicants are **almost similar**, but only some of them receive the treatment
- ERC applicants barely below the threshold can be used as a comparison group for the applicants just above.

# RDD: intuition

- Abrupt change in the probability of receiving the treatment at the known threshold.
- Learn about the local causal effect of the treatment on an outcome of interest,
- Around the threshold, units are **almost similar**, but only some of them receive the treatment
- Units with scores barely above the threshold can be used as a comparison group for units with scores barely below it.
- RDD is a **local** randomised experiment at the threshold




(a) Sharp RD (b) Fuzzy RD (one-sided compliance)

Skovron & Titiunik 2015, Fig. 1, p.5

## SHARP RDD

- Compliance with treatment assignment is perfect
- All eligible individuals receive the treatment if they are above (below) the threshold, all non eligible don't

## FUZZY RDD

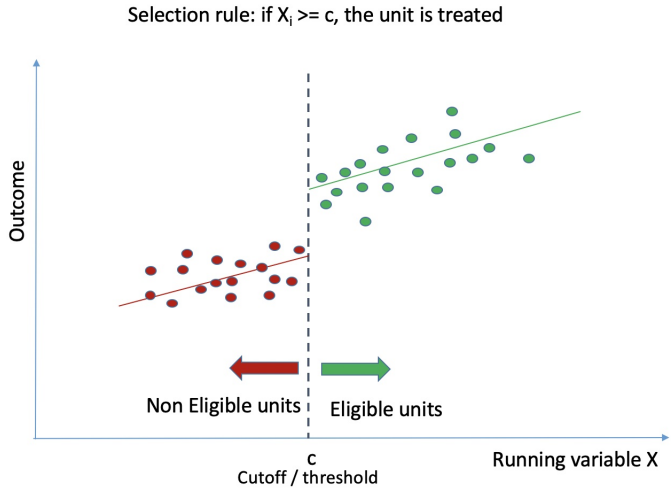
- Non compliance with treatment assignment
- Some eligible individuals may not received the treatment even if they are above (below) the threshold, some non eligible may receive receive the treatment

# Main conditions for RDD

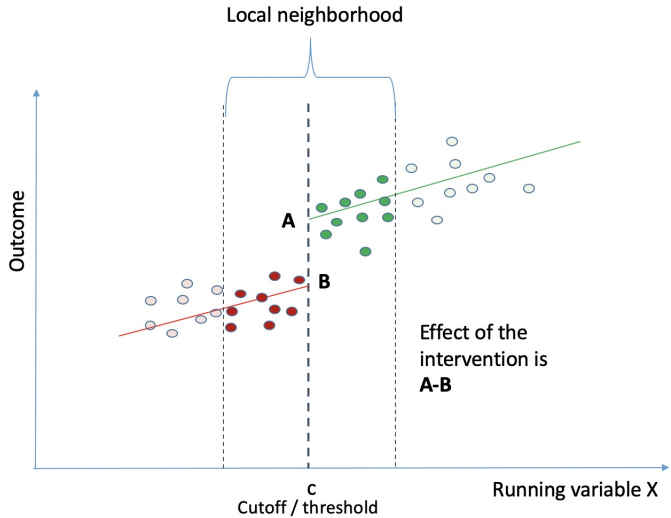
- Assignment to the treatment only occurs through a **known and measured** decision rule
- No manipulation of the value of the score to become eligible
- No other polices using the **same eligibility criteria** and which will affect the outcome
- Smooth relationship between the outcome and the score

# No intervention

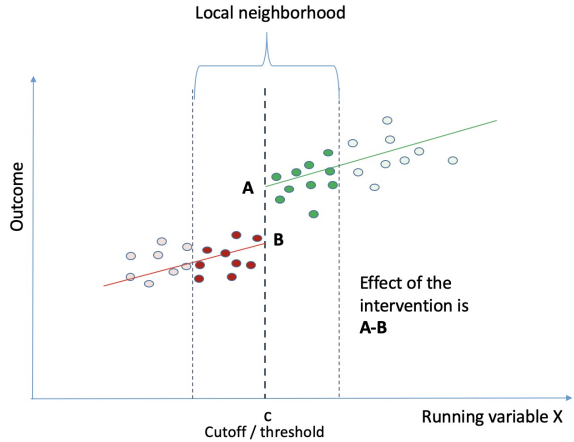
# With intervention



# With intervention



- Choice of the bandwidth
- Choice of the order of the polynomial
- Choice of weights



