Libraries

In [ ]:
```
import importlib

import config as Config
import constants.labels as Labels
import hangar.FirmBomberFactory as FirmBomberFactory
import hangar.FundBomberFactory as FundBomberFactory
import hangar.DisplayModule as DisplayModule
import models.firm as FirmModel
import models.funding as FundingModel
import models.investor as InvestorModel
import models.macro as MacroModel
import utils.file as FileUtils
import utils.industry as IndustryUtils
import utils.visualiser as Visualiser

def reload():
  importlib.reload(Config)
  importlib.reload(Visualiser)
  importlib.reload(DisplayModule)
```

Configuration

Configuration specified in src/config.js to avoid passing the same configurations around. Modify configurations as required, then run reload() to reload configurations without restarting the environment.

Load Data

Data obtained from Crunchbase and CEIC (for macro data), preprocessed, then stored in pickle files. This section retrieves the stored data and does additional preprocessing.

Load Investor Data

In [ ]:
```
# Read Investor Data
investors = FileUtils.read_pickle(f"investors")
```

In [ ]:
```
# Get Public Funded Investors
public_funded_investors = InvestorModel.get_public_funded_investors(investor
```

Load Company Data

In [ ]:
```
# Read Company Data
firms = FileUtils.read_pickle(f"{Config.country}_firms")
```

In [ ]:
```
# Read Domain Data
domain_created_year = FileUtils.read_pickle('domain_created_year')
```

In [ ]:
```
# Filter and Enrich Company Data
firms = FirmModel.filter_for_profit(firms)
firms = FirmModel.enrich_founded_year(firms, domain_created_year)
firms = FirmModel.filter_founded_year(firms, Config.start_year, Config.end_y
firms = FirmModel.enrich_public_funded(firms, public_funded_investors)
firms = firms.reset_index(drop=True)
```

Load Funding Data

In [ ]:
```
# Read Funding Data
funding = FileUtils.read_pickle(f"funding")
```

In [ ]:
```
# Filter and Enrich Funding Data
funding = FundingModel.filter_announced_year(funding, Config.start_year, Con
funding = FundingModel.enrich_public_funded(funding, public_funded_investors
```

Load Macro Data

In [ ]:
```
# Read Macro Data
real_gdp = FileUtils.read_pickle('real_gdp')
```

In [ ]:
```
# Filter Macro Data
real_gdp = real_gdp['United States']
real_gdp = real_gdp[real_gdp.index >= Config.start_year]
real_gdp = real_gdp[real_gdp.index < Config.end_year]
```

In [ ]:
```
# Store Macro Data
macro = MacroModel.Macro(real_gdp)
```

Load Other Data

In [ ]:
```
# Industry Labels
industry_groups = IndustryUtils.get_industry_groups()
industries = IndustryUtils.get_industries()
```

Preprocess Data

In [ ]:
```
# Load Bombers
firm_bomber_i = FirmBomberFactory.FirmBomber(firms, targets=industries, targ
firm_bomber_ig = FirmBomberFactory.FirmBomber(firms, targets=industry_groups
```

In [ ]:
```
# Load Bombers
fund_bomber_i = FundBomberFactory.FundBomber(funding, targets=industries, ta
fund_bomber_ig = FundBomberFactory.FundBomber(funding, targets=industry_grou
```

Research Question: "Trend-Following"

Have company founders and investors become more prone to trend-following over time, rather than relying on their own judgement? Trend-following results in bubbly behaviour, with sharp spikes and sharp drops, likely cause issues as bad causes are funded in the bubble and good causes are not funded in the crash. Note that a bubble can be defined as "trend-following" or as "deviation from fundamentals"; it's possible to follow a trend and by coincidence approach the fundamentals, though this likely does not hold true in the long run. I will use the "trend-following" definition, since "fundamental" are difficult to define, and understanding trend-following behaviour is still is empirically useful.

I use a year range of 1960 to 2019 as that is what is available on both Crunchbase and CEIC (for macro data), as well as to avoid what looks like incomplete data post 2019. There are dramatically few companies being reported to be founded around that time period, which may be partially attributed to the Coronavirus pandemic, but is still somewhat suspect.

Companies: Aggregate Rolling AR(1)

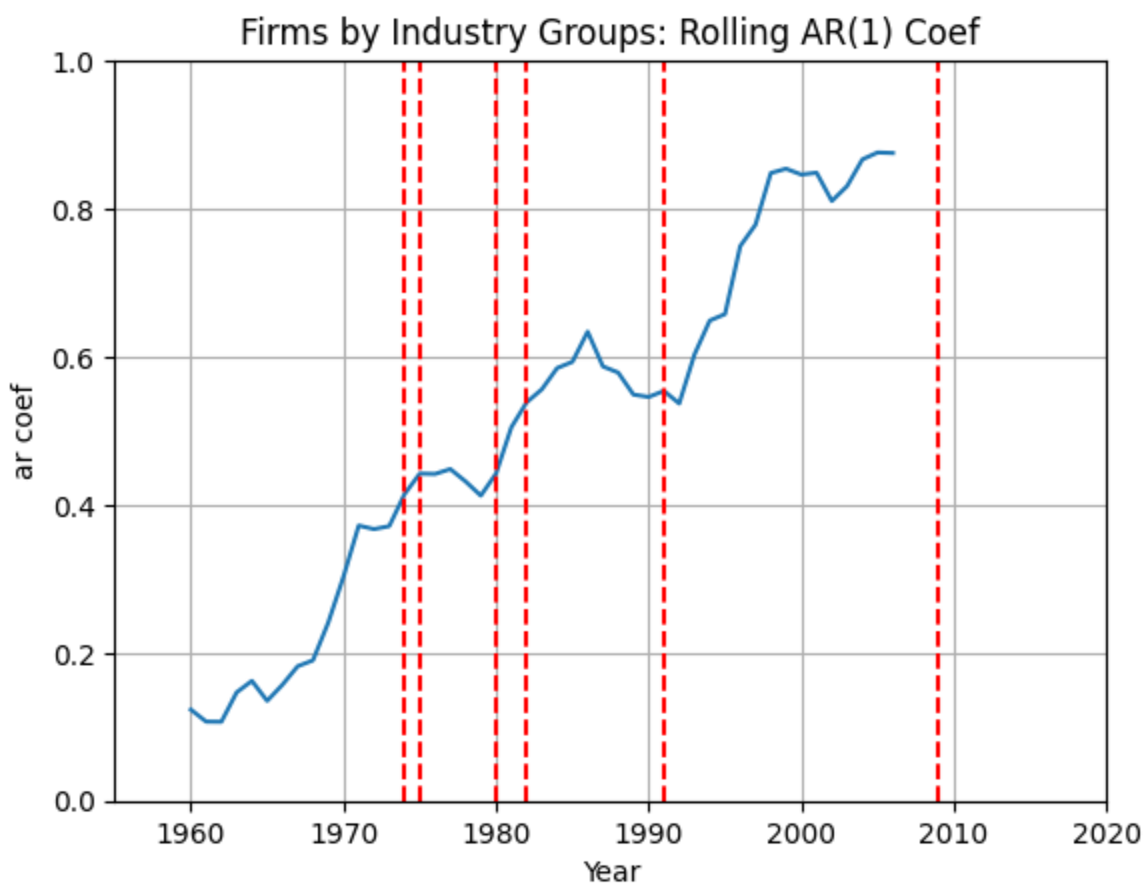
Rolling AR(1) calculated from % share of companies of a category **founded** in a year / total companies in that year. % share is used to avoid confounding by real gdp and interest rates, and also to avoid having to define an e.g. panel model using real gdp and interest rate as controls, since the relationship between real gdp and interest rate is very messy. A rolling window allows us to observe changes over time. There is a clear trend of increasing "trend-following" since the 1960s, that is robust to grouping by industries or industry groups, though there may be other composition effects.

For all subsequent analysis, thresholds have been applied before a particular category/ year is included in the analysis, to avoid issues witih small data sets. These thresholds, as well as the window size used, can be found in the configurations.

Red lines are recessions, defined as negative annual gdp growth.

In [ ]:
```
# Uncomment reload to make changes to the code without having to restart the
# reload()
firm_bomber_i=FirmBomberFactory.FirmBomber(refurb=firm_bomber_i)
firm_bomber_i.identify()
firm_bomber_i.report(macro)

firm_bomber_ig=FirmBomberFactory.FirmBomber(refurb=firm_bomber_ig)
firm_bomber_ig.identify()
firm_bomber_ig.report(macro)
```



Funding: Aggregate Rolling AR(1)

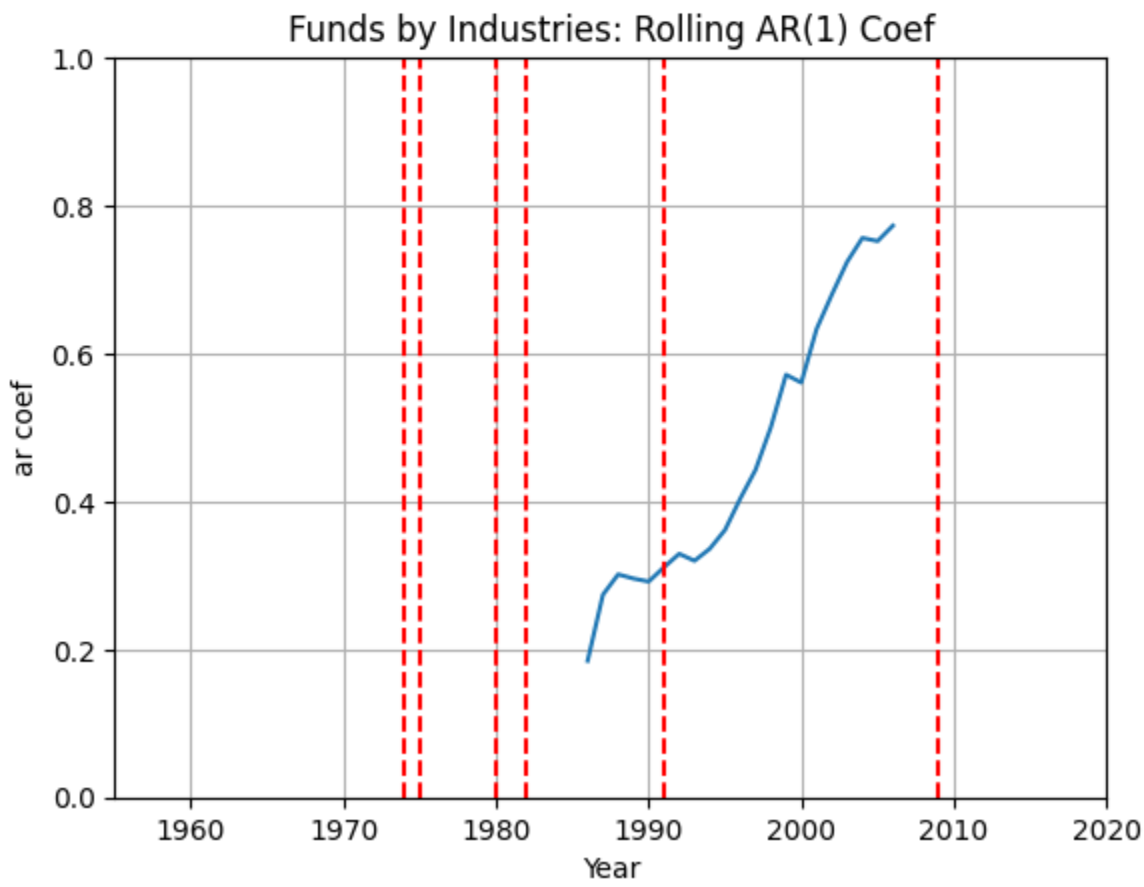
Rolling AR(1) calculated from % share of funding rounds of a category **announced** in a year / total funding rounds in that year. The time range is more restricted for funding data than company data, starting only around the 1980s. There is also clear trend of increasing "trend-following" since the 1960s, that is robust to grouping by industries or industry groups, though there may be other composition effects.

