SFU

Replication Assignment

ECON 836 – Applied Econometrics

Professor John Gilraine

Nastaran Khorram

Sprin 2025

SFU

Paper Chosen to Replicate:

"Income Segregation and the Rise of the Knowledge Economy" by Enrico Berkes and César S. B. Gaetani (2023).
Link:https://www.researchgate.net/publication/369707395_Income_Segregation_and_the_Rise_of_the_Knowledge_Economy

Data Source:

Replication package from the American Economic Journal: Applied Economics.

Link: https://www.openicpsr.org/openicpsr/project/152903/version/V1/view

Empirical Strategy:

The authors use the instrumental variable (IV) method to estimate the causal effect of innovation on income segregation. Since patent growth might be influenced by other local factors, they use predicted patent growth based on a pre-1990 patent citation network as an instrument. This helps isolate changes in innovation that aren't caused by recent local trends, making it more likely the results capture a true causal link between innovation and segregation.

Author's Main Results:

The authors find that cities with more innovation (measured by growth in patents) became more income segregated over time. Specifically, a one standard deviation increase in patent growth led to a 0.65-point increase in income segregation between 1990 and 2010. That's about 31% of the overall increase in segregation during that period. The result suggests that as innovation grows, people become more sorted by income, especially in cities with more knowledge-based jobs.

My Replication:

As my code is attached, I replicated the main IV regression using Python and got a very similar result to the paper. The coefficient on patent growth was about 1.3049 and multiplying it by the standard deviation of patent growth (0.499) gives 0.6512, meaning a one-standard-deviation increase in innovation raises income segregation by 0.6512 Gini points. Dividing 0.6512 by the standard deviation of the change in income segregation (2.12)[i] gives about 0.31 (0.3071), which matches the main finding reported in the abstract. I also replicated the scatter plots shown in Figure 3 (patent growth vs. changes in segregation), which is also attached.

Empirical Weakness/ Robustness:

1. Concerns on Robustness:

One thing that might be affecting the main result is that it could be mostly driven by a few rich, high-tech cities like San Francisco, Boston, or New York. These places were already highly innovative and economically unequal before the 1990s, and they also had strong patenting networks that make them central to the instrument used in the paper. In fact, the authors themselves mention that innovation is concentrated in a handful of large cities, which also happen to have higher levels of income segregation. So, it's possible that the link between innovation and segregation mainly reflects what's happening in those specific places, not something that applies more broadly across the country. If that's the case, the main finding wouldn't be very general or representative of lower-income or smaller cities.

2. Changes I Made in Data:

To check this, I ran the same main IV regression again, but this time only for cities that had below-average income levels in 1990. I used the median of the log household income variable to split the sample and focused on the lower half, these are cities that were less wealthy at the start of the period and likely had less exposure to tech-based innovation. The goal was to see whether the relationship between patent growth and segregation still holds in places that weren't already rich or considered innovation hubs. I kept the same regression setup as the original paper, using predicted patent growth as an instrument for actual growth, and included the same control variables, I just changed the sample to test the result in a different context.

3. Differences / Changes in the Results:

In this smaller group of lower-income cities, the coefficient on patent growth was about 1.1577, meaning the effect of innovation on segregation was still positive, but smaller and not statistically significant. This suggests that the main result is stronger in richer cities, where people might sort more easily by income. The takeaway is that the original findings are still meaningful, but the effect seems especially relevant for wealthier urban areas, not necessarily for all cities.

---

[i] The values 0.499 (standard deviation of patent growth) and 2.12 (standard deviation of the change in income segregation) are summary statistics mentioned in the paper. As they were not part of the original calculations and were based on the same dataset I used, I didn't re-calculate them and just used the values as given. (Page 86 of the paper)