# RDD pros & cons

Pros

- Credibility: RDD is considered has the most credible method in terms on internal validity of the results obtained
- Transparency: RDD can be illustrated with simple graphs
- Weak conditions which can be checked

Cons

- External validity is limited: the estimated treatment effect is **local** to the discontinuity

# ERC grant

- Collaboration between JRC.I.1 and **DG RTD.G2**
- Evaluation of the European Research Council grant, in the framework of the FP7 program (2007-2013)

# FP7 program (2007-2013): European Research Council grant

- Premier European funding organisation for excellent frontier research. It funds **creative researchers** of any nationality and age, to run projects based across Europe.
- Since 2007, ERC funded more than 12,000 projects and over 10,000 researchers.
- Every year, there is a call of proposal devoted to three ERC programmes:
  1. Starting Grant (SG) 2-7 years since PhD, max 1.5 millions
  2. Consolidator grant (CG) 7-12 years, max 2 millions
  3. Advanced grant (AG), at least 10 years of research experience, max 2.5 millions
- For each call and type of grant there are 25 *panels*, covering one of three *domains*: Life sciences (LS), Physical and Engineering Sciences (PE), Social sciences and Humanities (SH).

- Two steps selections:
  1. From the pool of applicants, in each panel, roughly the 30% of proposal are selected to go to the second phase.
  2. Proposals are **ranked**, according to **scientific excellence** and roughly the 50% of proposals who passed the first step are funded.
- *Ground-breaking nature, ambition, and feasibility of the research project, and the intellectual capacity, creativity, and commitment of the Principal Investigator(s).*
- The two steps selection is done by a panel composed by renowned scientists and scholars selected by the ERC Scientific Council.

## Research Question

- **What is the causal impact of receiving a ERC grant on researcher's productivity and network?**
- **Data**: Universe of ERC grant applications 2007-2013.
- **Outcomes**: Scopus bibliometric data (number of publications, H-index, etc...)

**Econometric Strategies**:

1. **RDD (exploiting the assignment mechanism at the threshold)**.
   - Retrieve the effect of the grant close to the threshold (local effect).
2. DiD (exploiting the availability of outcomes over time, before and after the assignment).
   - This allows to estimate heterogeneous effects also of best performers - at the top of the distribution. ( not today)

# Literature

- Literature on effect of grant on research production at university level (mostly positive).
- Literature on effect of grant on individual researchers (positive or null results, depending on short/long term horizon, type of outcomes, country), e.g.:
  - Denmark and Norway (Langfeldt et al., 2015), US (Jacob and Leffgren, 2011, Wang et al. 201), former Soviet Union (Ganguli, 2017) Chile (Benavente et al. 2019), Argentina (Chudnovsky et al. 2008).
- Typical explanation for absence of effect on productivity: access to alternative fundings for control group.
- Existing evidence on ERC is mostly descriptive, **no causal study yet**.

## Data

- Data on all applicants for the period 2007-2013
- Ranking available only for the ones who passed the first step.
- Pool data of the various year, as we wish to do the analysis by domain.
- We do not consider the **Interdisciplinary** panel and **CoG**.

**Table:** Distribution of rejected and winner applicants by type of grant and domain

| Grant type | Domain | Rejected | Winners | total |
|---|---|---|---|---|
| Advanced | LS | 627 | 591 | 1,218 |
| Advanced | PE | 749 | 722 | 1,471 |
| Advanced | SH | 315 | 252 | 567 |
| Starting | LS | 498 | 778 | 1,276 |
| Starting | PE | 632 | 977 | 1,609 |
| Starting | SH | 272 | 364 | 636 |
| | Total | 3,093 | 3,684 | 6,777 |

## Outcomes: Scientific Productivity & network

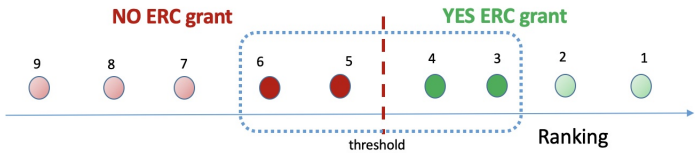
- We link each applicant to Scopus (disambiguation done via algorithm and manual checks)
- We use the database of publications downloaded from Scopus -as of April 2021- to construct the following outcomes **for each year after the call year, up to 9 years**:

1. Cumulative number of articles after receiving the grant.
2. Share of publications/articles in top 10% and 1% journals. The ranking of each journal is to be constructed by Scopus ranking list.
3. h-index
4. Number of coauthors
5. Number of funds acknowledged (Matthew effect)
   - EU(ERC, marie curie,FP7, H2020 )
   - All other ( other EU, national, unversity, etc.)

