Research Assistantship - Data Collection

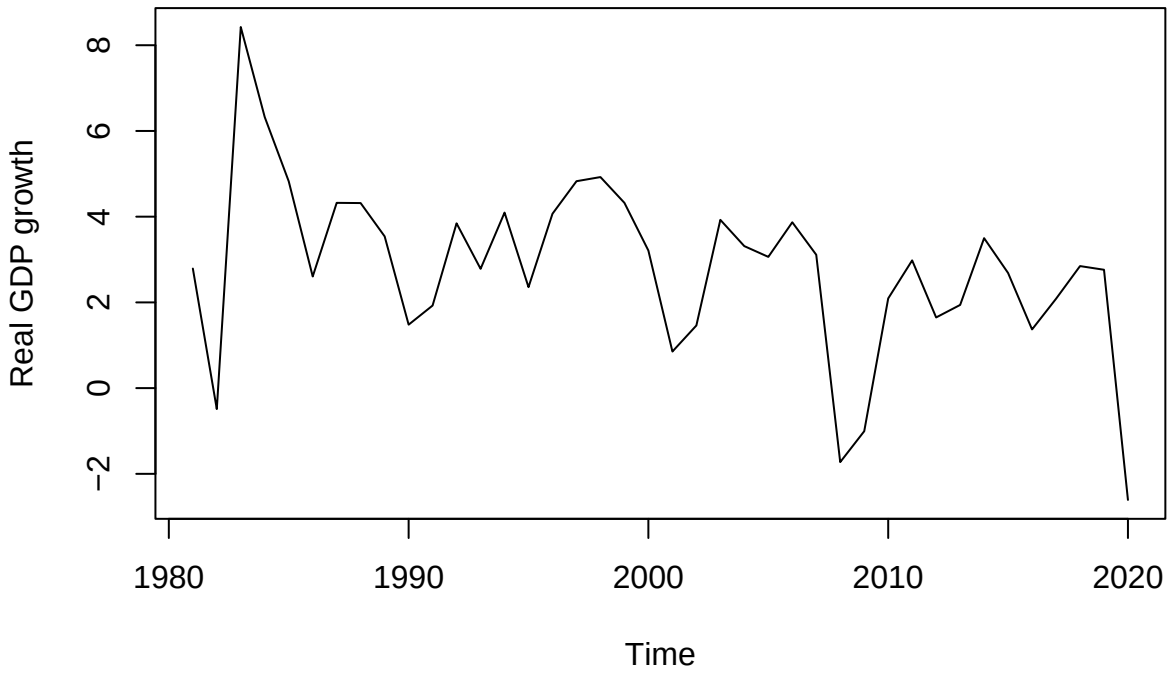
Federico Vicentini

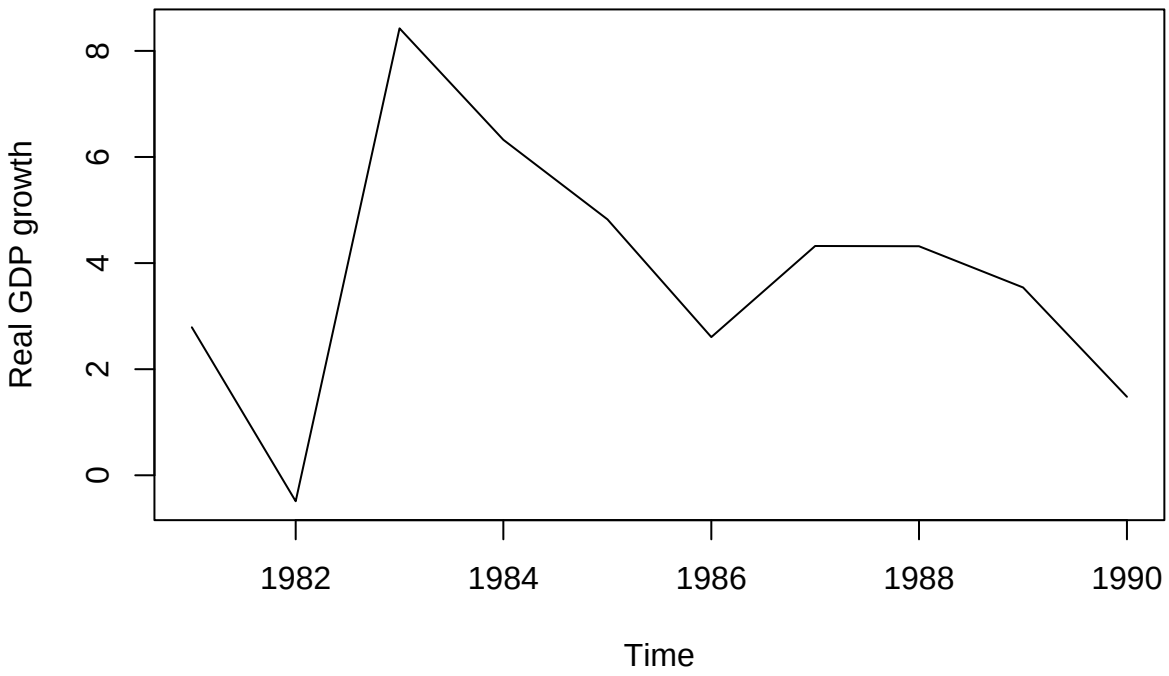
01/05/2023

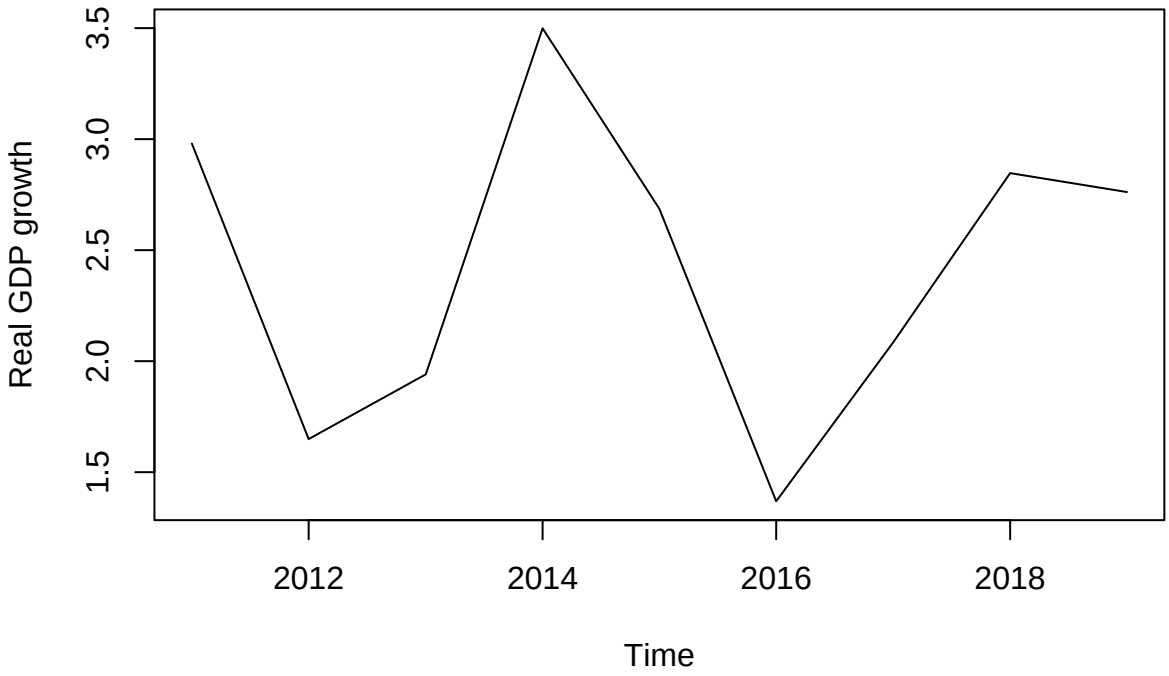
DATA COLLECTION

Real GDP Data

In this first part, we download nominal GDP data and GDP deflator data from 1980 to 2020 and then we divide it into the 2 section (1980s and 2010s) in order to calculate means and thus compare it to the data we got in the TTD Presentation.