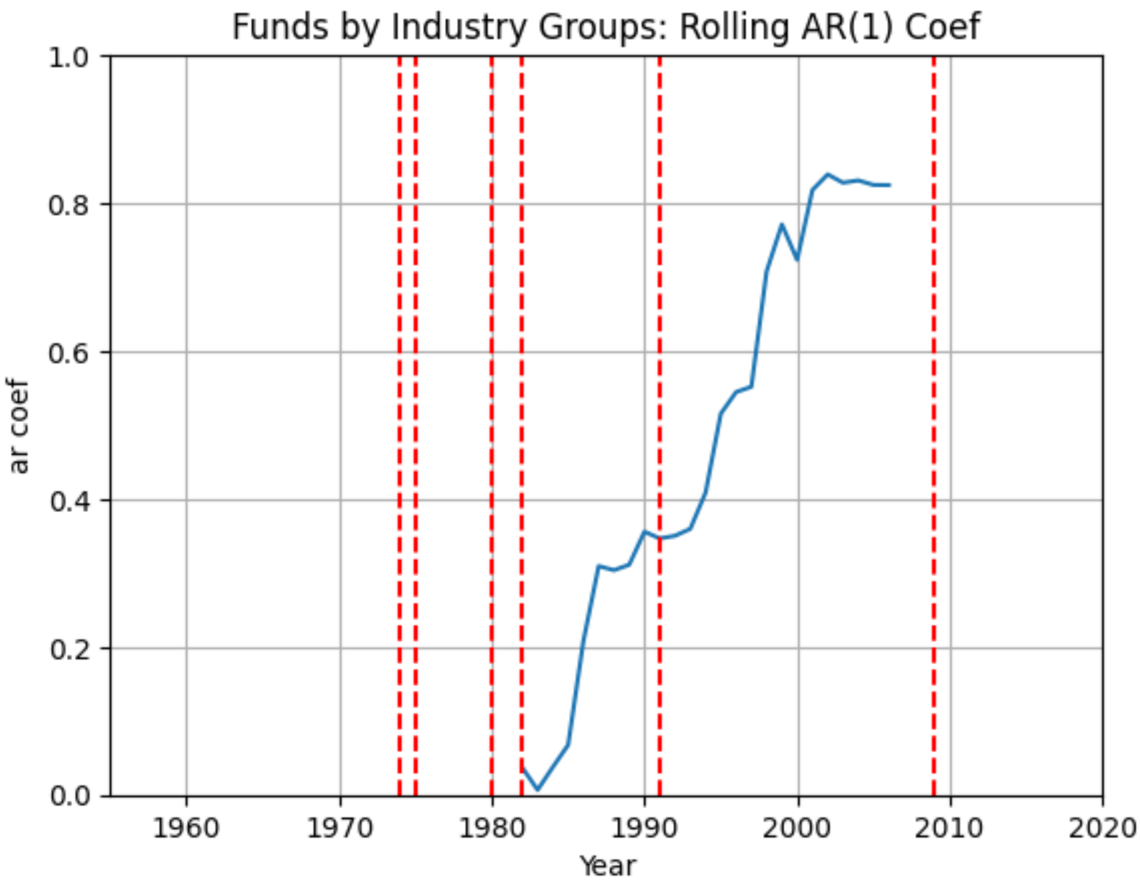
In [ ]:

```
# Uncomment reload to make changes to the code without having to restart the
# reload()

fund_bomber_i=FundBomberFactory.FundBomber(refurb=fund_bomber_i)
fund_bomber_i.identify()
fund_bomber_i.report(macro)

fund_bomber_ig=FundBomberFactory.FundBomber(refurb=fund_bomber_ig)
fund_bomber_ig.identify()
fund_bomber_ig.report(macro)
```





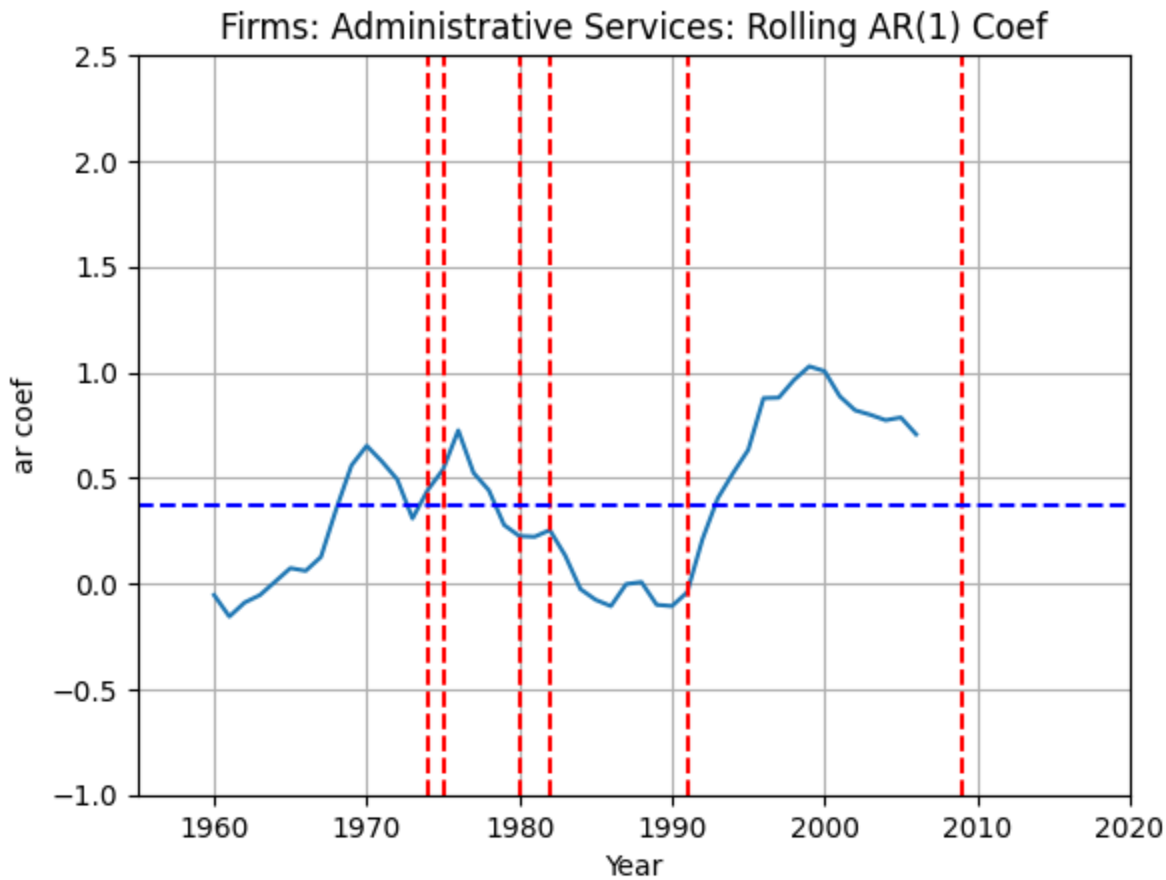
Companies: Industry Group Rolling AR(1)

Rolling AR(1) for each industry group, to check for composition effects. The graphs look slightly odd because I chose the same y-axis scale for all industry groups, resulting in less extreme industry groups being "squished". Some industry groups have high coefficients on average (~1) without necessarily looking like it.

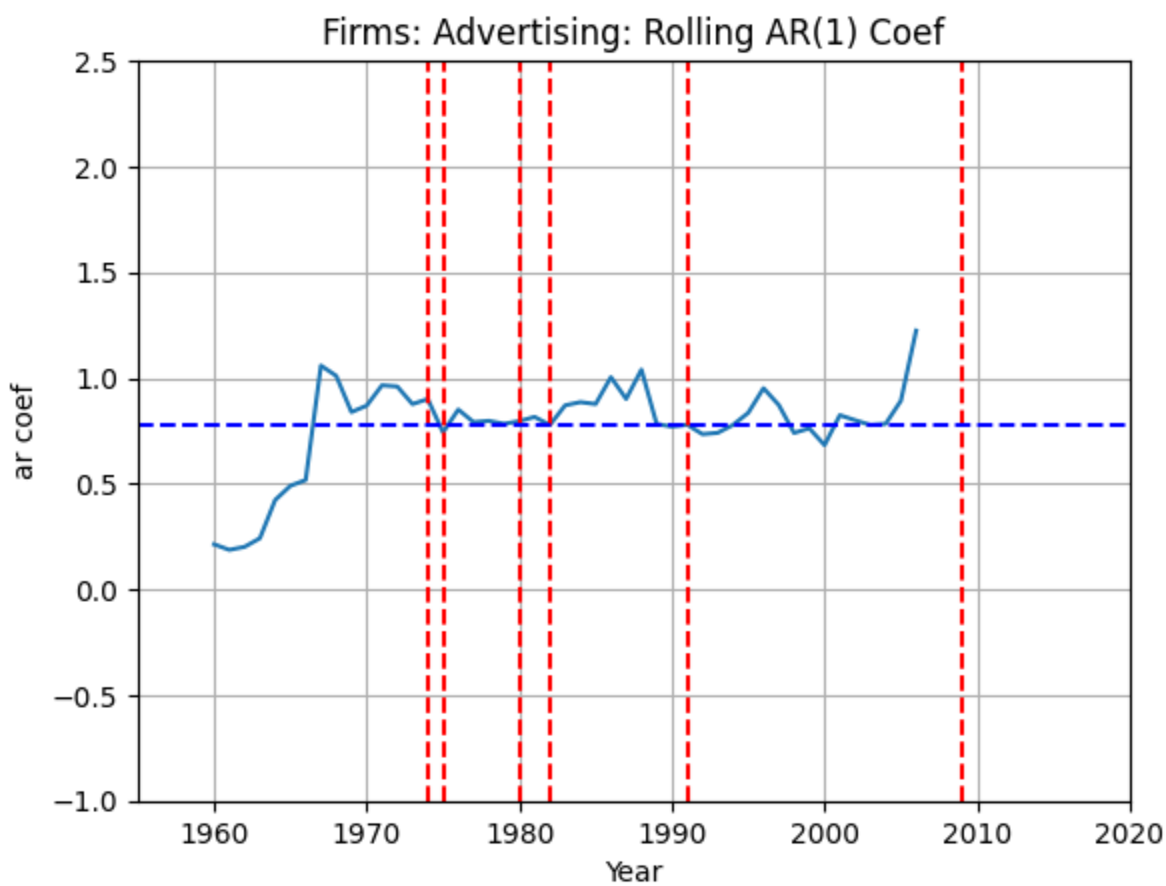
Overall, the graphs seem to indicate that the analysis above is legitimate. "Slow" industries like "Administrative Services" and "Education" have low coefficients on average, while "Fast" industries like "Blockchain and Cryptocurrency" and "Information Technology" have high coefficients on average. There are highs in the graph of "Real Estate" corresponding to known bubbles in the 1980s and 2000s, as well as "Financial Services" in the 2000s. While the validity of this method is not fully established (there are also some oddities in the graphs like the lows in the graph of "Financial Services" in the 1980s and the consistently high coefficient of "Manufacturing") it does at least seem to capture something real, and almost all groups exhibit increasing "trend-following" since data is available for them.

Blue lines are the average coefficient of a category.

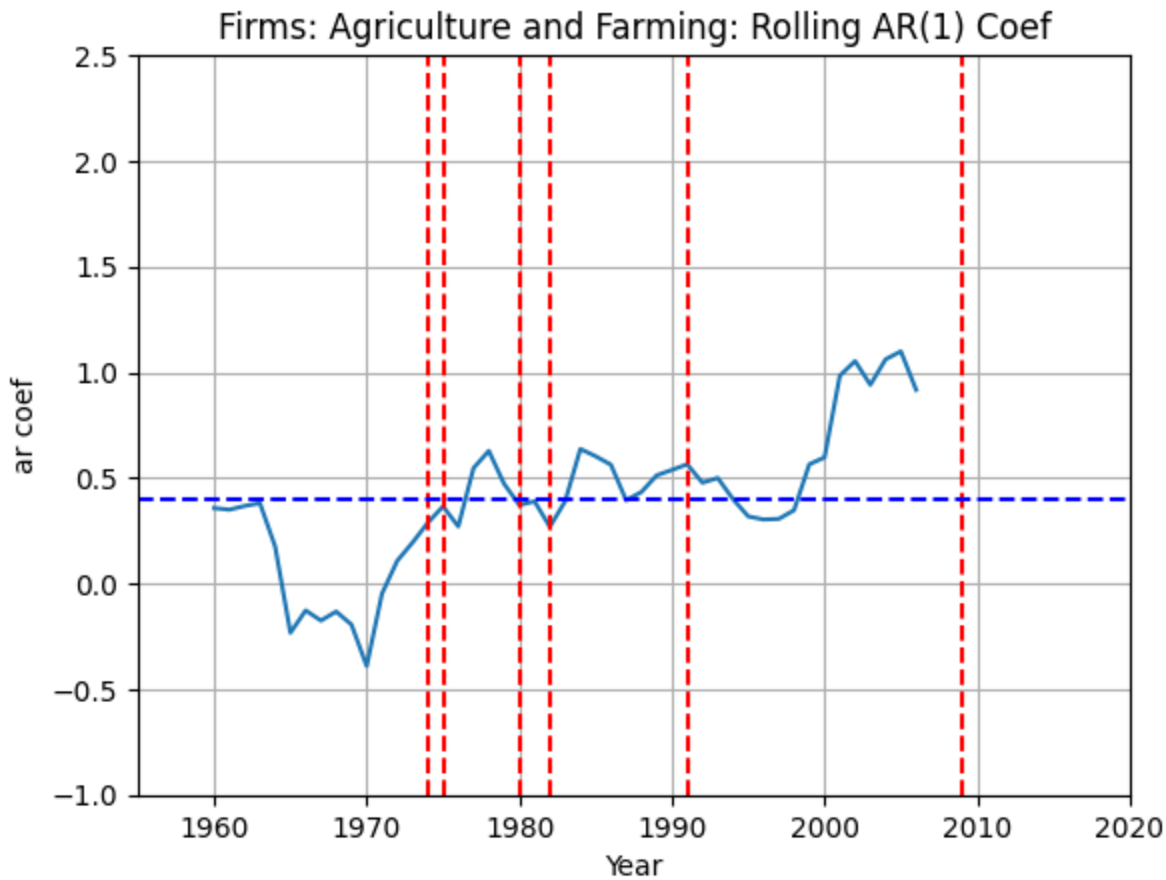
In [ ]:
```
firm_bomber_ig.detailed_report(macro)
```