- 2 types * 3 domains * 7 outcomes * 9 years = **378** point estimates

# RDD

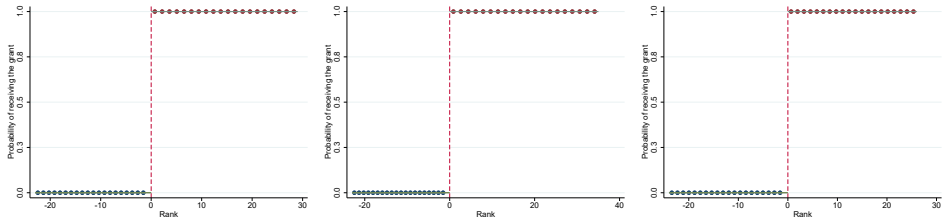
- **Design**
  - Proposal are ranked, only some of them are funded.
  - The ranking can be used as the running variable.
  - The threshold is the rank of the last proposal funded.
- **Data**
  - Data on the ranking **all** applicants (also the rejected ones)
  - Data on outcomes



# RDD - methodological details

- Sharp RDD
- Since we are pooling different *panels* and *call years* together we add **panel * call year fixed effect**
- Non parametric local polynomial estimators, optimal bandwidth one common MSE. (Calonico et al. 2014).
- Robust variance estimator

# Sharp RDD



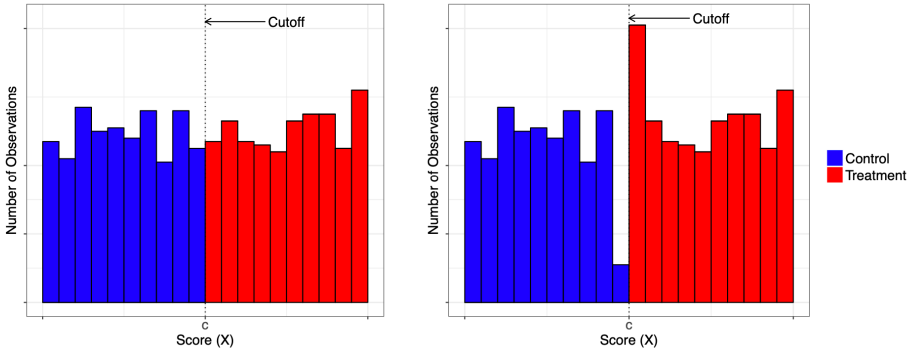

Probability of receiving the grant conditional on the running variable, StG LS, PE ,SH

# RDD assumptions

1. Absence of manipulation of running variable
2. No effect of pre-determined covariates or placebo outcomes.

# Absence of manipulation of running variable

Is the number of observations below the cutoff surprisingly different from number of observation above it?