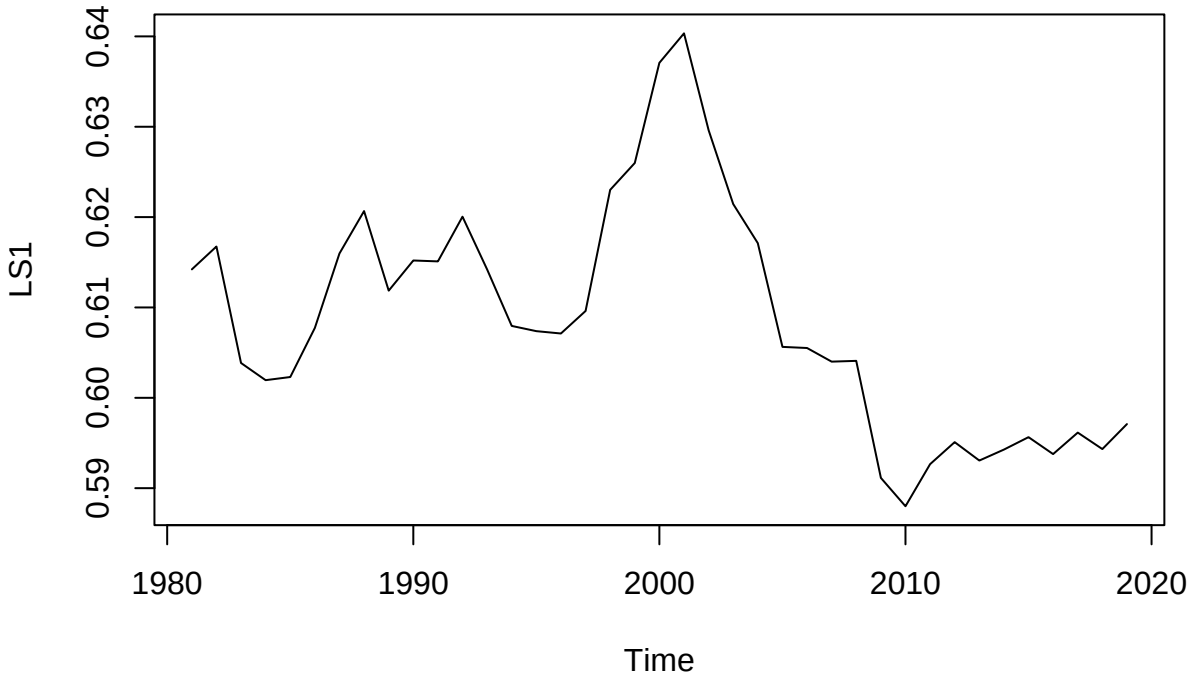

Real GDP growth
0
2
4
6
8
1982
1984
1986
1988
1990
Time

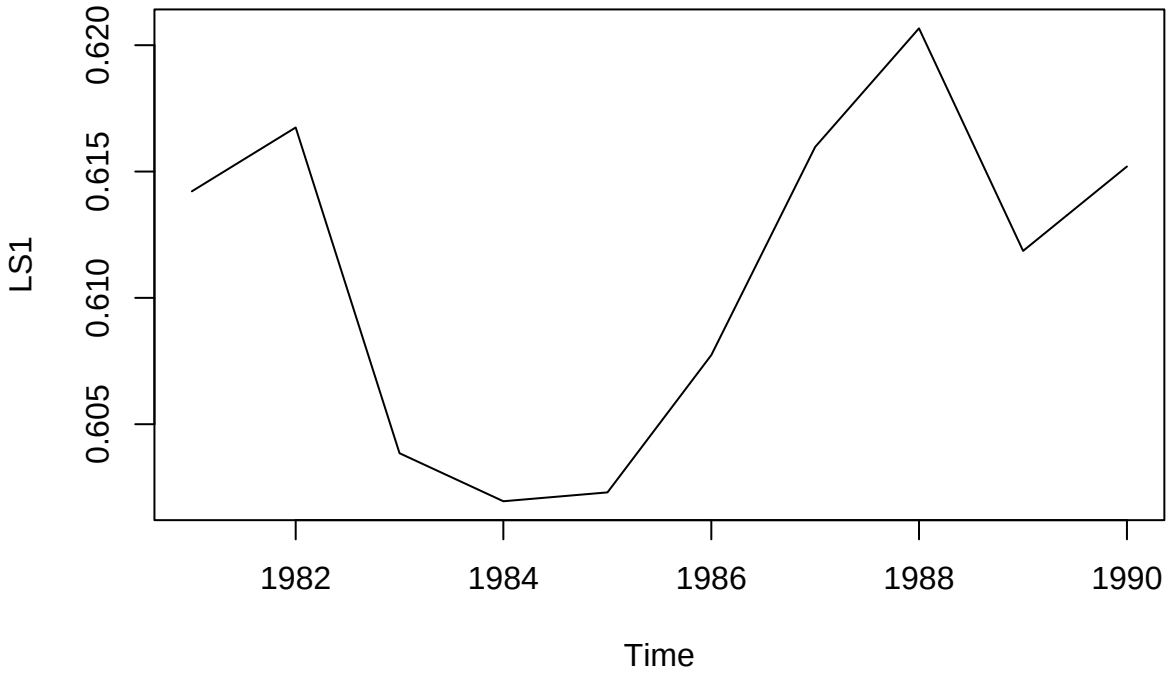


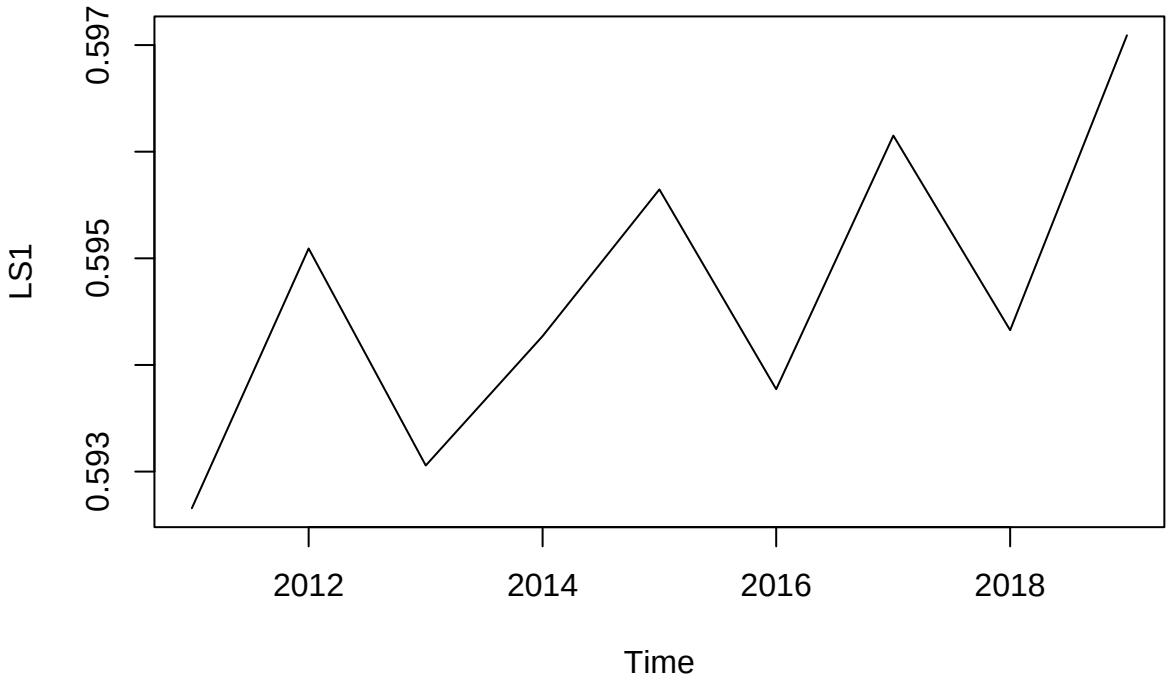
Labor Share of Output

"Naive" Labor Share (LS1)

First attempt here is to download the labor share timeseries from fred. You can see from the values of the plots that the level of the labor share is underestimated if we use LS1. In fact, this approach neglects some important aspects, such as the role of self-employed workers and correction to value added in the forms of indirect taxes and consumption of fixed capital.


LS1
0.59
0.60
0.61
0.62
0.63
0.64
1980
1990
2000
2010
2020
Time


LS1
0.605
0.610
0.615
0.620
1982
1984
1986
1988
1990
Time



Guerriero Index

This is precisely why we tried to replicate the labor share measure provided by {Guerriero (2019)}, defined as:

$$LS6 = \frac{compensation\ of\ employees * \left(\frac{workforce - employers}{employees}\right)}{value\ added - ind.\ taxes - fixed\ cap.\ cons.}$$

Data is retrieved from Ilostat R package API and (United Nations System of National Accounts (2023))

For lack of available data on the structure of the labor force in the US prior to the introduction of the ICSE93 standard, we will use LS5 instead of the LS6 index, defined as:

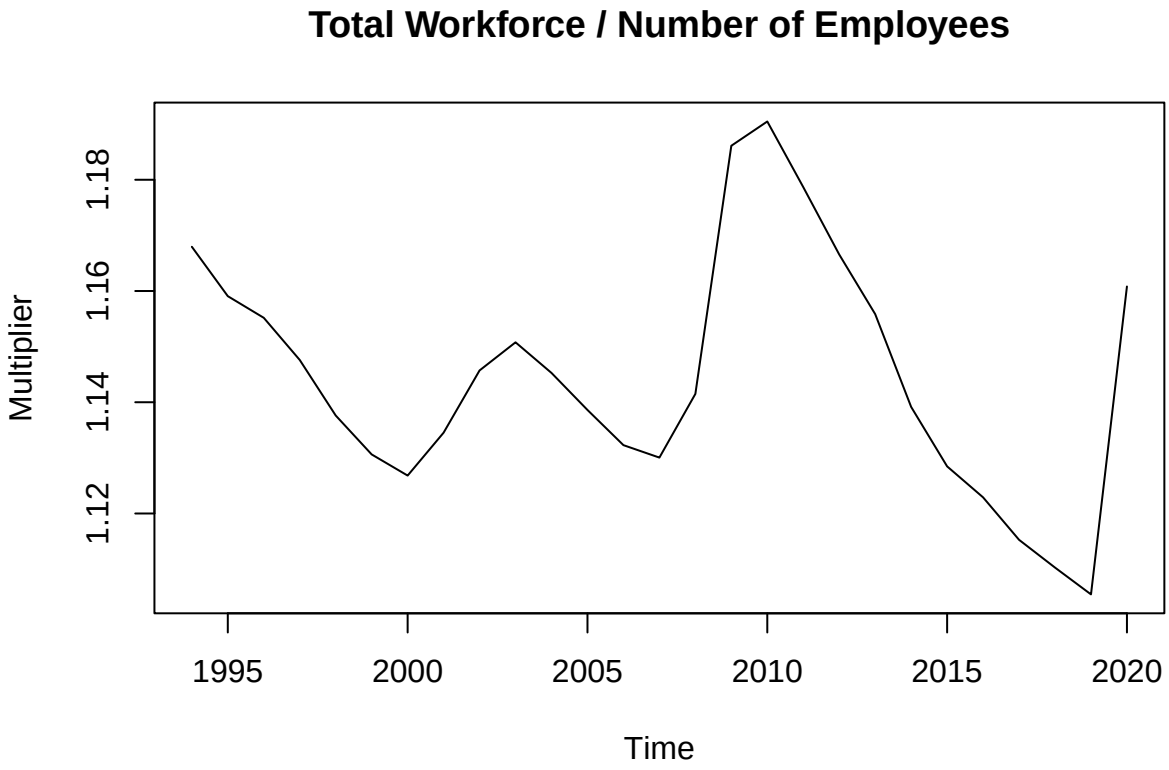
$$LS6 = \frac{compensation\ of\ employees * \left(\frac{workforce}{employees}\right)}{value\ added - ind.\ taxes - fixed\ cap.\ cons.}$$