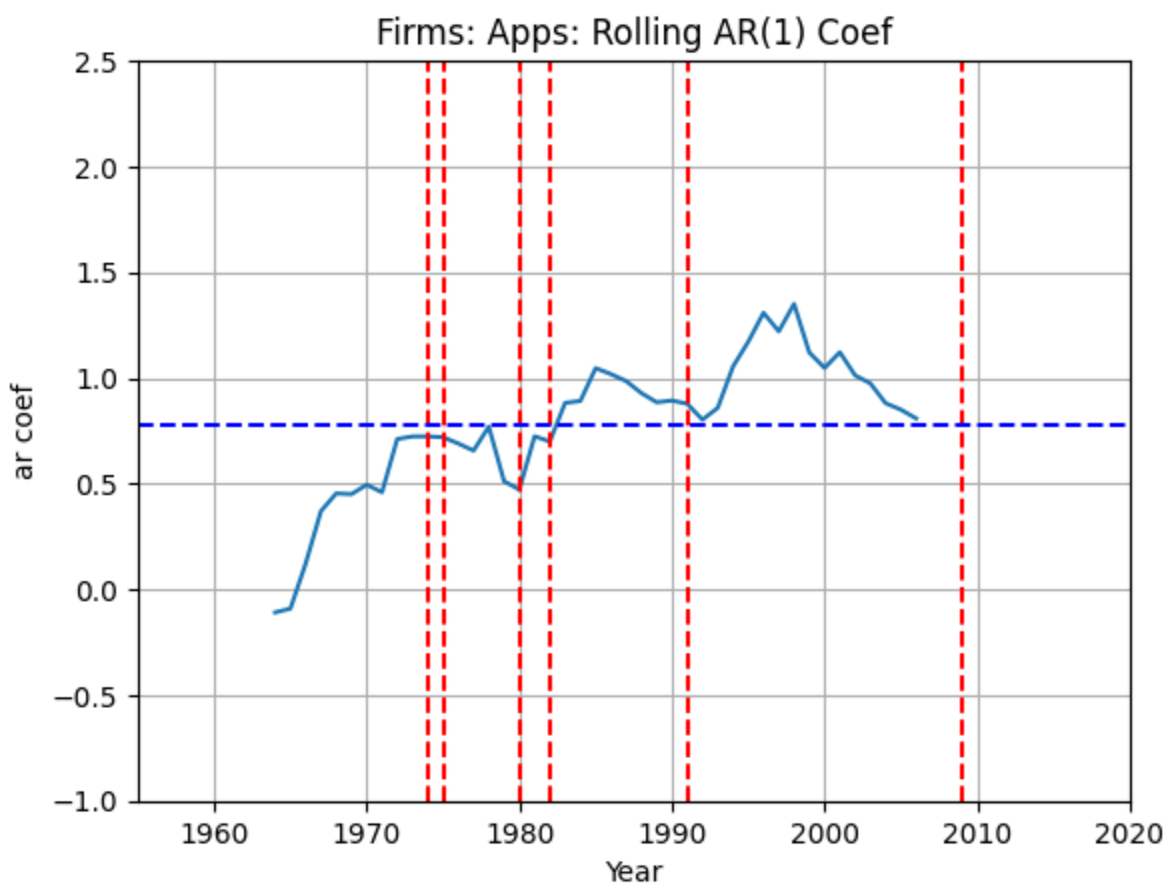


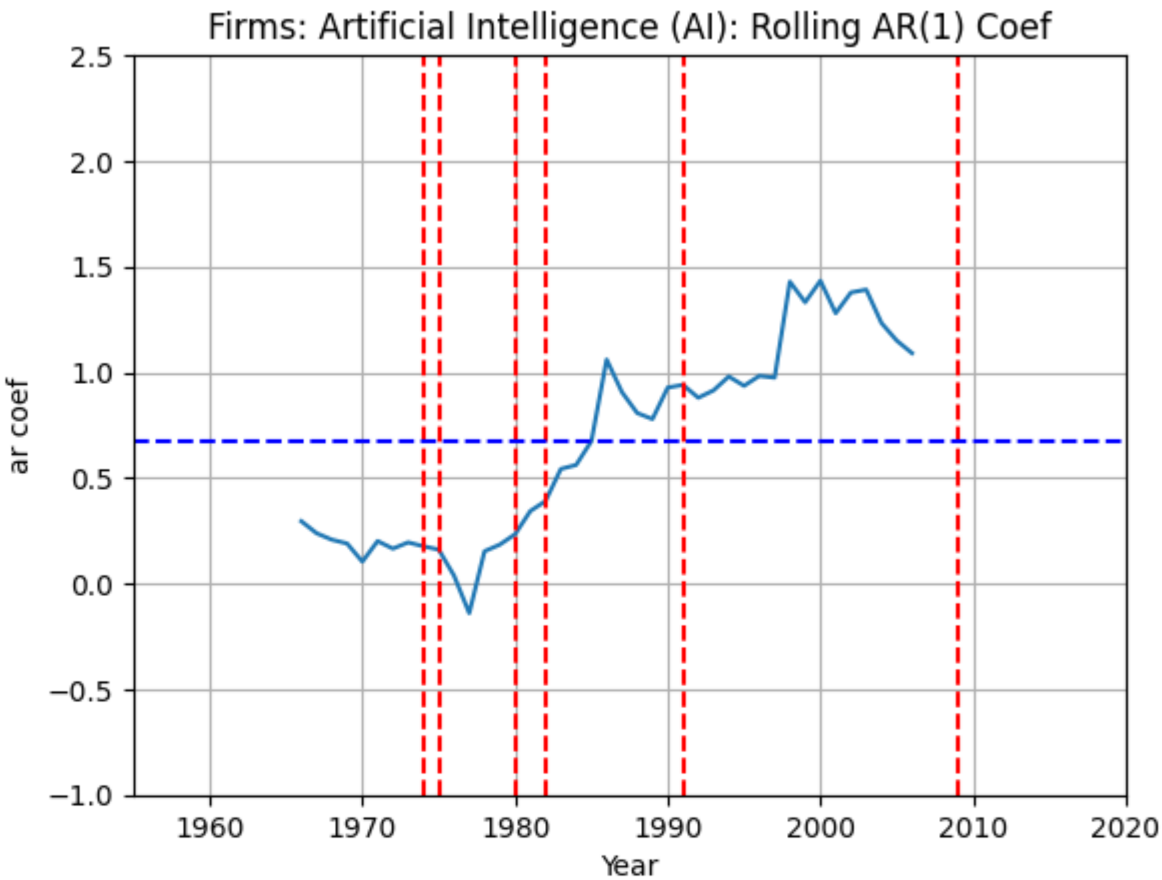
Firms: Administrative Services: Rolling AR(1) Coef
ar coef
Year

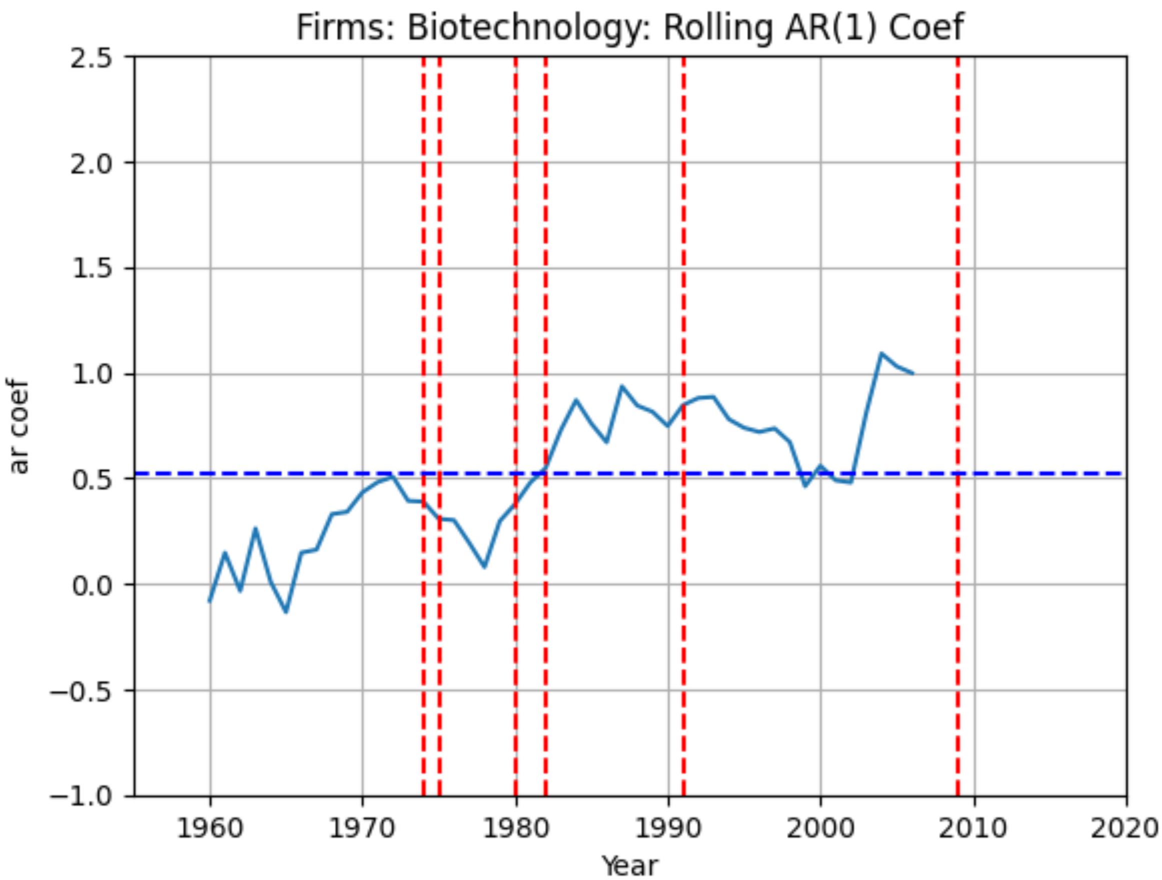


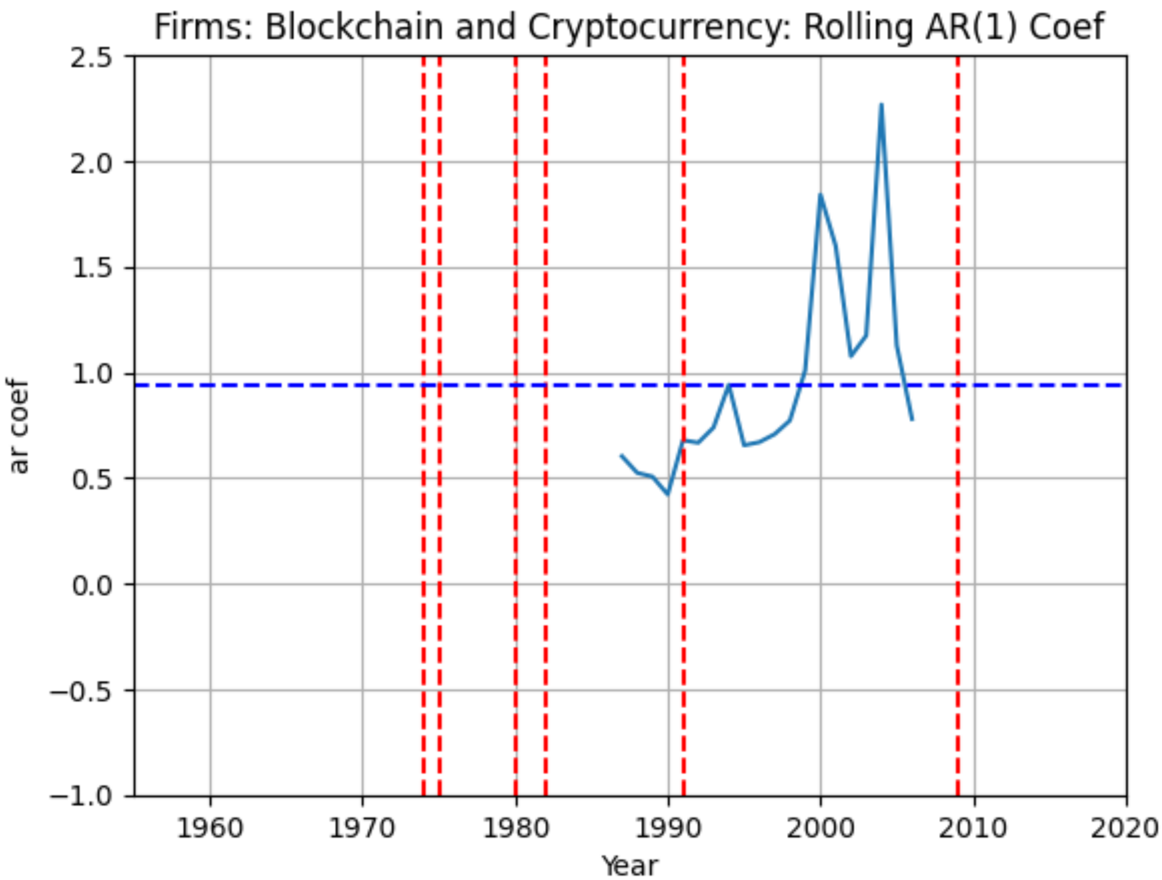
Firms: Advertising: Rolling AR(1) Coef
ar coef
Year