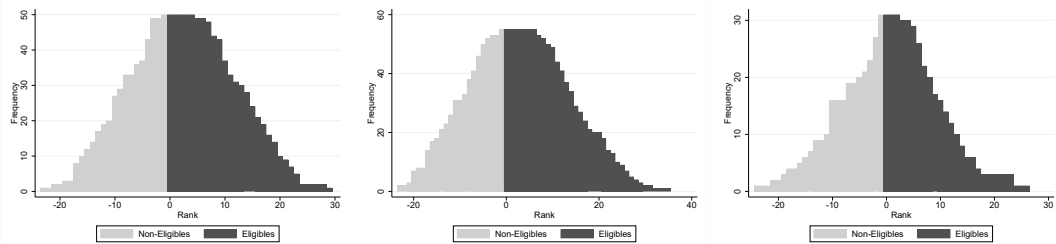



(a) No Sorting

(b) Sorting

*Cattaneo & Idrobo & Titiunik 2020, Fig. 18, p.97*

# Absence of manipulation of running variable

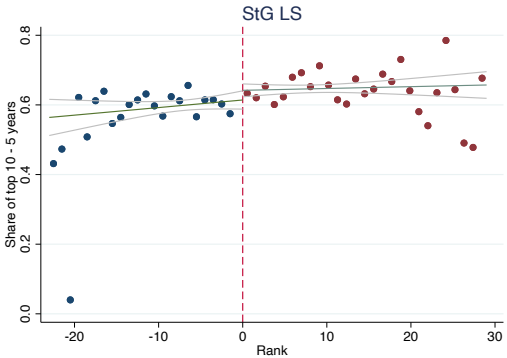


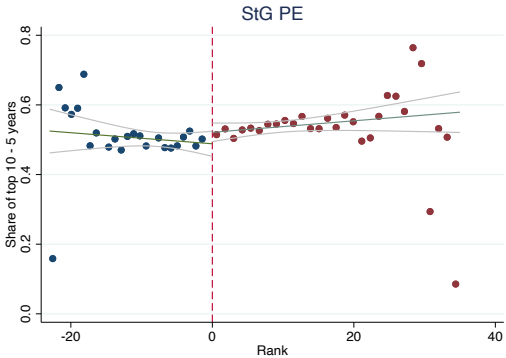

Distribution of the running variable (rank) , StG LS, PE, and SH

# Pre-determined covariates or placebo outcomes.

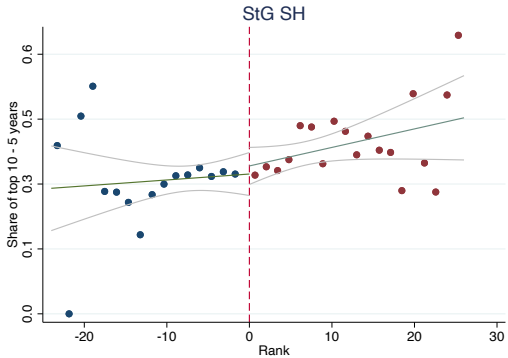
| **SdG LS** | Year first paper | Female | # publications | # top1 | Share top1 | # top 10 |
|---|---|---|---|---|---|---|
| Robust | 0.310 | -0.040 | -3.917 | 0.404 | 0.055* | -0.159 |
| | (1.027) | (0.102) | (3.794) | (0.752) | (0.027) | (2.204) |
| Observations | [498:778] | [498:778] | [498:778] | [498:778] | [498:778] | [498:778] |
| Bandwidth | [5:5] | [6:6] | [6:6] | [6:6] | [5:5] | [7:7] |
| Effect. observations | [147:249] | [220:348] | [220:348] | [220:348] | [147:249] | [220:348] |
| | Share top 10 | h-index | # coauthors | Any fund | EU fund | |
| Robust | 0.053 | 1.257 | 4.475 | 1.386* | 0.198 | |
| | (0.050) | (1.493) | (18.285) | (0.688) | (0.109) | |
| Observations | [498:778] | [498:778] | [498:778] | [498:778] | [498:778] | |
| Bandwidth | [7:7] | [5:5] | [6:6] | [6:6] | [4:4] | |
| Effect. observations | [220:348] | [182:299] | [182:299] | [182:299] | [147:249] | |

# Results - graphical inspection: share of articles in top 10% journals



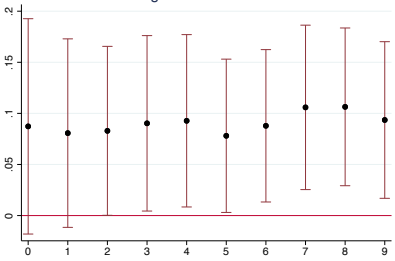



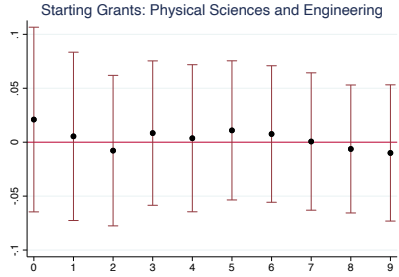
# Results - graphical inspection: share of articles in top 10% journals

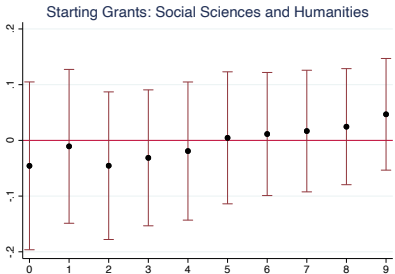


# Results - RDD estimation

| | (1) | (2) | (3) |
|---|---|---|---|
| Share articles in top 10, 5 years | LS | PE | SH |
| | 0.078* | 0.011 | 0.005 |
| | (0.038) | (0.033) | (0.060) |
| Observations | [498:778] | [632:977] | [272:364] |
| BW Type | mserd | mserd | mserd |
| Bandwidth | [4:4] | [8:8] | [6:6] |
| Effect. observations | [191:250] | [388:490] | [122:182] |
| Order Loc. Poly. (p) | 1 | 1 | 1 |
| Order Bias (q) | 2 | 2 | 2 |