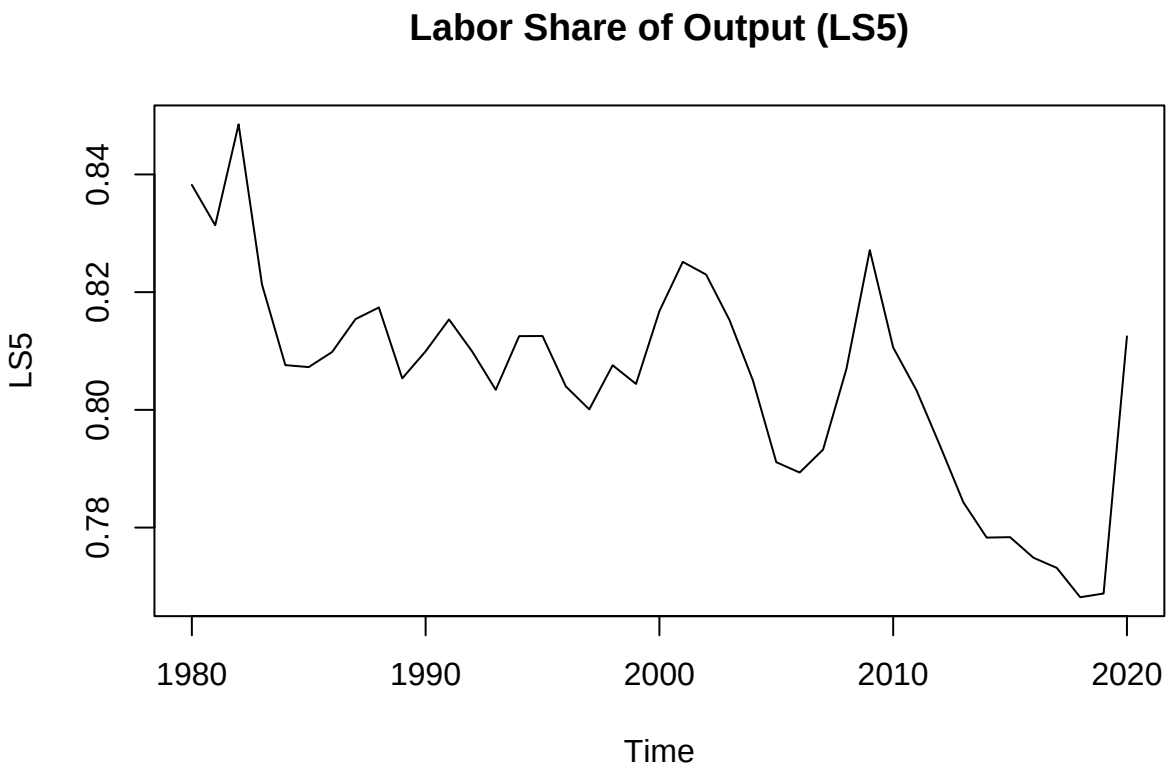
Below you can find also an example of the result we would get by trying to compute LS6 too.

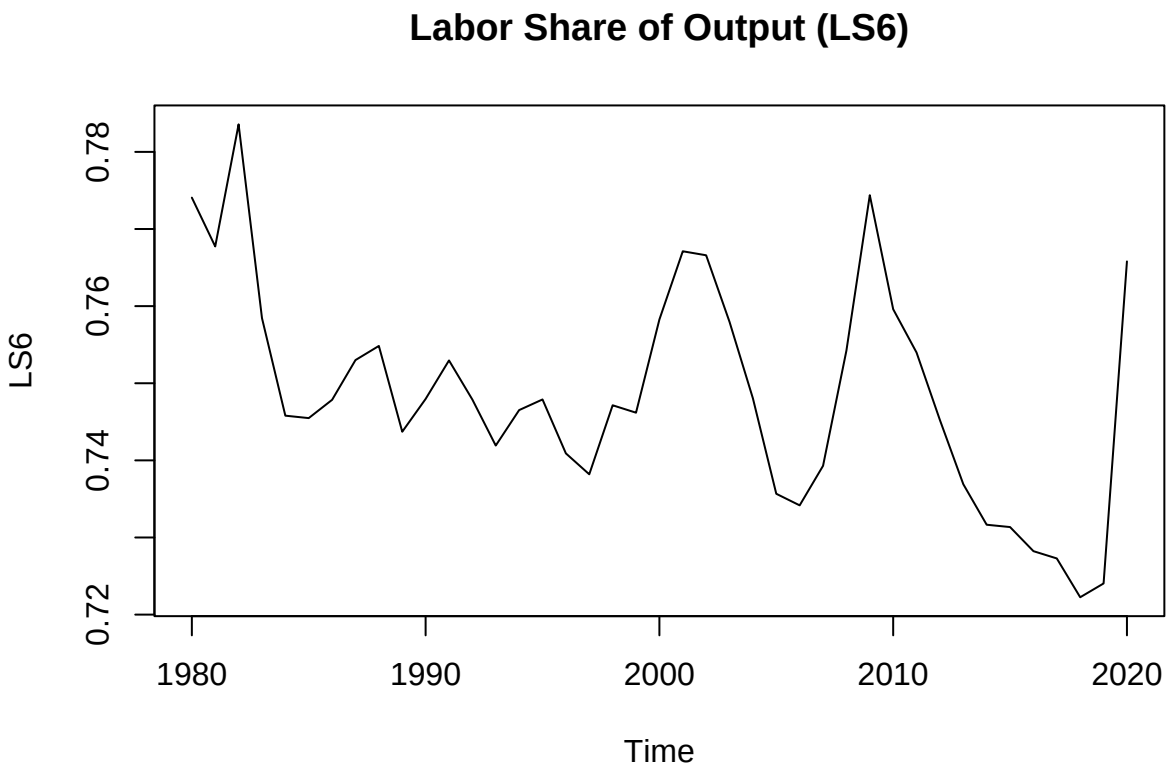
Note: since (Guerriero (2019)) says that "Data on the composition of the workforce is not always available for every year. When absent, it is assumed to be the same as in the previous year (Gollin 2002). This is a realistic assumption (Askenazy 2003), given that the composition of the workforce is relatively constant over time.", and since we have data on the composition of the workforce from 1994 onward, we will assume that in years preceding 1994, workforce composition is the same. Below, you can find a plot of workforce/employees for 1994 onward, where you can see that the figure is pretty stable. However, instead of taking 1994 value

and projecting backward, I found it better to make an average for the period 1994-2000 and then project that figure backward.

Note2: since there are 2 different time series for SNA data, coming from two different accounting standards (SNA93 and SNA08), we tried 2 approaches. First, we tried to find a conversion factor between those 2 standards (using first lags in order not to have spurious regressions), but this approach leads to a big jump in the data around the cutoff year. Thus, the second method was just to keep the newer standard for years after 1995, and the old one for years prior. This is probably what Guerriero (2019) did in the paper, since the results look similar (albeit different since we use LS5 and not LS6). This is the method I selected in the end.

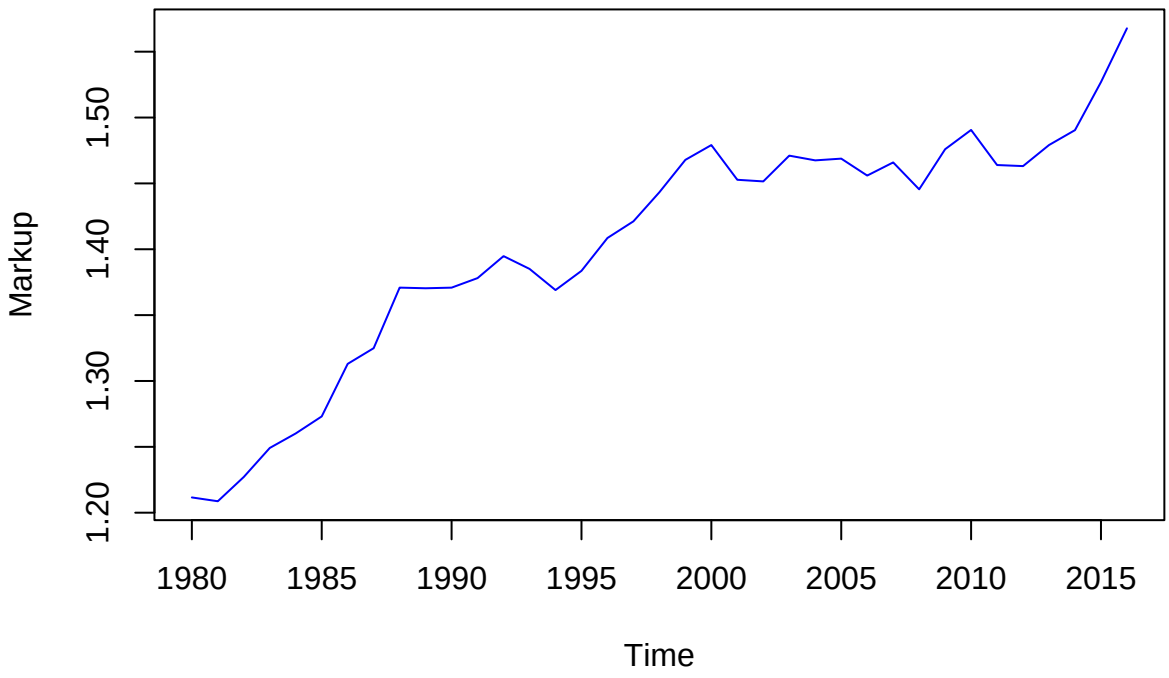



Labor Share of Output (LS5)
LS5
0.78
0.80
0.82
0.84
1980
1990
2000
2010
2020
Time