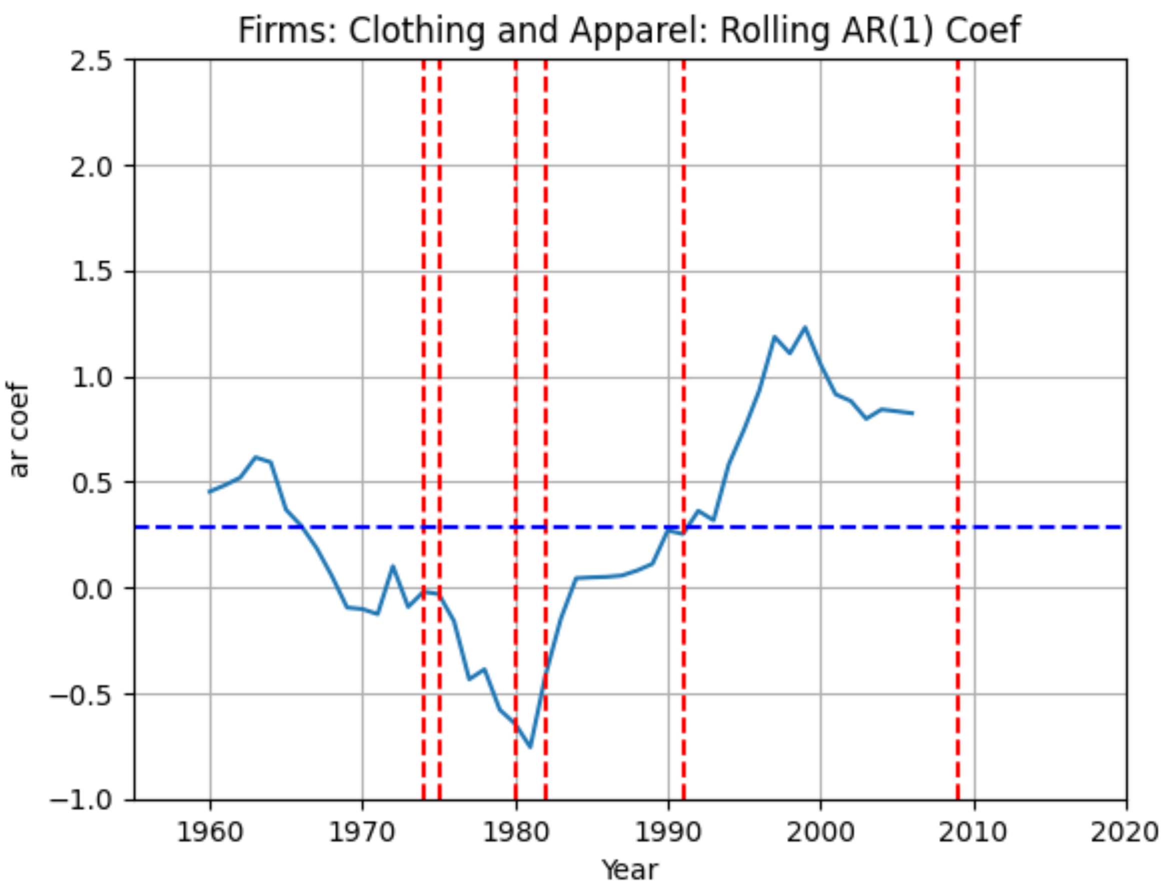

Firms: Agriculture and Farming: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Apps: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

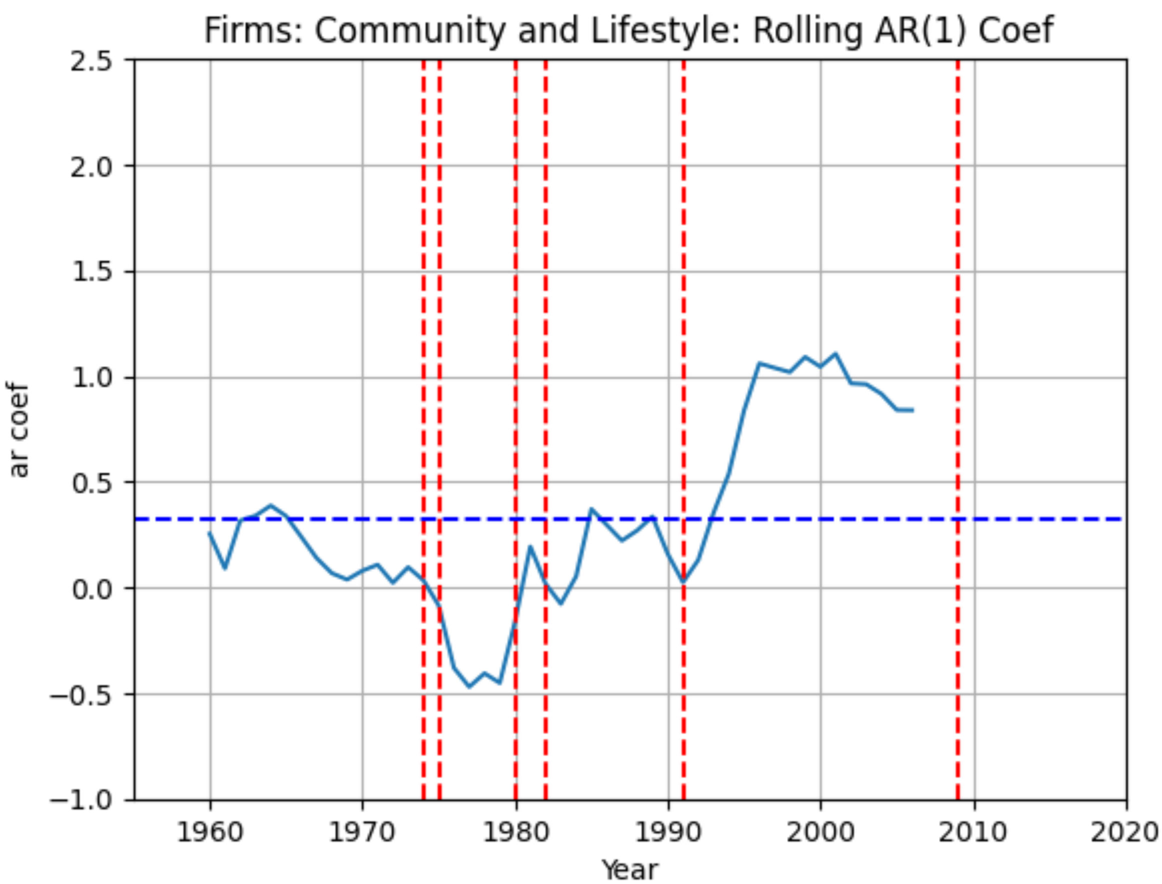

Firms: Artificial Intelligence (AI): Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

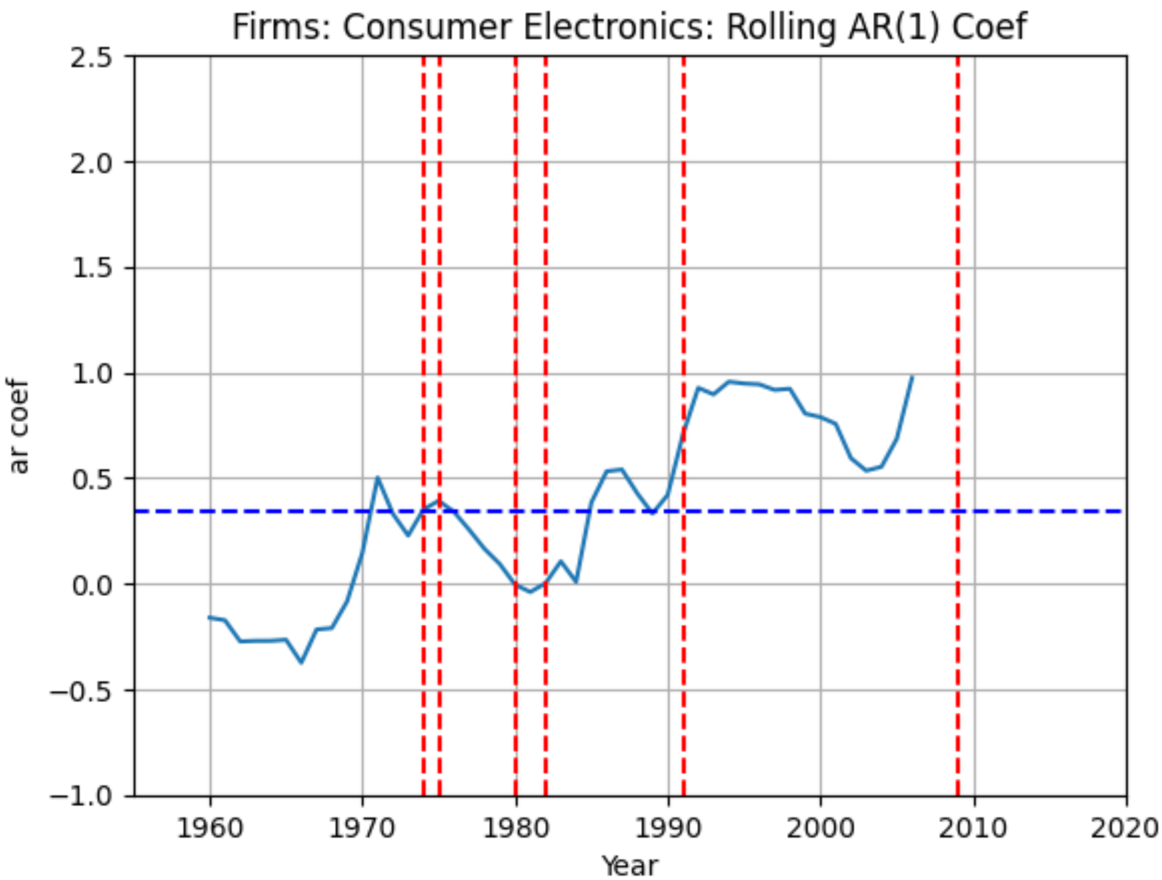

Firms: Biotechnology: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

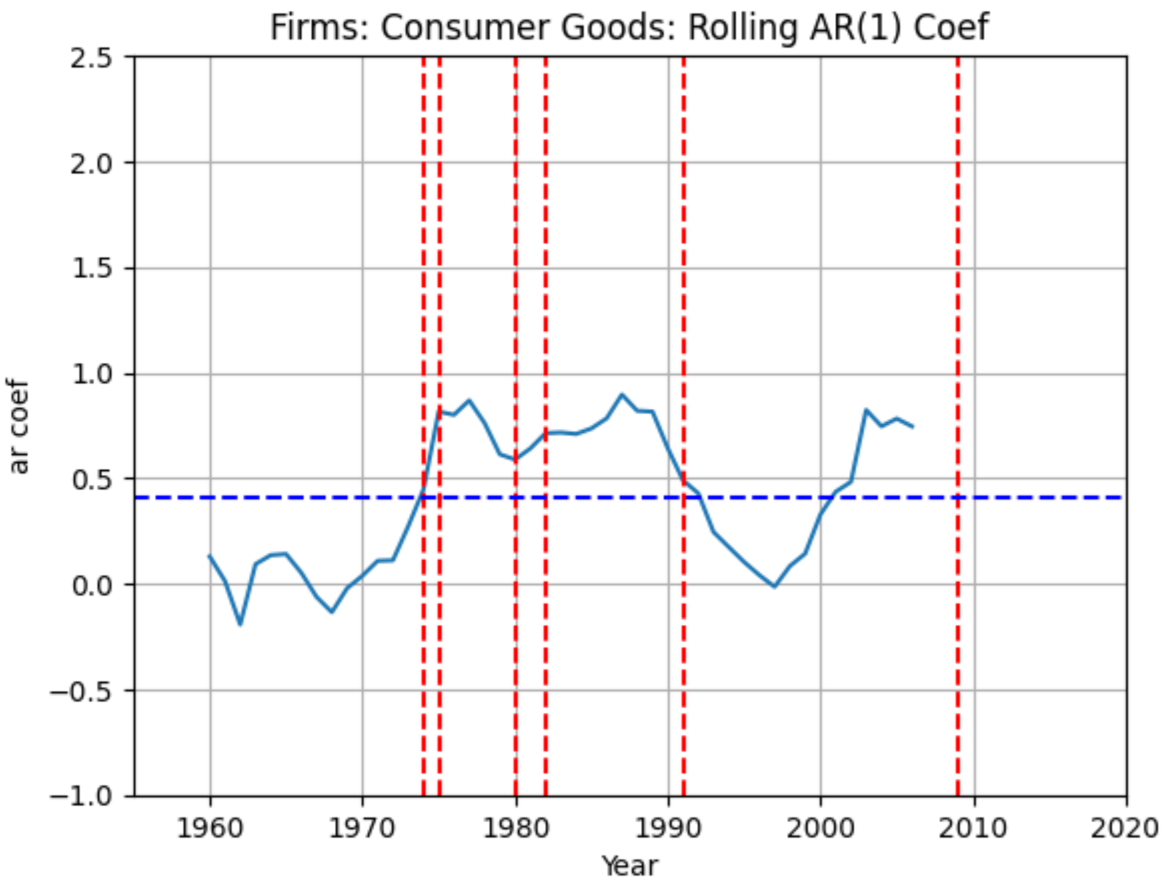

Firms: Blockchain and Cryptocurrency: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

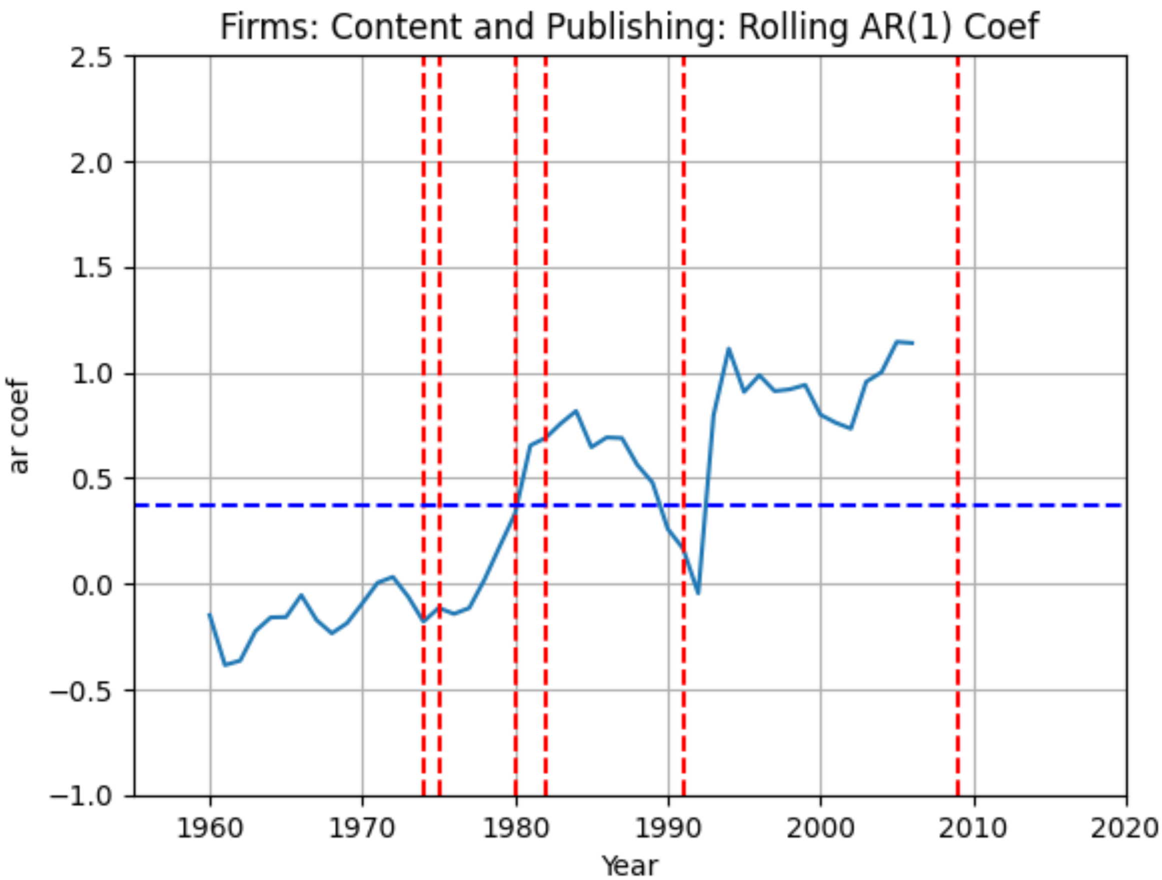

Firms: Clothing and Apparel: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

