Data lab #5

Giorcelli, Michela and Petra Moser, Copyrights and Creativity: Evidence from Italian Operas. Journal of Political Economy, Oct 2020

Question for discussion:	1. How did the creation of copyrights in Lombardy and Venetia affect the creation of new operas? 2. How did extensions in copyrights influence output?
Deadline:	Lecture on Thursday, November 19, 11:00 am
Datacamp:	No exercises this week, your grade is only based on the report

Introduction

Copyrights protect an increasing share of what we consume today, including news, images, film, and music. Despite their importance, economic analyses continue to be rare and there is little proof so far that copyrights fulfill their primary purpose: to encourage creativity. A key reason for the dearth of empirical evidence is that nearly all changes in copyright laws today are a result of lobbying by the owners of particularly valuable copyrighted works. These lobbying efforts are reflected in the nicknames of modern copyright laws, for example, the 1998 "Mickey Mouse Protection Act."

To identify the causal effects of copyrights on creativity, this paper exploits variation in the adoption of copyright laws as a result of the timing of Napoléon's military victories. In 1796, Napoléon began his Italian campaign by invading the Kingdom of Sardinia at Ceva. Although he was unable to subdue Sardinia at the time, two other states, Lombardy and Venetia, came under French rule in 1801, and adopted all French laws, including France's copyright law. The French law granted composers, who had no protection before the law, the right to sell or rent their work, plus charge for repeat performances of their work. In this lab, you will investigate whether the adoption of copyrights encouraged the creation of new works, measured by the creation of new operas that premiered across eight Italian states between 1780 and 1820.

Instructions

0. In your introduction, explain why copyrights are important and why it is difficult to investigate the causal effects of copyrights on creativity with modern data.
1. Reproduce Figure 2 from the paper. Tip: Focus on the content and critical features, like properly labelled axis, but do over-invest in formatting. Formatting is not crucial, simply make sure your tables and figures follow the criteria in our syllabus.
2. Reproduce Table 1 in the paper. Please report, for 1781-1800 & 1801-1820 and together for the whole period:
    a. The total number of operas (the top panel in Table 1)

b. The number of popular operas. Use two definitions of popular operas:

   i. Historically popular operas that appear in Annals of Opera

   ii. Durable operas that were available for purchase on Amazon in 2014

c. How many popular operas are there *in total*? Remember, the two definitions overlap, and so the total number is *not* going to be 62 + 42 = 104.

3. We want to adopt the diff-in-diff strategy and run the following diff-in-diff regression:

$$y_{ct} = a_t + \gamma_c + \beta_0 Post_t + \beta_1 Treat_c + \beta_2 Treat_c \times Post_t + \varepsilon_{ct}$$

Here, a_t are time fixed effects and γ_c are state fixed effect. The variable $Post_t$ is an indicator for years after 1800, the variable $Treat_c$ is an indicator for states that adopted copyrights.

a) For each regression, report coefficient estimates, standard errors and p-values.

   i. Run the regression for all operas with a known composer and a known title.

   ii. Discuss your results. Which coefficient identifies the treatment effect? Is your estimate statistically significant? Why is it important to include state and year fixed effects?

b) Redo part a), running two regressions using only the data on the **popular** operas, as defined in question 2 (separately for the two definitions).

   i. Are the sizes of these coefficients plausible? What do you think explains the magnitude of the estimates?

c) Re-run the regression from part a) without the fixed effects

   i. Explain how the results change and explain why. You can report results for a), b) and c) in one table with four columns.

d) *Optional:* Calculate robust standard errors and compare with the standard errors you calculated above. How are they different? Why?

e) *Optional:* Calculate standard error with clustering at the level of states. Compare with both robust and non-robust standard errors. Explain.