Price Markup

Price markup is computed according to De Loecker, Eeckhout, and Unger (2020), using the replication file related to the paper and extracting the time series of aggregate markup presented as the main finding of the paper itself.



Market Concentration

Here we present a slideshow of graphs of various concentration measures:

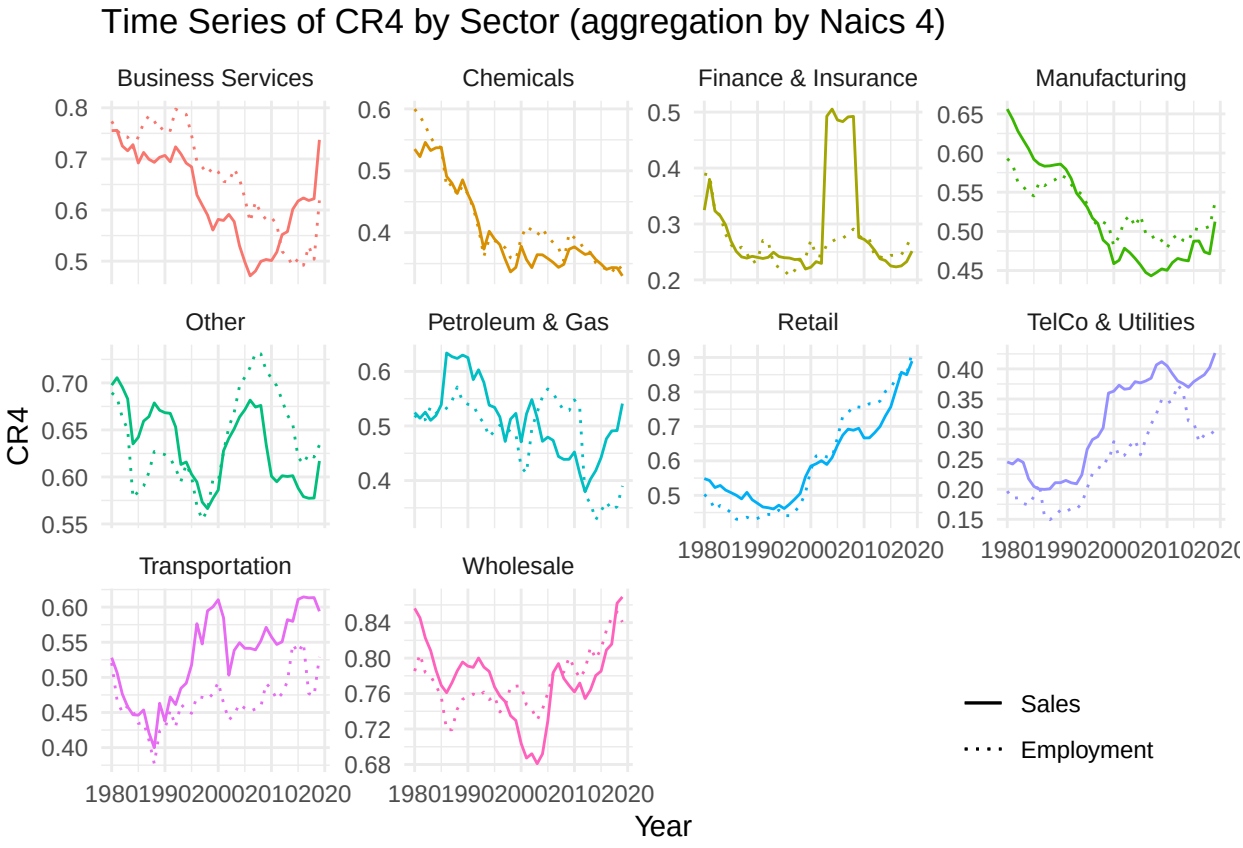
- CR4
- CR20
- Herfindal-Hirschman (HH) Index

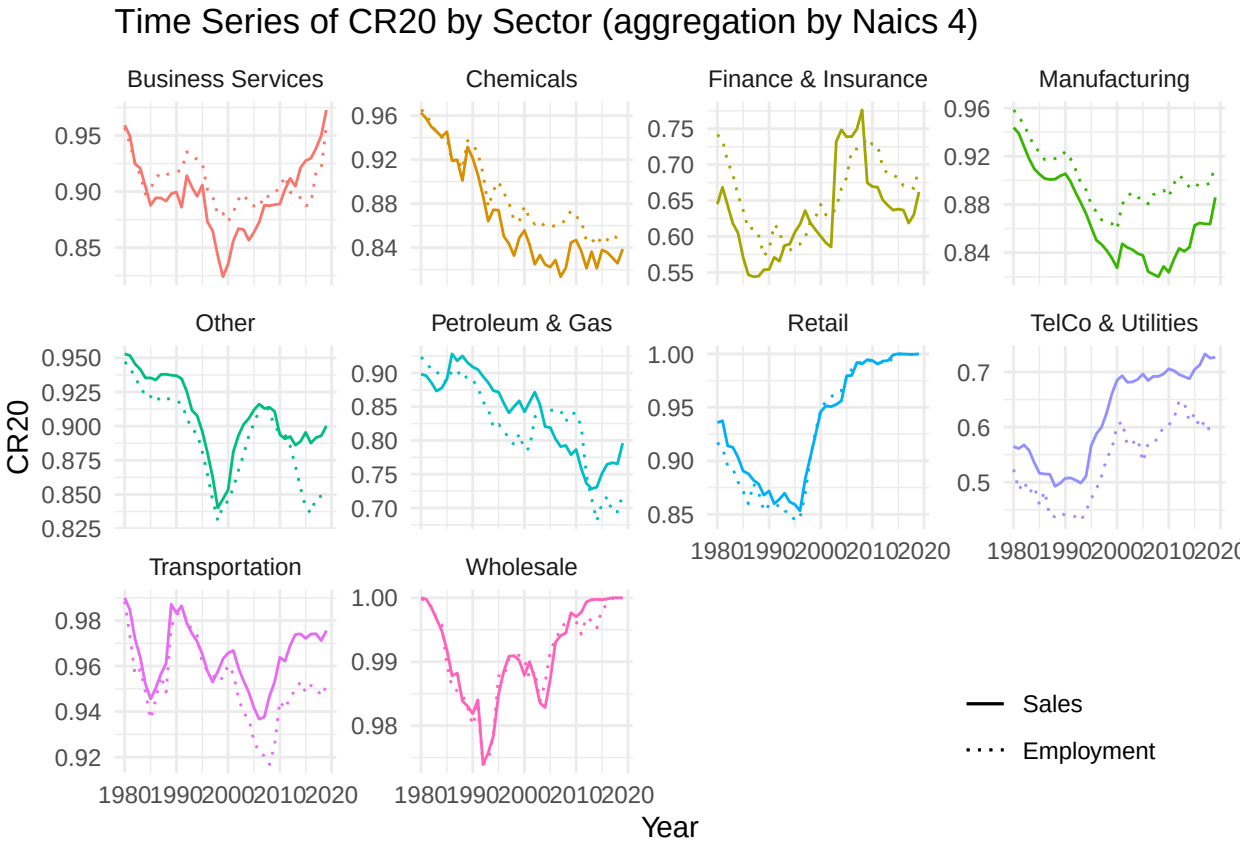
All measures are computed using Compustat firm-level data, so they rely on data exclusively on US publicly traded firms. Compustat data has already been cleaned in Stata (the file is saved as "compustatclean.do"). Concentration measures are computes firstly at the Naics 4-digit level, then at the 3-digit one. Aggregation by sector is done according to the sectorization provided in Autor et al. (2020). For every SIC code employed by Autor, we tracked down the corresponding Naics code and replicated the procedure. In the end, the best results came from CR4 and HHI computed at the Naics 3 level. Nonetheless, the upward trend is not that clear in all sectors: that could be due to multiple factors:

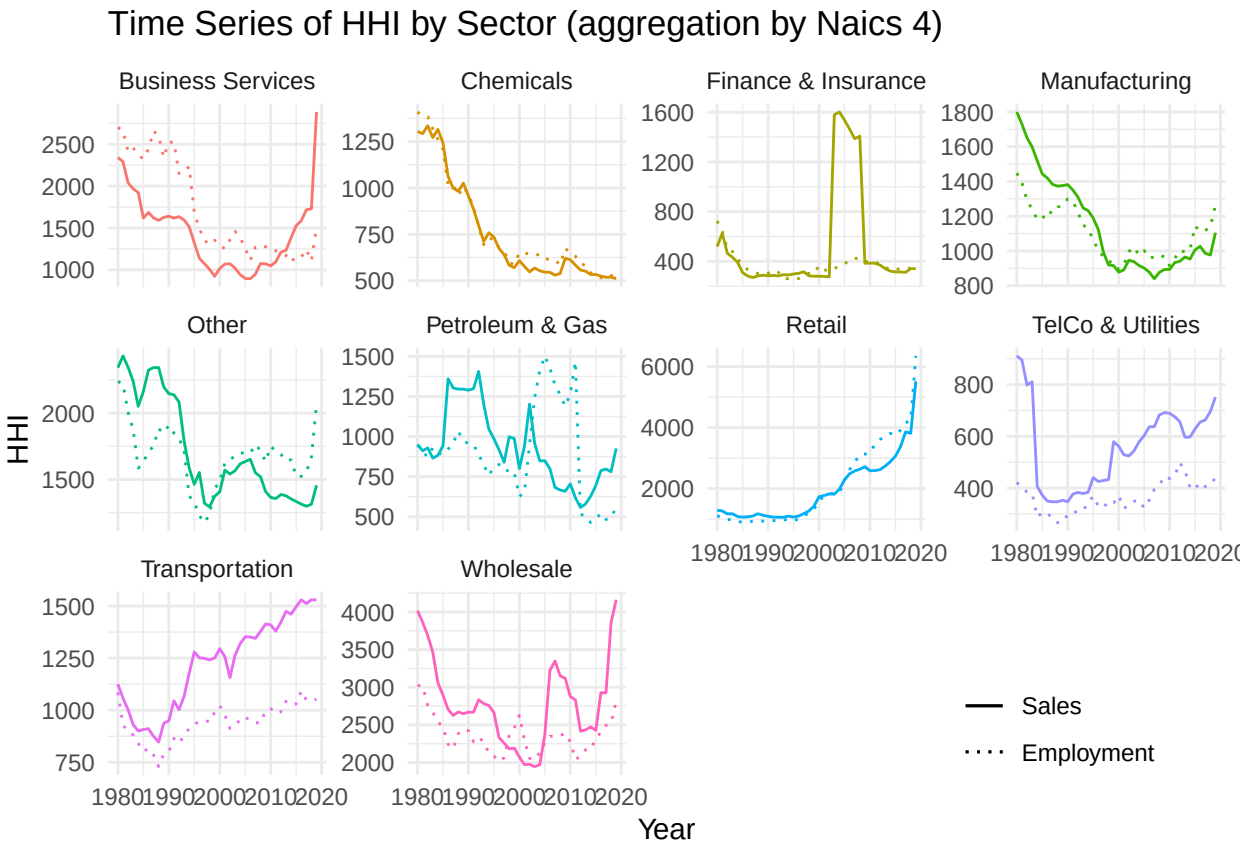
- We deal only with publicly traded firms, thus it is possible that concentration measures computed are higher, since the denominator is actually smaller than the real size of them corresponding industrial sector if privately owned companies are accounted for.
- The second reason is related to the first one. Since in the 1980s the stock market was smaller than today, both in absolute value and as a share of the total economy, it is possible that the concentration measures are biased upward more at the start of the time series than they are at the end.

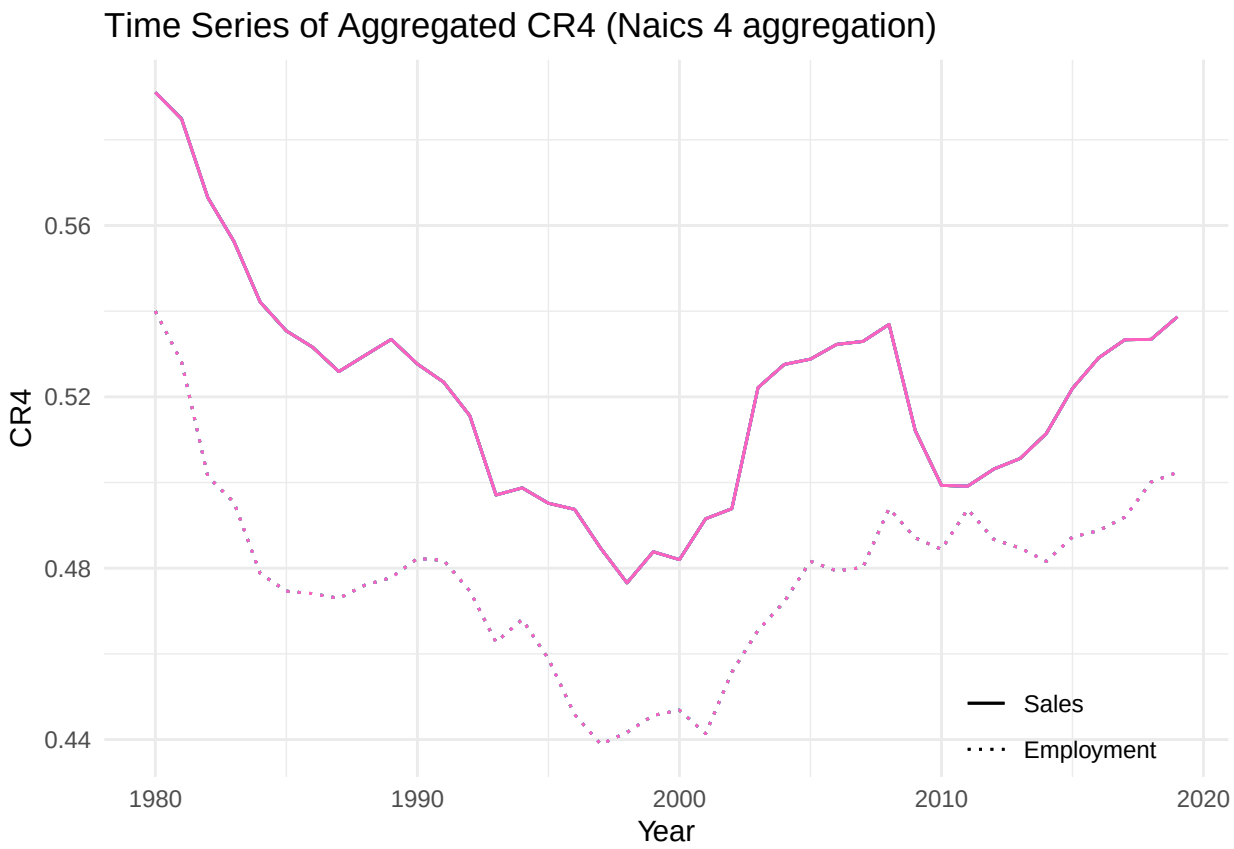
- Specifically when using highly detailed sectorization like naics4, in our sample the population of firms in the sector shrinks down to even less than 4 in some cases. It is clear that in those cases, the concentration measure is equal to 1, thus biasing our weighted mean upward. Computing measures at the naics3 level partially accounts for this bias, but not for all sectors, since economic sectors like agriculture are not widely represented in the sample of publicly traded firms.

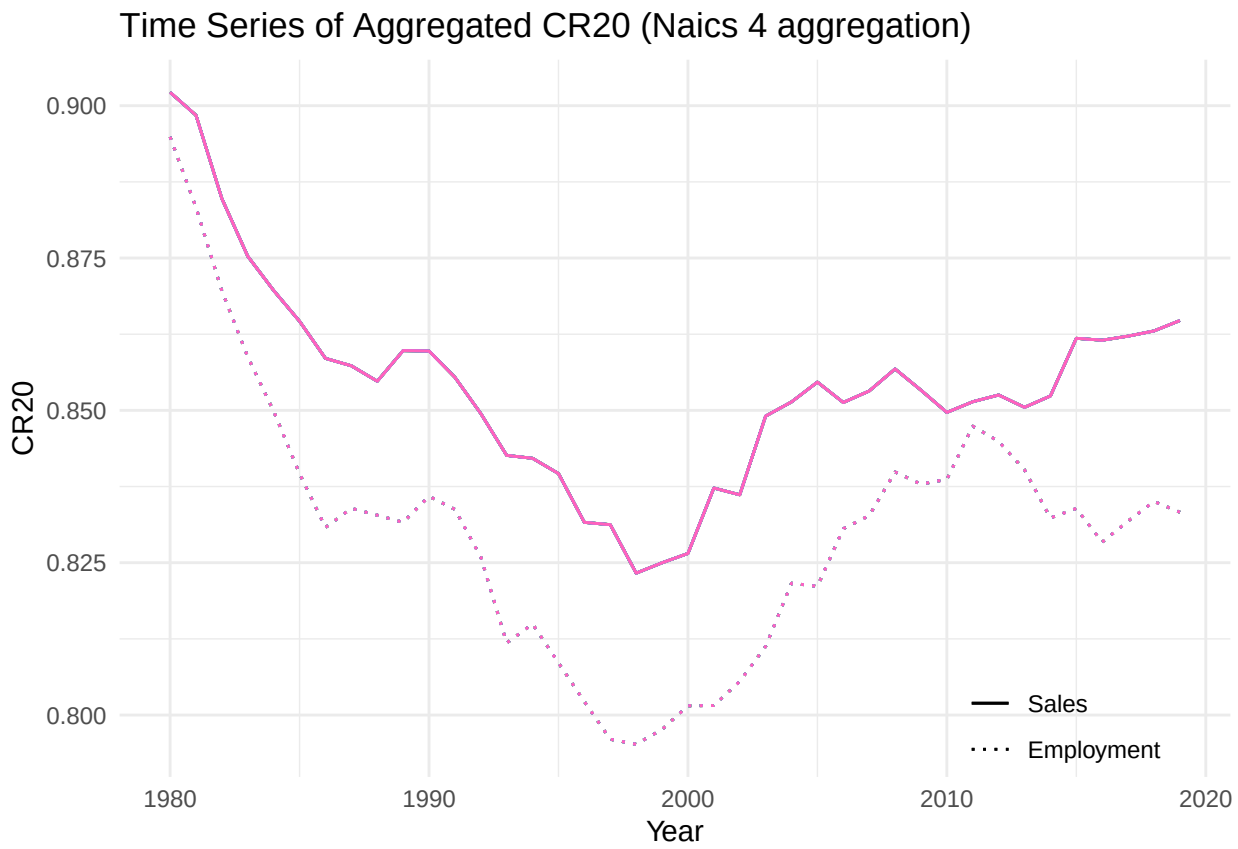
These are just some of the possible reasons why our upward trend seem to be not so strong in every sector, and not so strong even at the aggregated level for some of the indices.