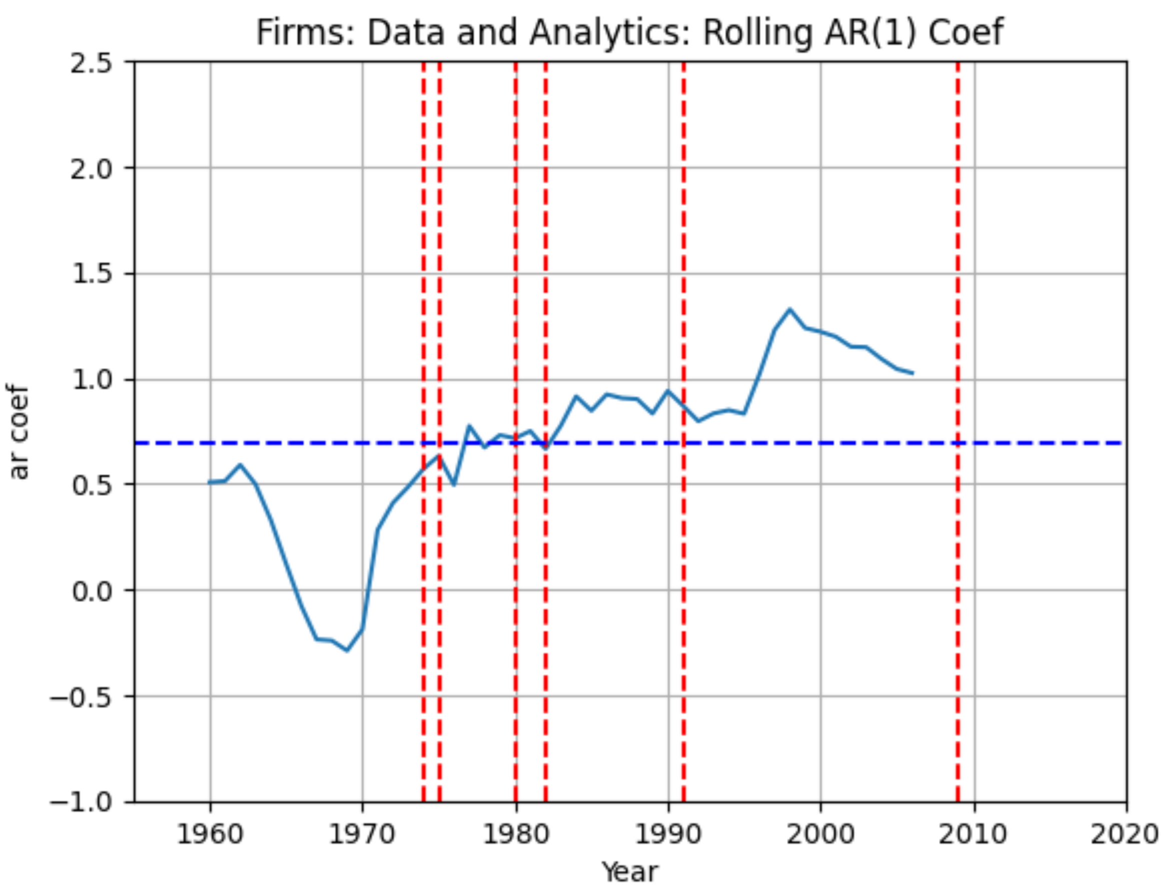

Firms: Commerce and Shopping: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

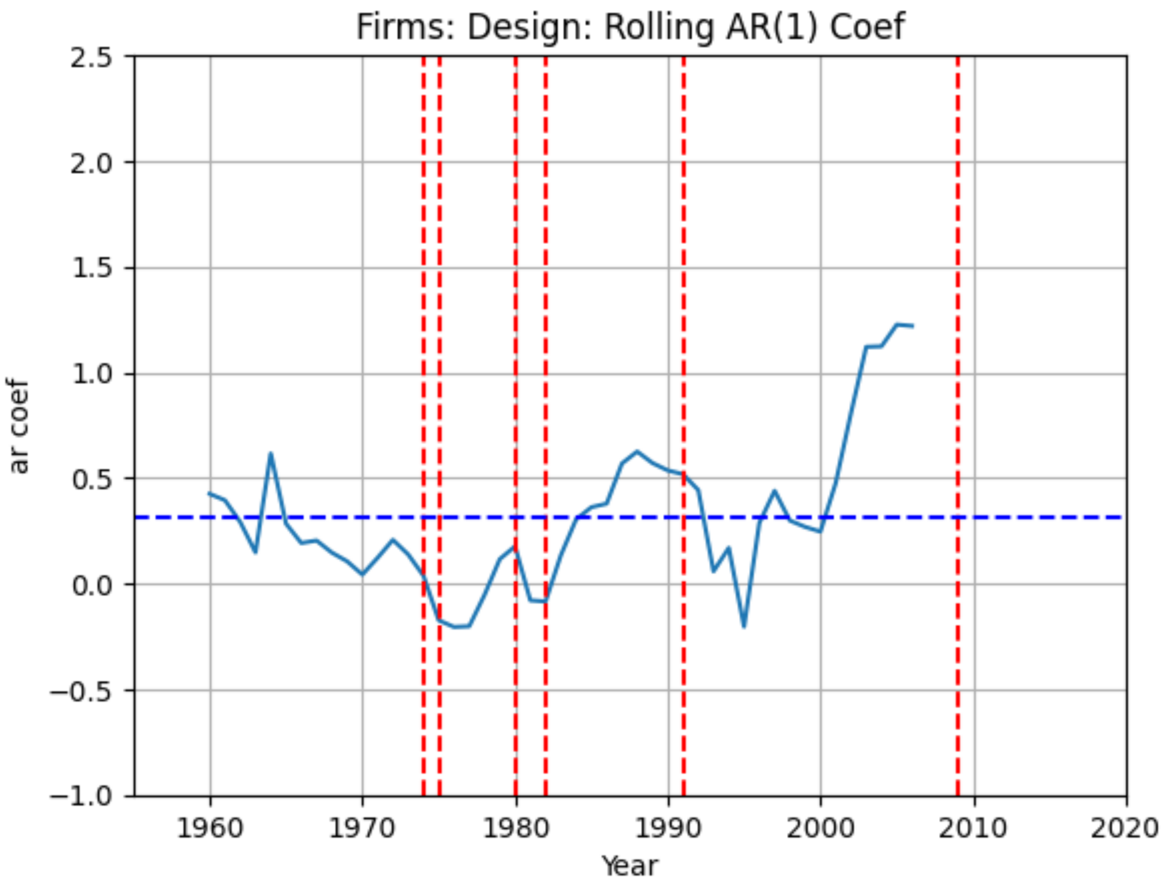

Firms: Community and Lifestyle: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year


Firms: Consumer Electronics: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

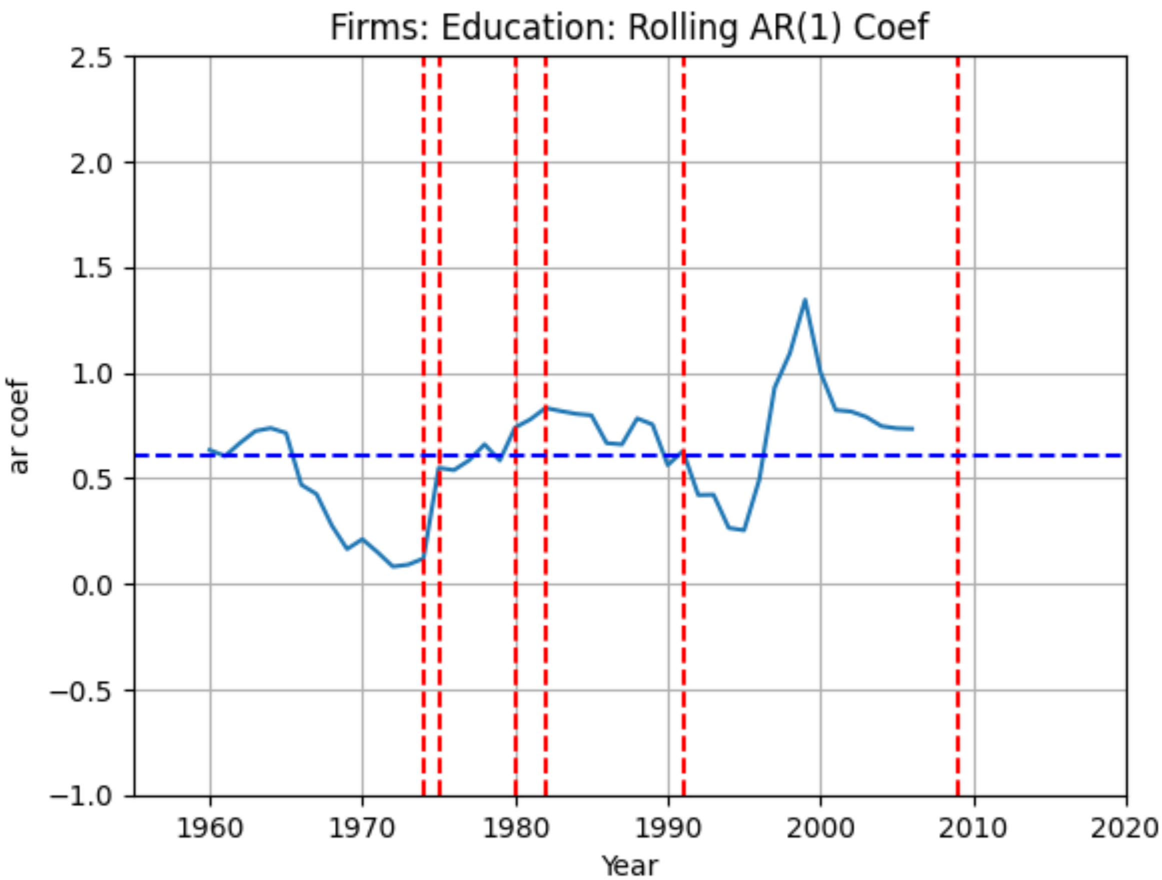

Firms: Consumer Goods: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year


Firms: Content and Publishing: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