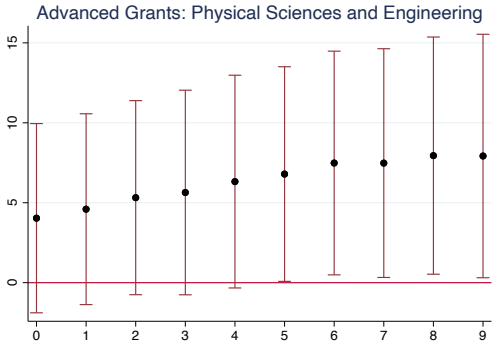



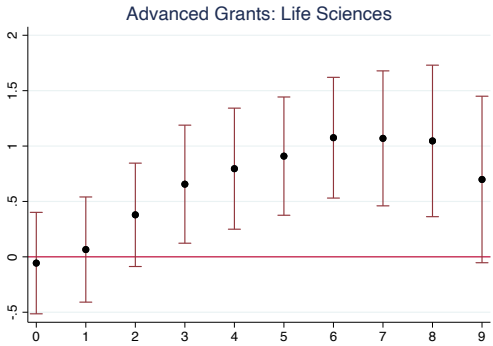

Share of top 10 articles, in StG, from the call year up to 9 years after the call

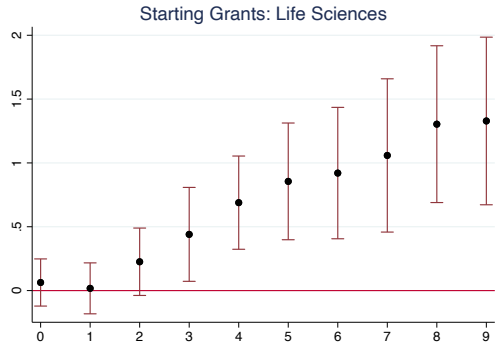
European Commission

# h-index

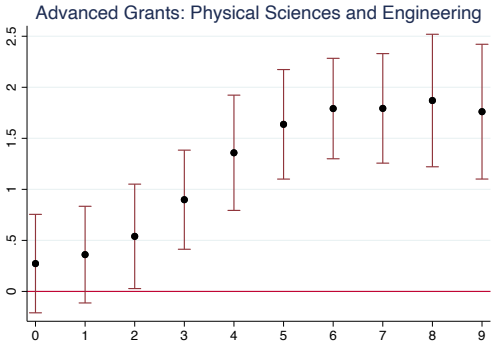

Advanced Grants: Physical Sciences and Engineering
0
5
10
15
0
1
2
3
4
5
6
7
8
9

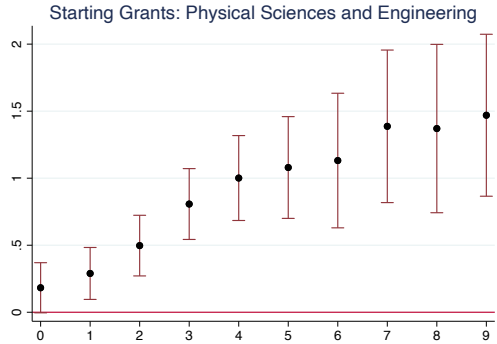
# EU funds



# EU funds


Advanced Grants: Physical Sciences and Engineering
0
.5
1
1.5
2
2.5
0
1
2
3
4
5
6
7
8
9


Starting Grants: Physical Sciences and Engineering
0
.5
1
1.5
2
0
1
2
3
4
5
6
7
8
9

# Heterogeneity by field

- We divided the 3 domains into 11 sub-fields, aggregating the 25 panels:
  - LS: Biology and Chemistry, Medicine,Applied life sciences
  - PE: Math, Physics, Chemistry, Engineering,Universe and Earth Sciences
  - SH: Individuals and Institutions, Institutions and behavior, Human Mind
- Positive effects in LS come from the **Medicine** field in **StG** (share of top 1 and top 10)
- Positive effects in PE come from the **Physics** field in **StG** (# of articles and h-index)

# Conclusions & ongoing work

- **Conclusions**
  - Around the thresholds, winning the ERC grant doesn't cause an increase in productivity
  - Some localized effect for the Medicine and Physics sub-fields
  - Matthew effect: money brings money
- **Ongoing work**
  - Using an alternative approach : differences-in-differences can allow **going beyond the local estimates** at the threshold
  - Investigating how to build **other outcomes**, to better capture the real focus of the ERC : scientific excellence, novelty, innovation, etc..
  - Effects on the career progression?
  - Heterogeneity by country
  - Marie-Curie grant
  - Replicate analysis for H2020

# Thank you