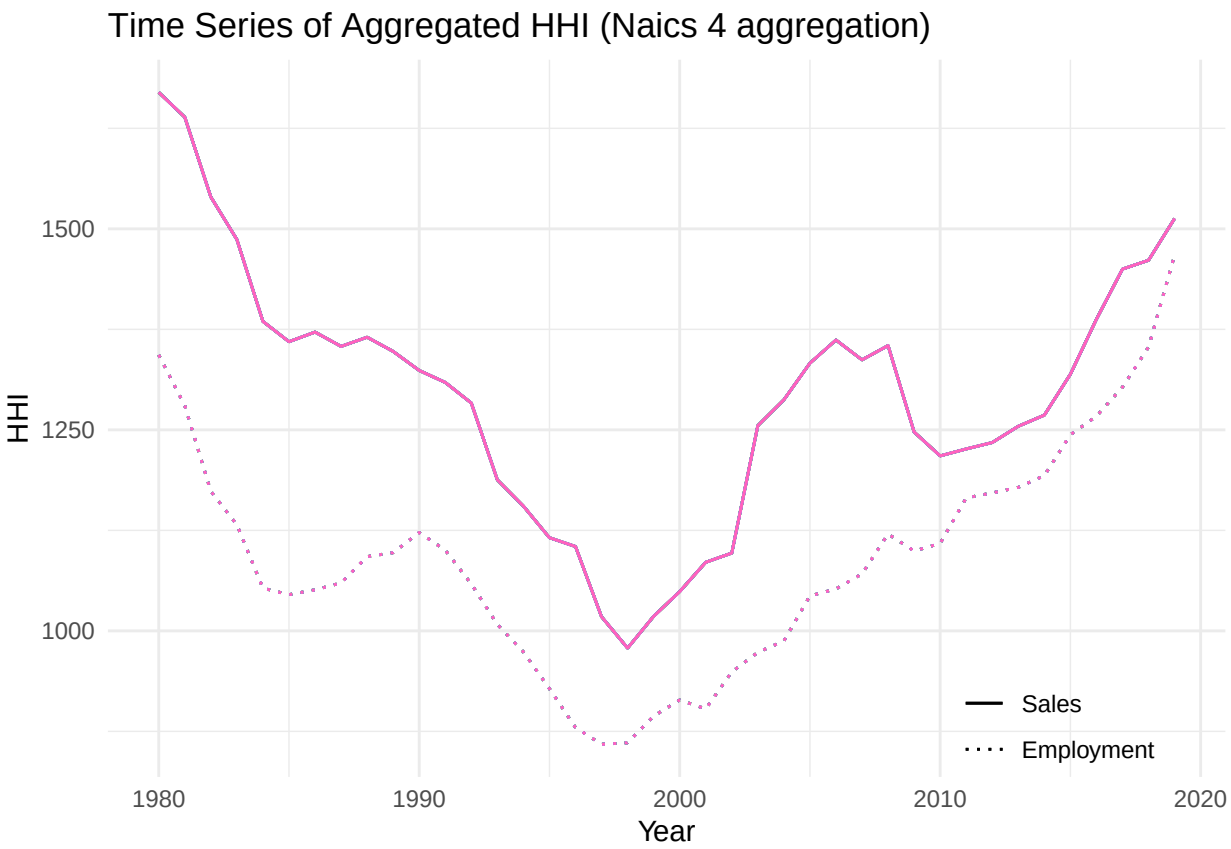


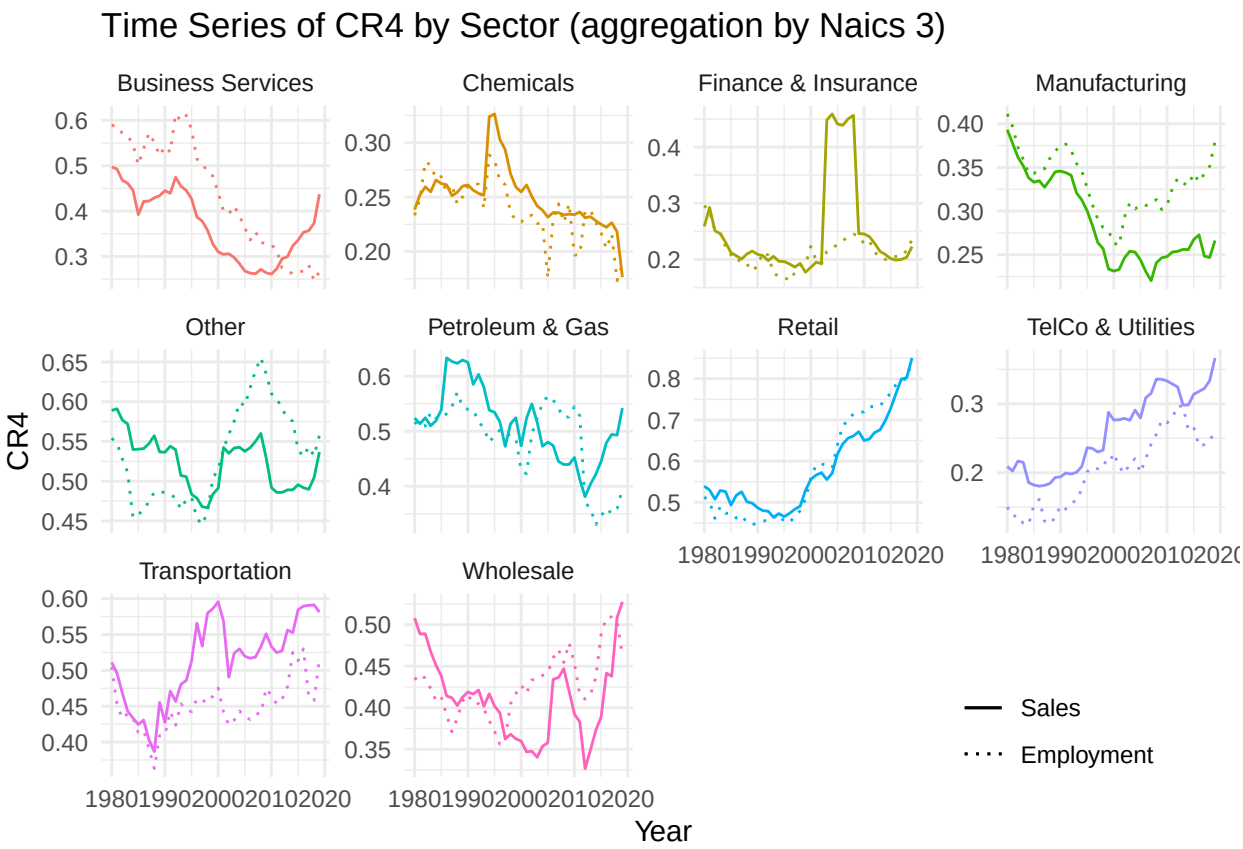

Time Series of CR20 by Sector (aggregation by Naics 4)
Business Services
Chemicals
Finance & Insurance
Manufacturing
Other
Petroleum & Gas
Retail
TelCo & Utilities
Transportation
Wholesale
CR20
Year
Sales
Employment