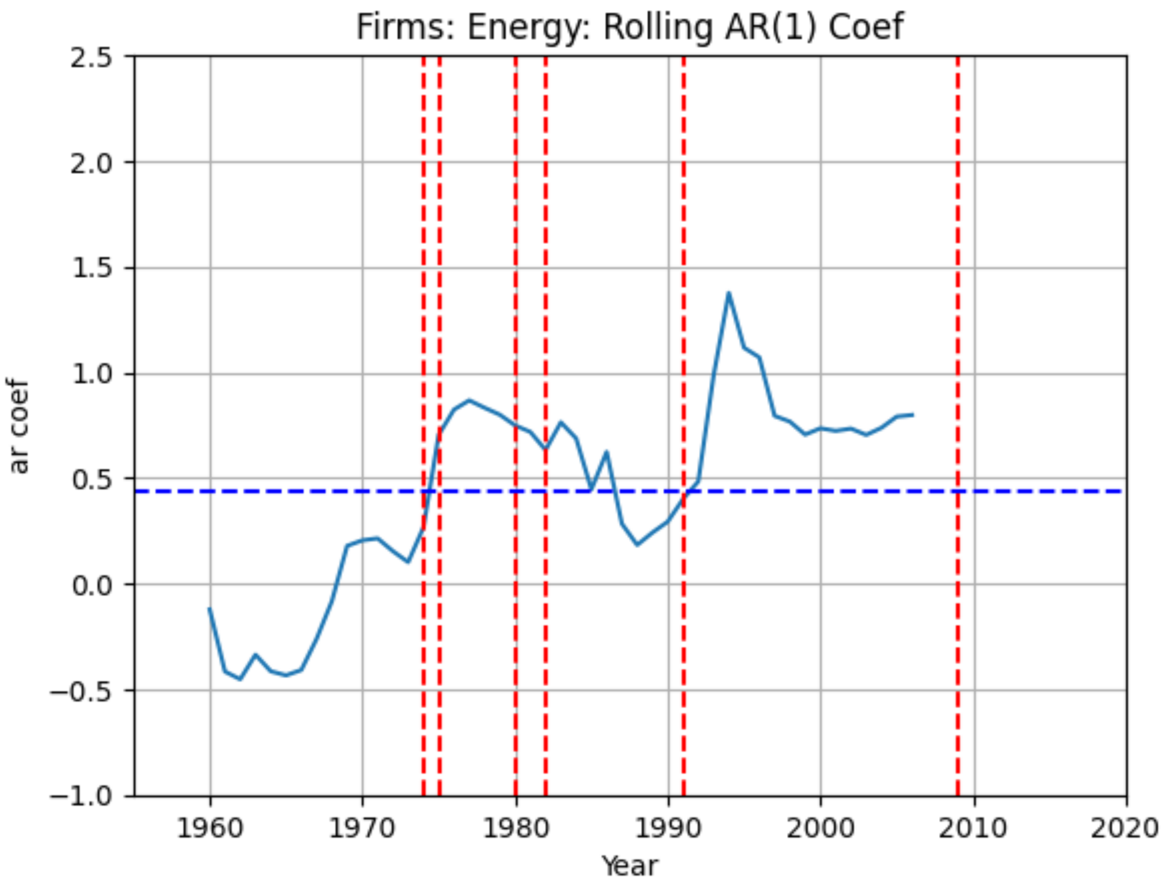

Firms: Data and Analytics: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

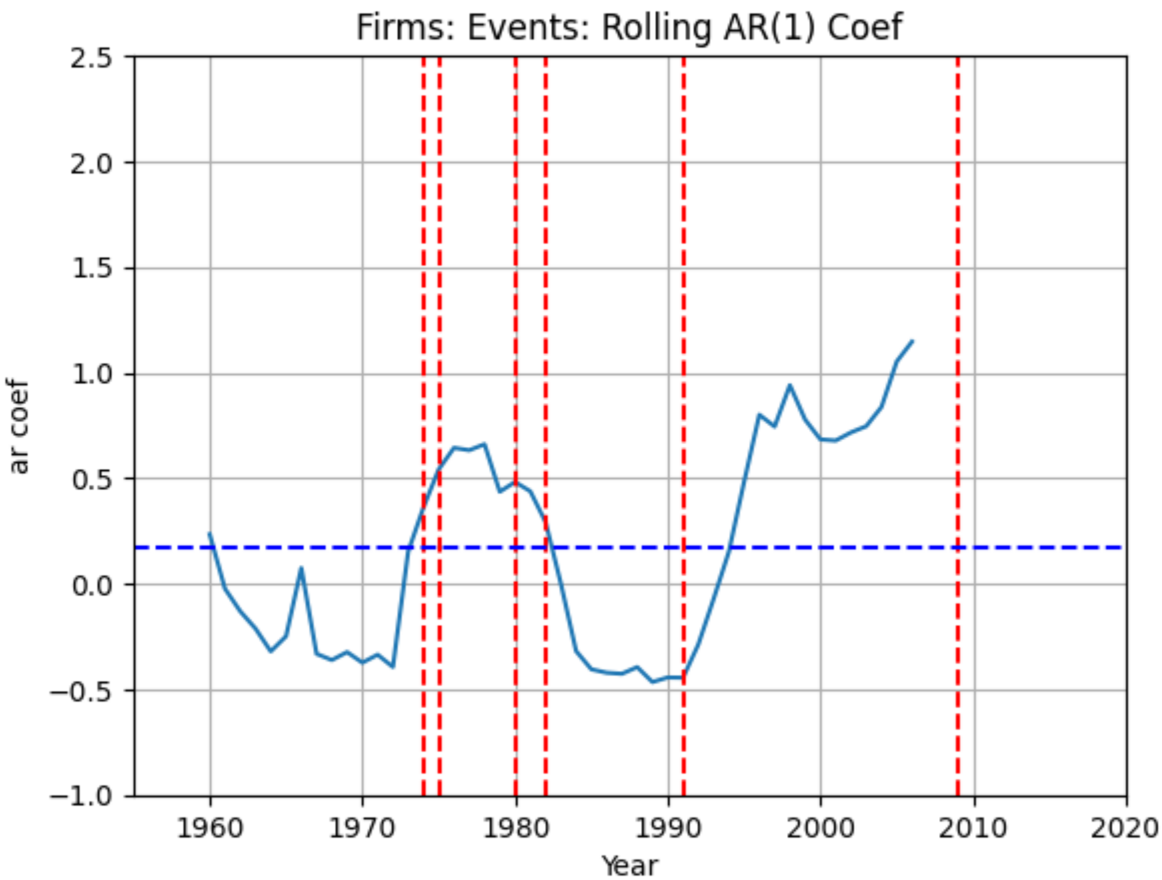


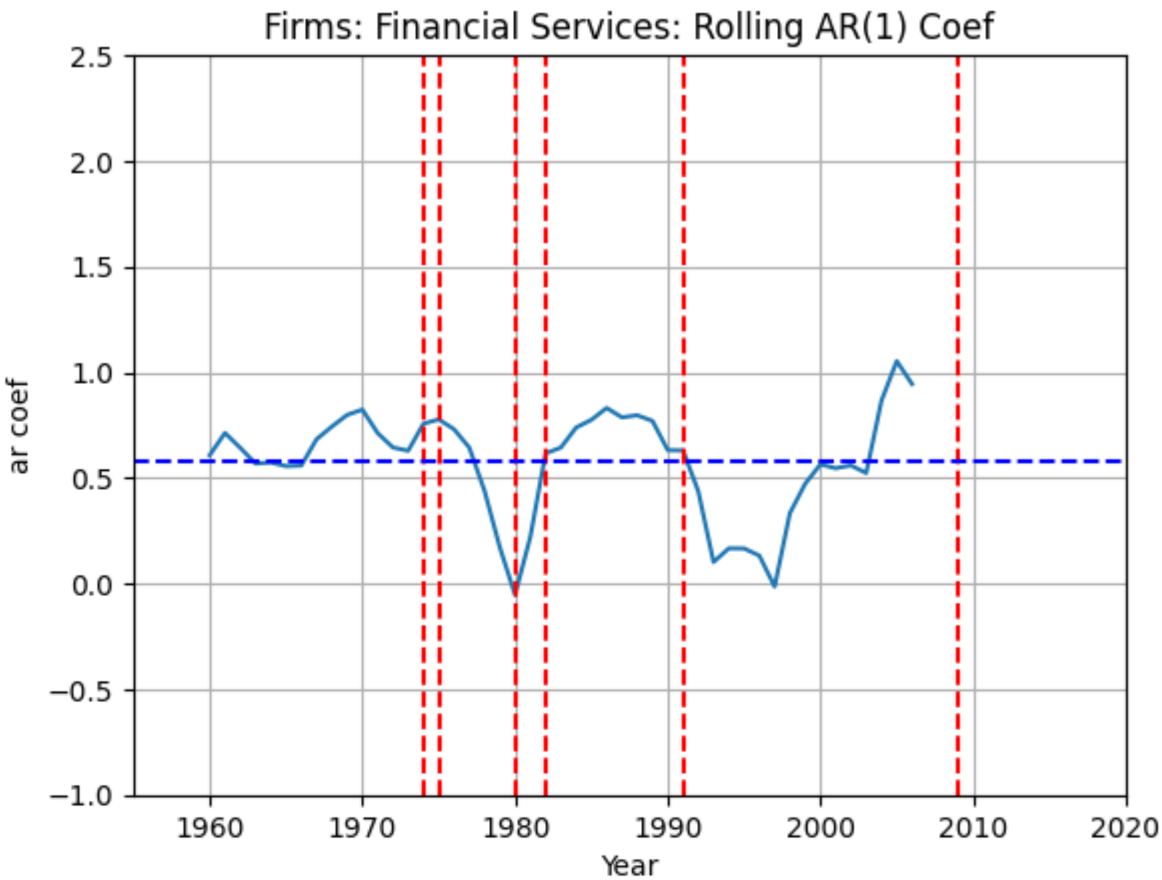
Firms: Design: Rolling AR(1) Coef
ar coef
Year



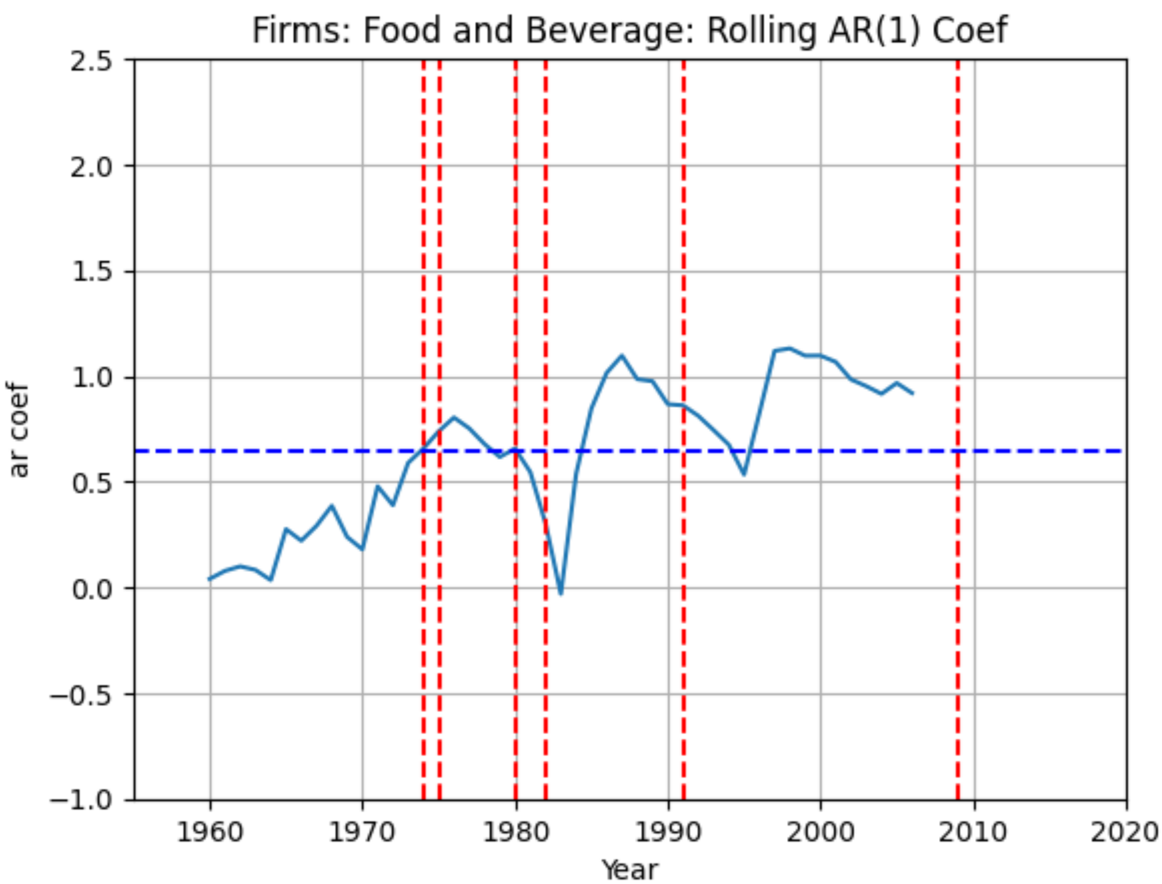
Firms: Education: Rolling AR(1) Coef
ar coef
Year