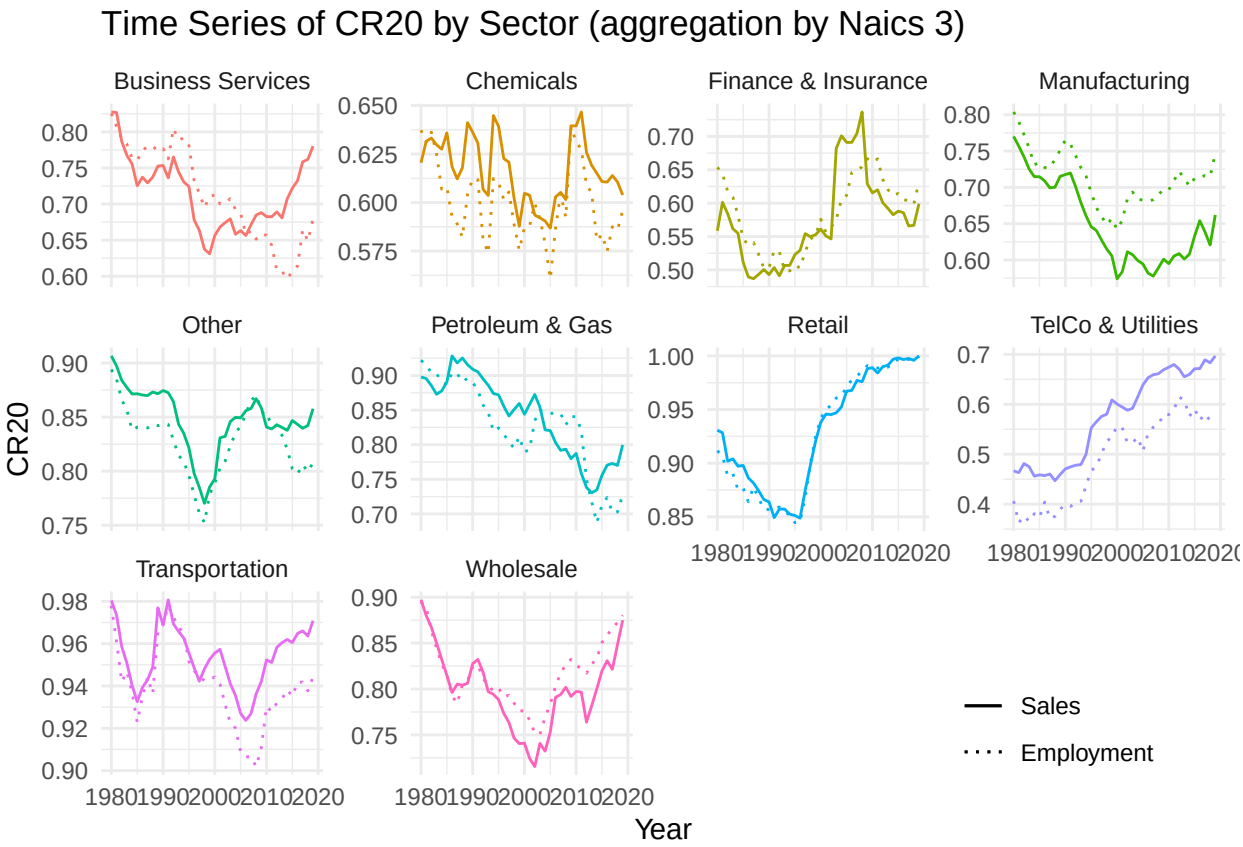

Time Series of HHI by Sector (aggregation by Naics 4)
Business Services
Chemicals
Finance & Insurance
Manufacturing
Other
Petroleum & Gas
Retail
TelCo & Utilities
Transportation
Wholesale
HHI
Year
Sales
Employment

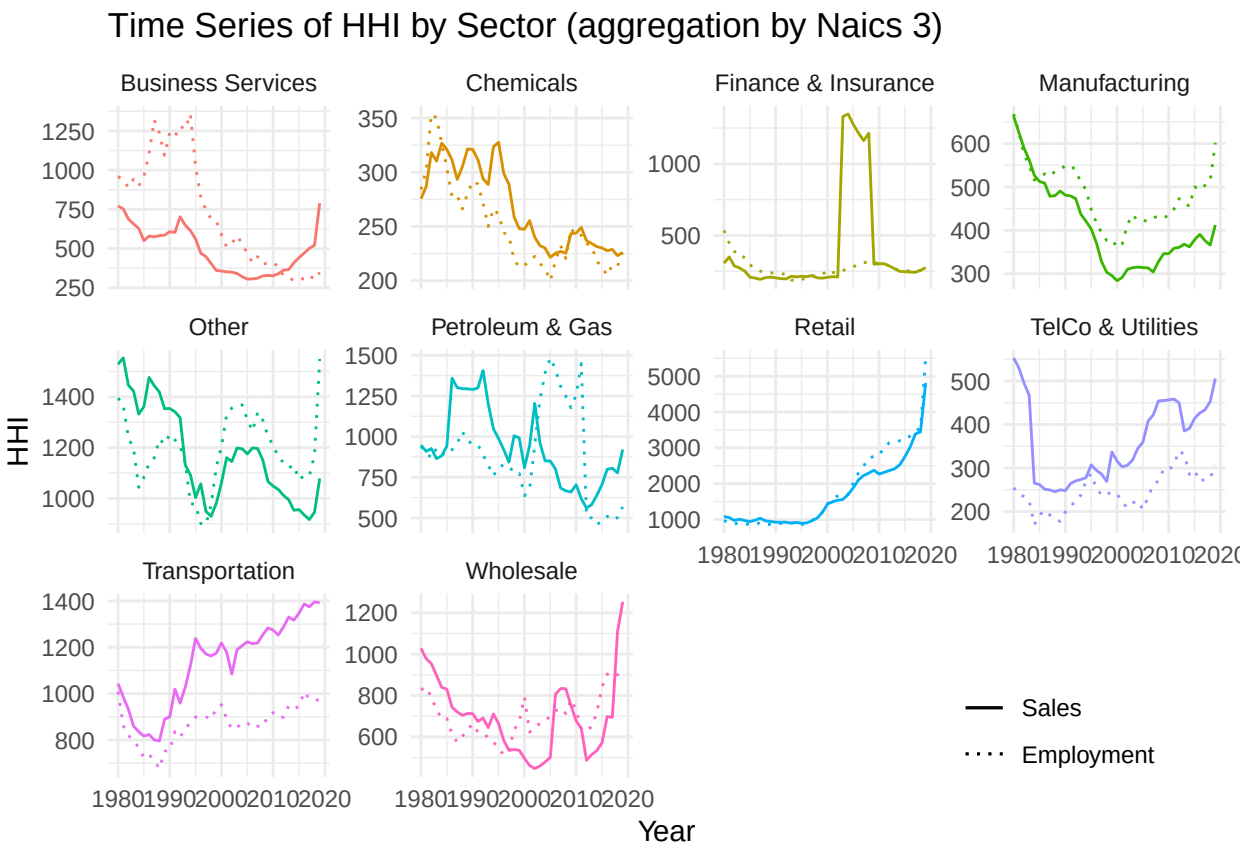

Time Series of Aggregated CR4 (Naics 4 aggregation)
CR4
0.56
0.52
0.48
0.44
1980
1990
2000
2010
2020
Year
Sales
Employment

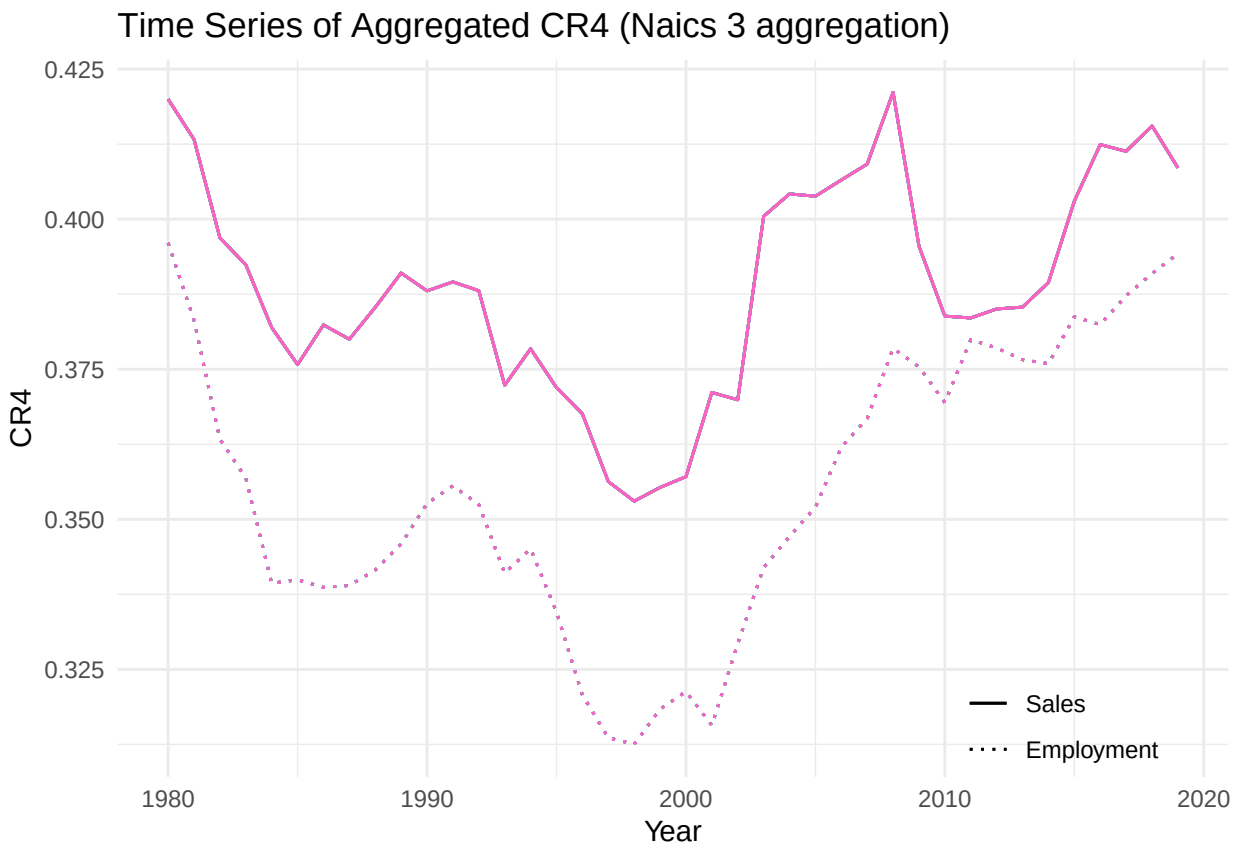

Time Series of Aggregated CR20 (Naics 4 aggregation)
CR20
0.900
0.875
0.850
0.825
0.800
1980
1990
2000
2010
2020
Year
Sales
Employment

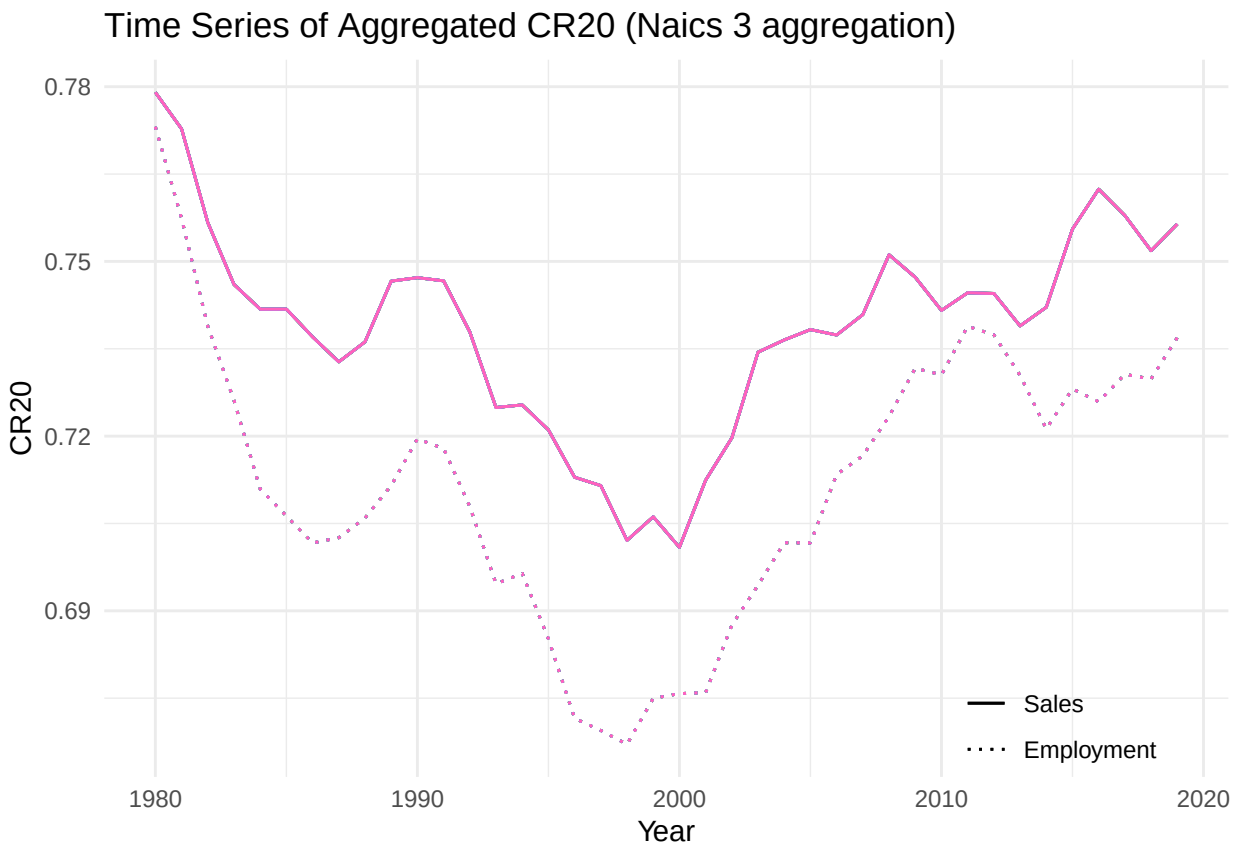

Time Series of Aggregated HHI (Naics 4 aggregation)
HHI
1500
1250
1000
1980
1990
2000
2010
2020
Year
Sales
Employment

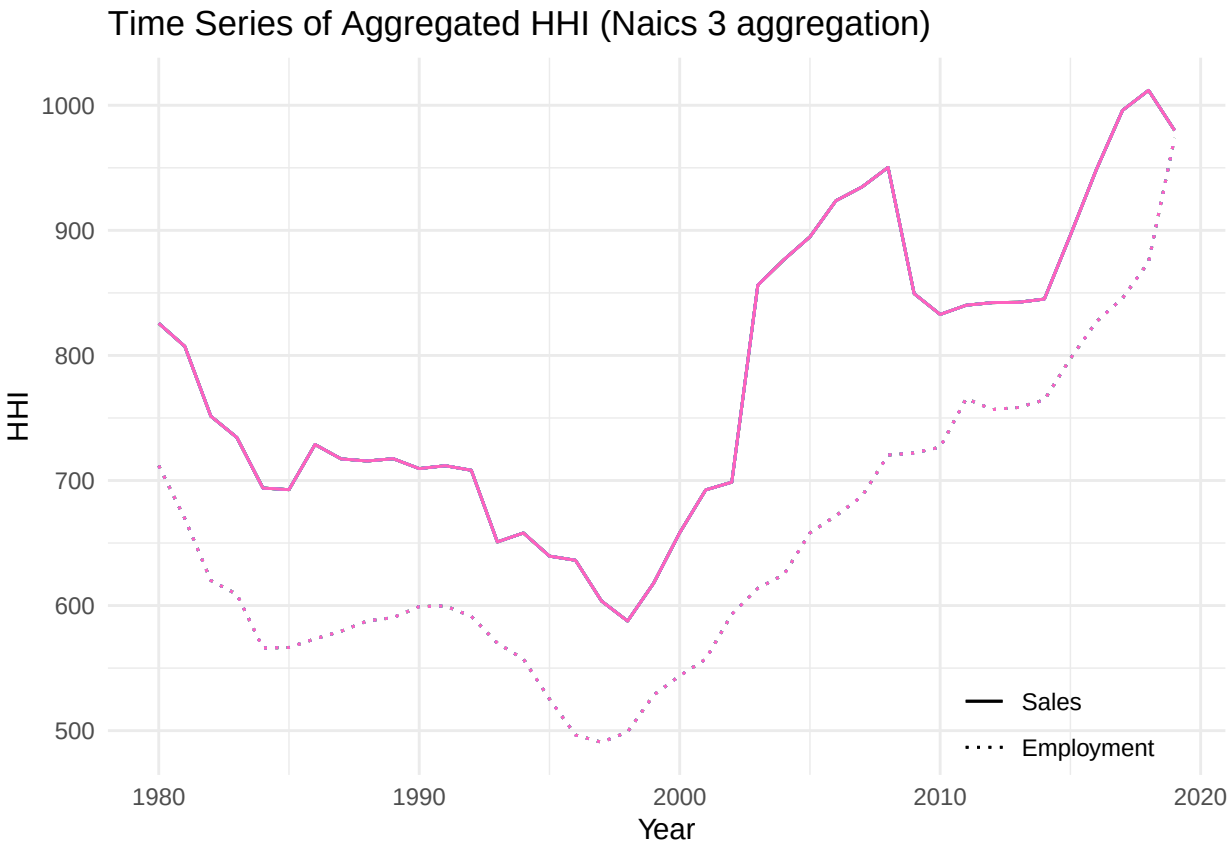

Time Series of CR4 by Sector (aggregation by Naics 3)
Business Services
Chemicals
Finance & Insurance
Manufacturing
Other
Petroleum & Gas
Retail
TelCo & Utilities
Transportation
Wholesale
CR4
Year
Sales
Employment


Time Series of CR20 by Sector (aggregation by Naics 3)
Business Services
Chemicals
Finance & Insurance
Manufacturing
Other
Petroleum & Gas
Retail
TelCo & Utilities
Transportation
Wholesale
CR20
Year
Sales
Employment


Time Series of HHI by Sector (aggregation by Naics 3)
Business Services
Chemicals
Finance & Insurance
Manufacturing
Other
Petroleum & Gas
Retail
TelCo & Utilities
Transportation
Wholesale
HHI
Year
Sales
Employment


Time Series of Aggregated CR4 (Naics 3 aggregation)
0.425
0.400
0.375
0.350
0.325
CR4
1980
1990
2000
2010
2020
Year
Sales
Employment


Time Series of Aggregated CR20 (Naics 3 aggregation)
0.78
0.75
0.72
0.69
CR20
1980
1990
2000
2010
2020
Year
Sales
Employment



Finalised Output

In this table you can find reported all the average values for the two time frames analyzed (1980s and 2010s) for each variable considered. I reported HHI too for concentration, but I do not know what is the computed HHI for the model, so, i left it blank.

Table 1: Empirical and Simulated Data

Variable	E.1980s	E.2010s	E.Delta	S.1980s	S.2010s	S.Delta
GDP growth	3.81	2.42	-1.39	3.97	2.28	-1.69
CR4	35.00	38.30	3.32	7.73	17.30	9.54
HHI	596.00	818.00	222.00	0.00	0.00	0.00
Markup	1.30	1.50	0.20	1.33	1.47	0.15
Labor share	75.50	73.30	-2.14	78.00	74.10	-3.82

References

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *The Quarterly Journal of Economics* 135 (2): 645–709. https://doi.org/10.1093/qje/qjaa004.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *The Quarterly Journal of Economics* 135 (May). https://doi.org/10.1093/qje/qjz041.

Guerriero, Marta. 2019. "The Labor Share of Income Around the World: Evidence from a Panel Dataset." *ADBI Working Papers*. https://doi.org/10.2139/ssrn.3358006.

United Nations System of National Accounts. 2023. "UN - SNA." http://data.un.org/Data.aspx?d=SNA&f=group_code%3a401.