Firms: Energy: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

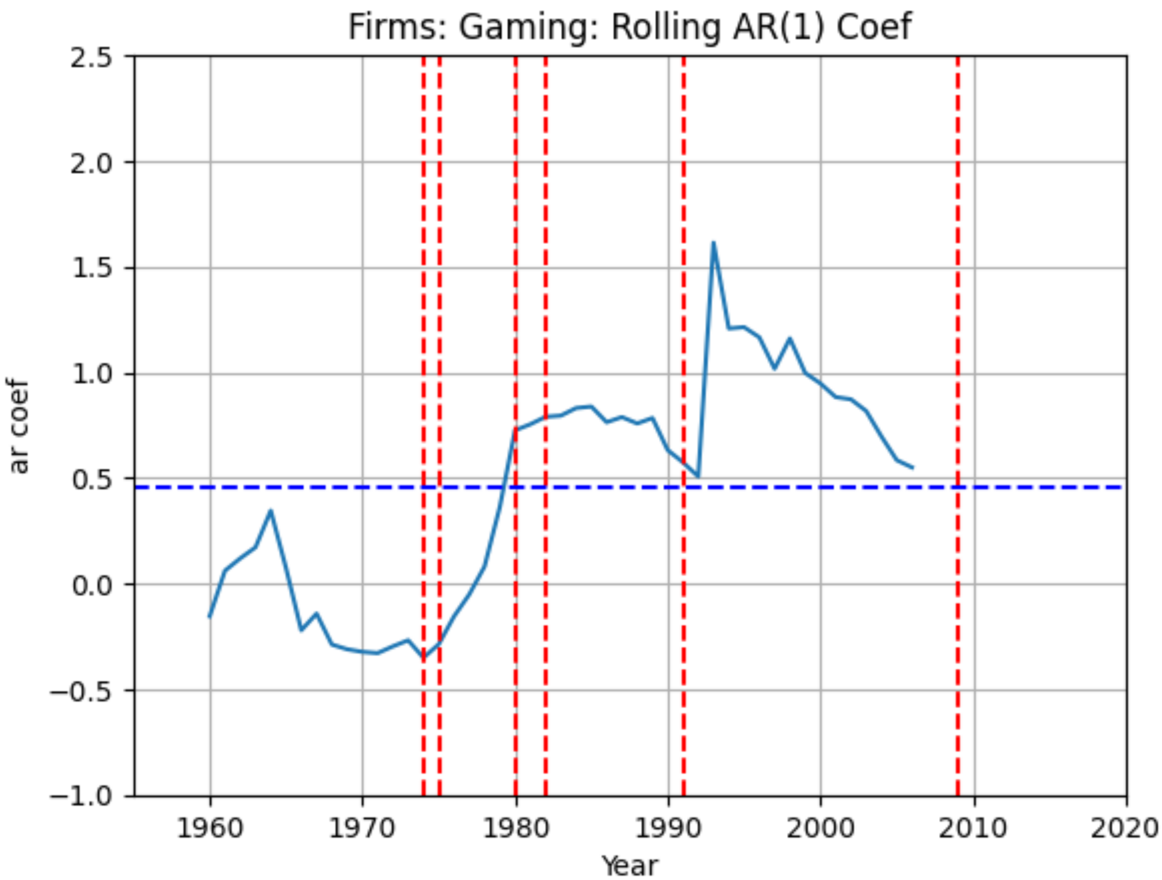

Firms: Events: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

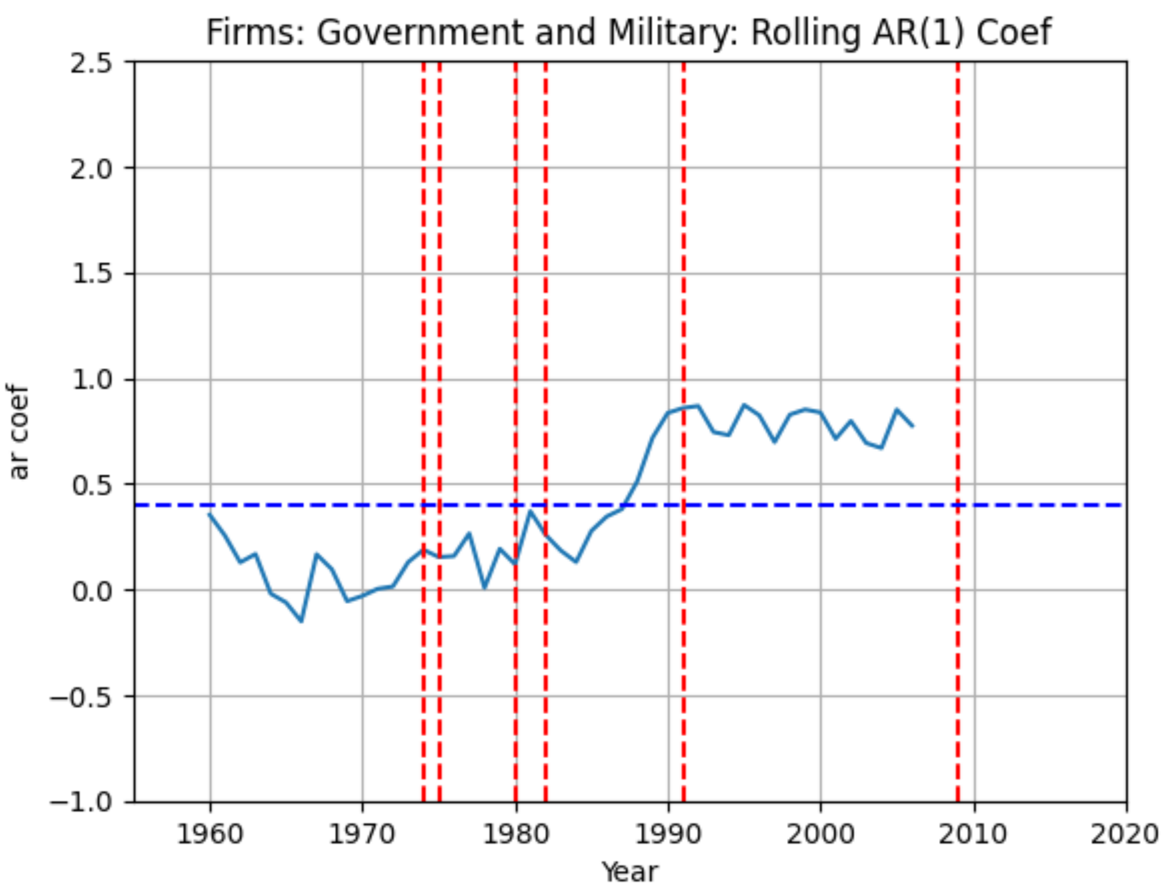


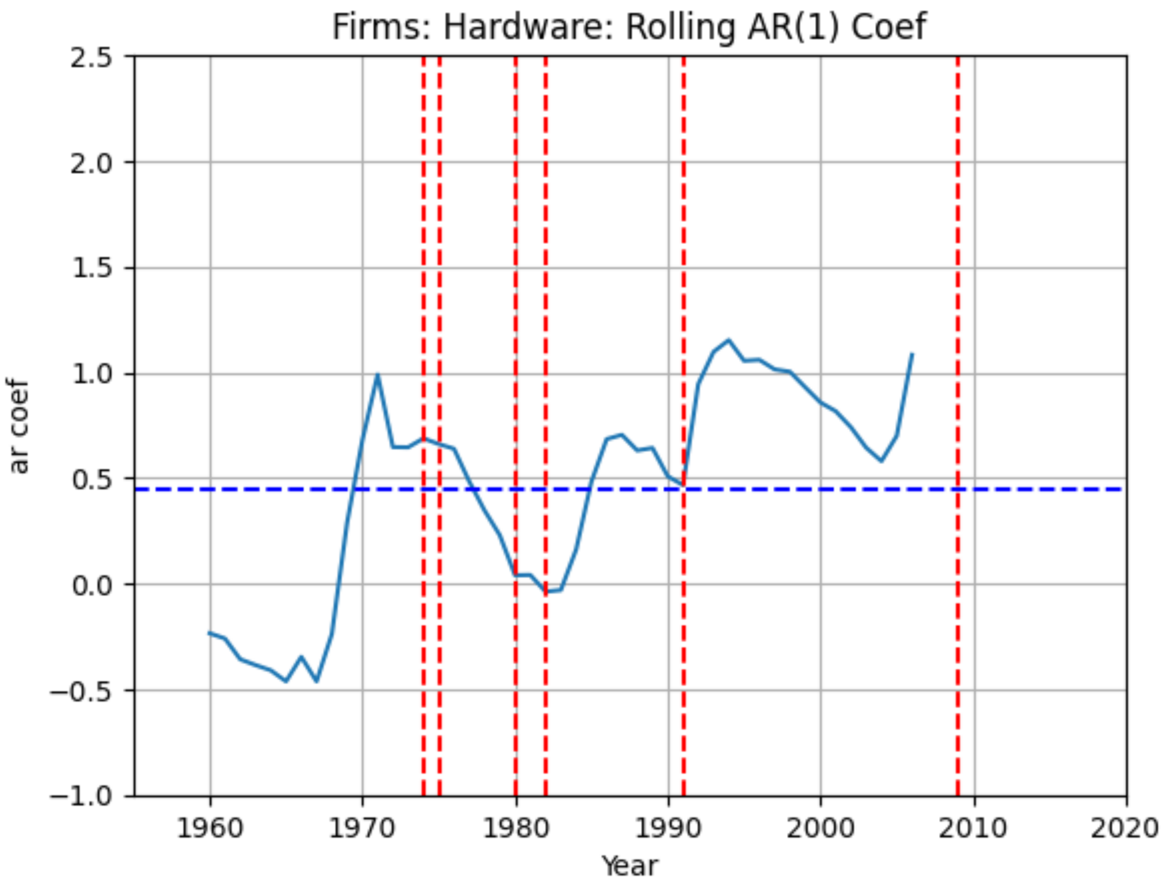
Firms: Financial Services: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020



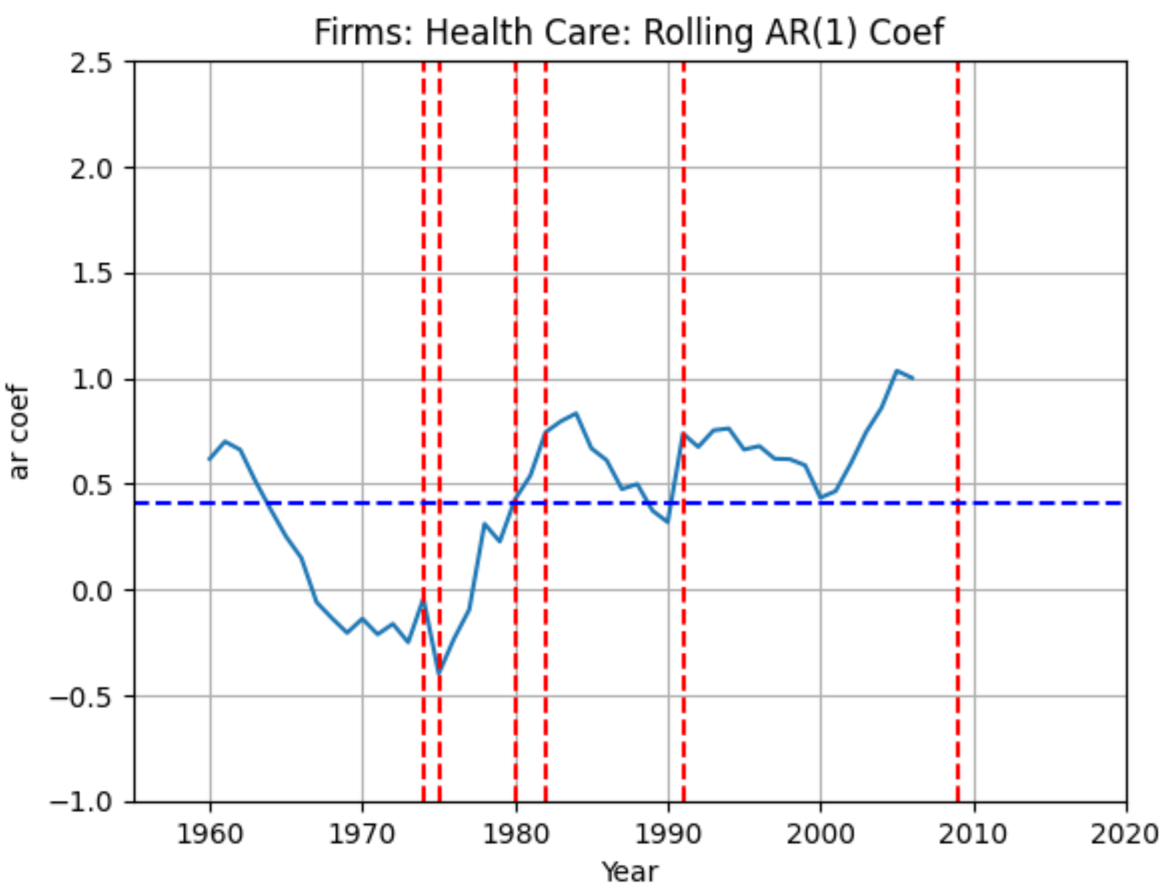
Firms: Food and Beverage: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Gaming: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

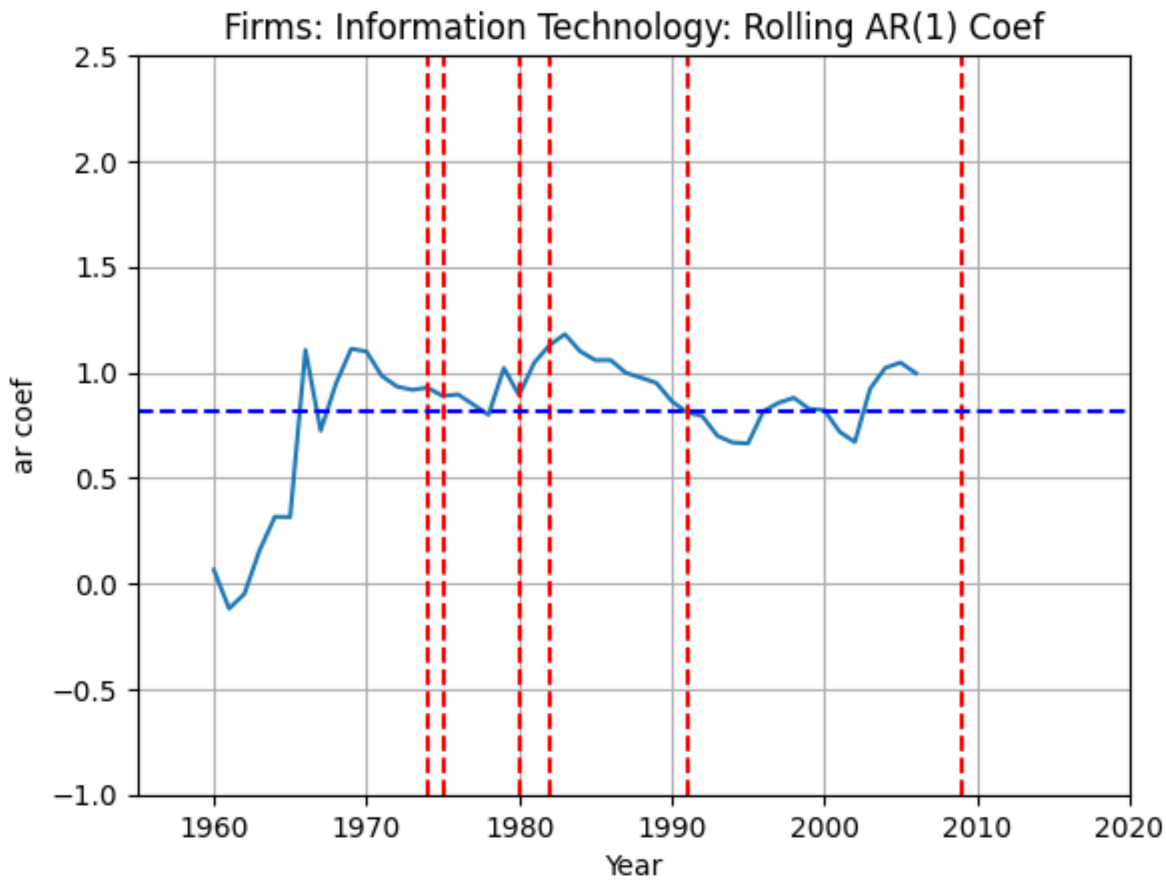

Firms: Government and Military: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

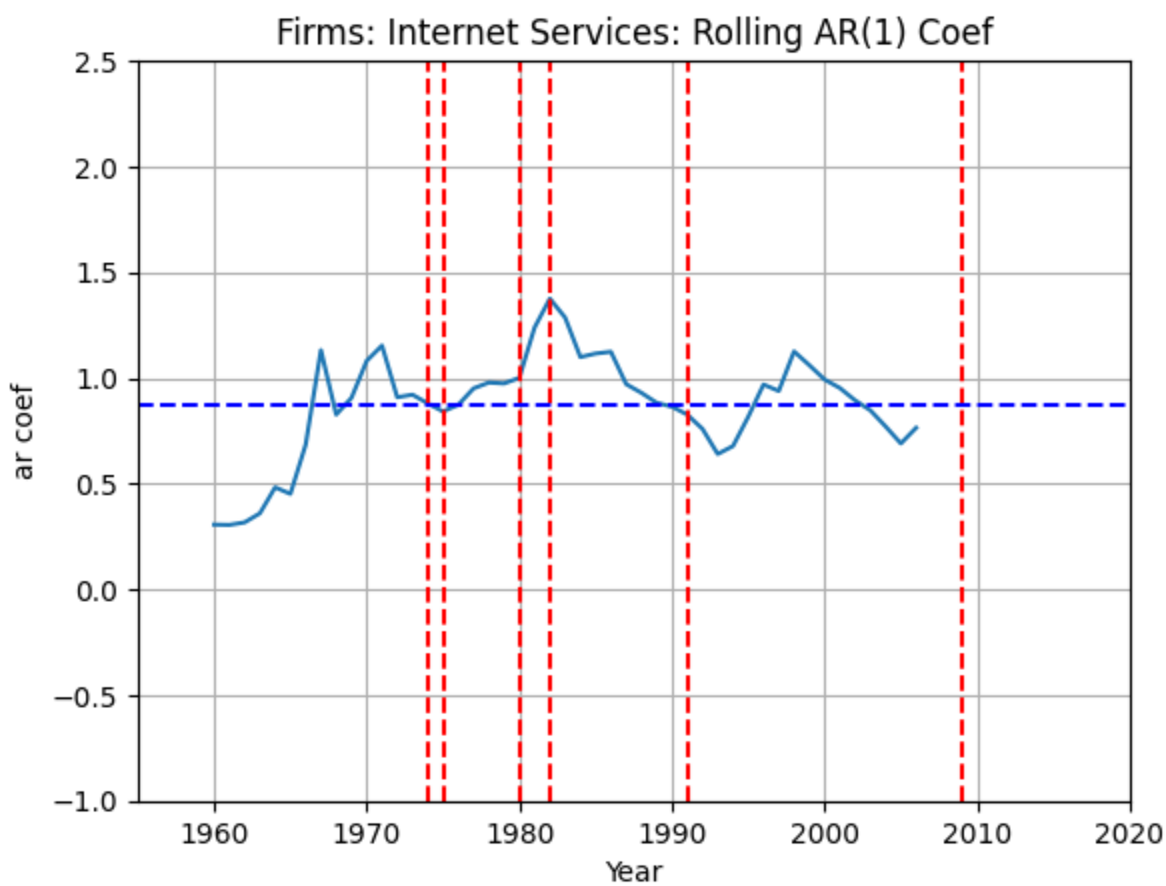


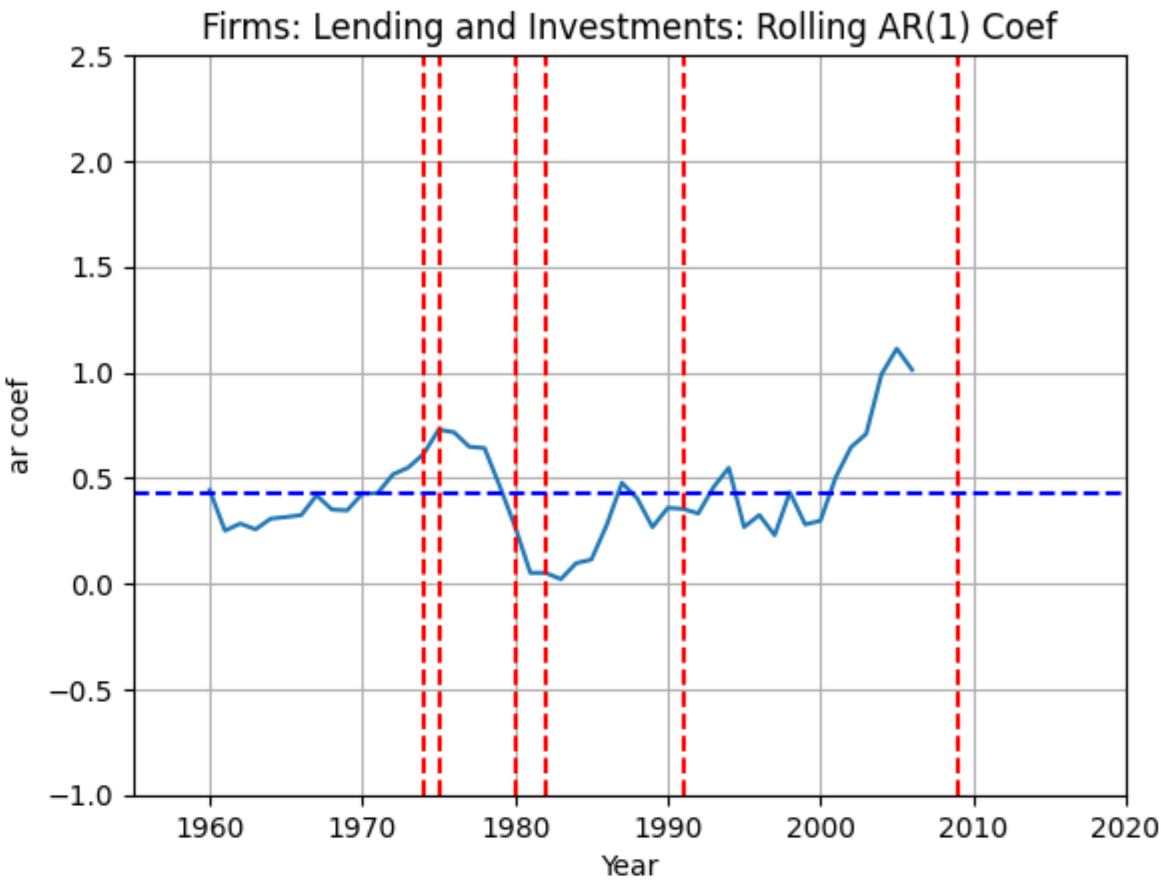
Firms: Hardware: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

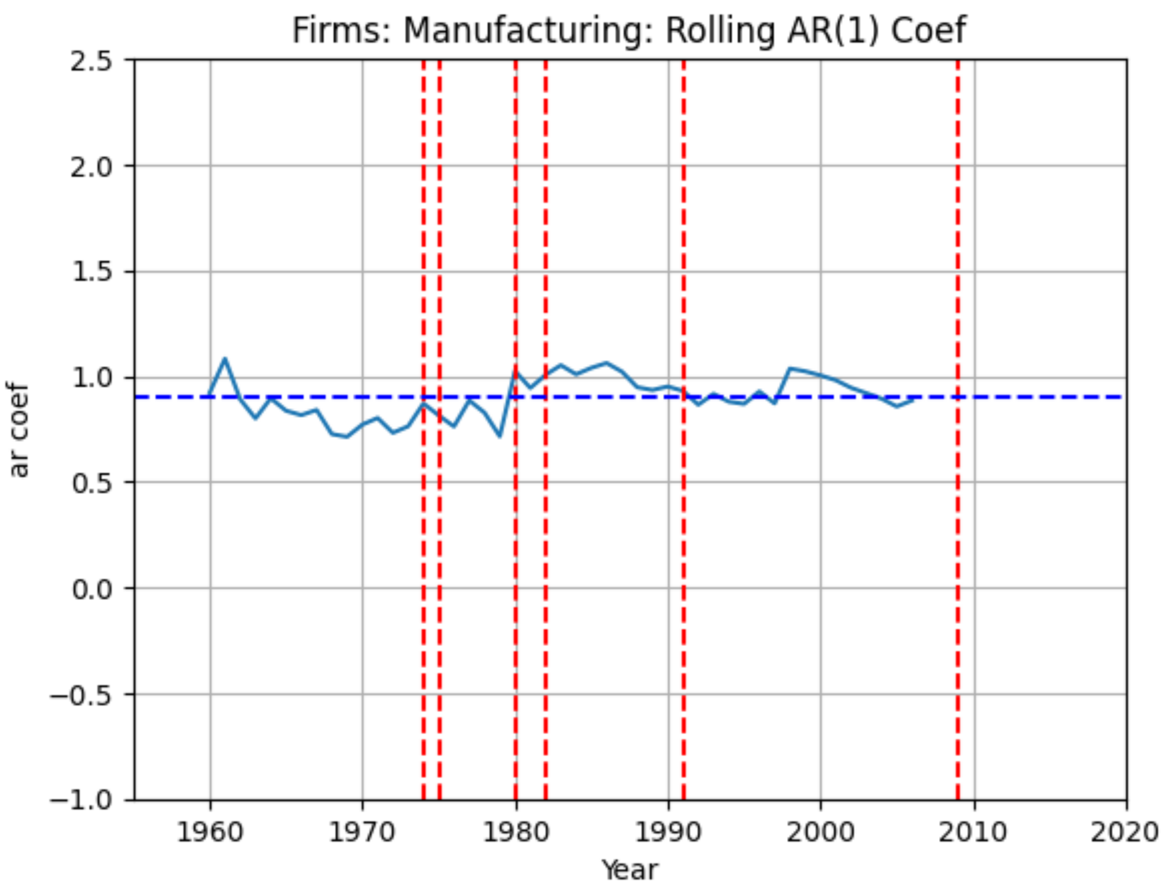


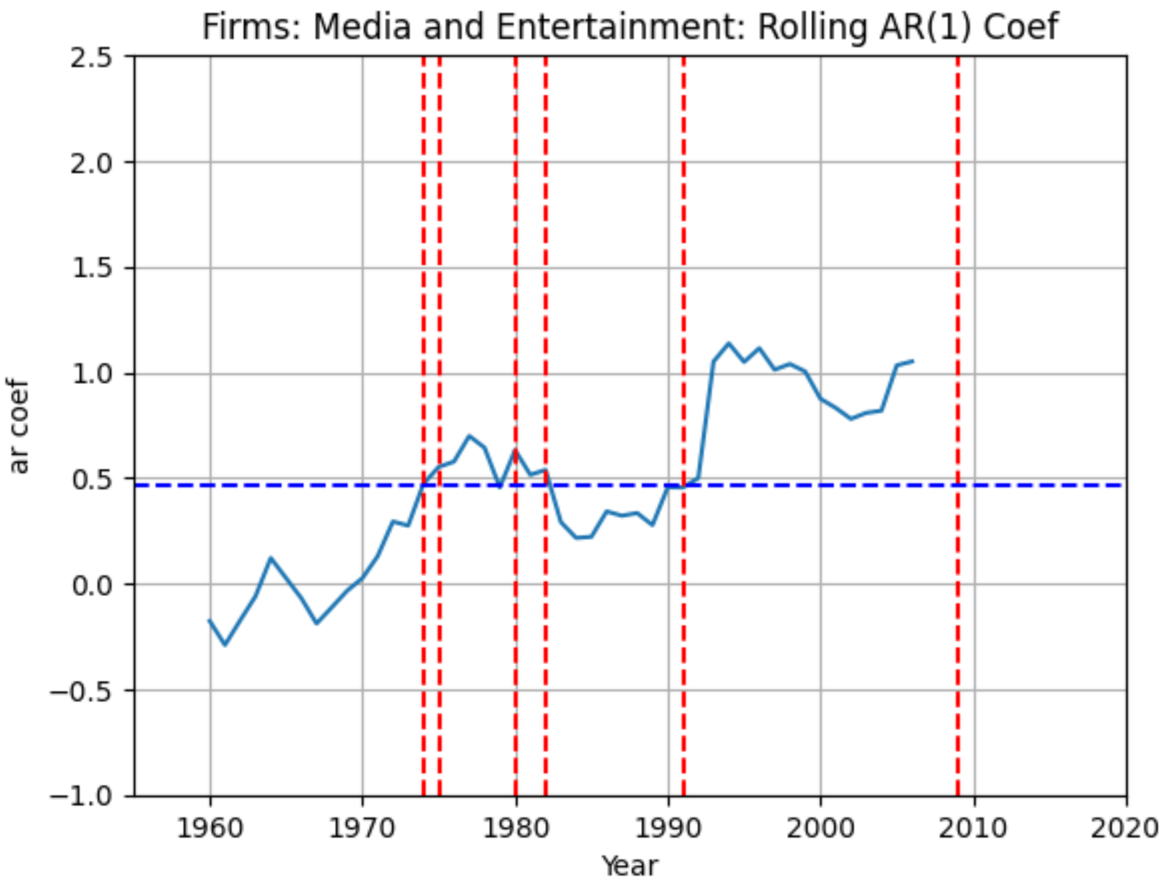
Firms: Health Care: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

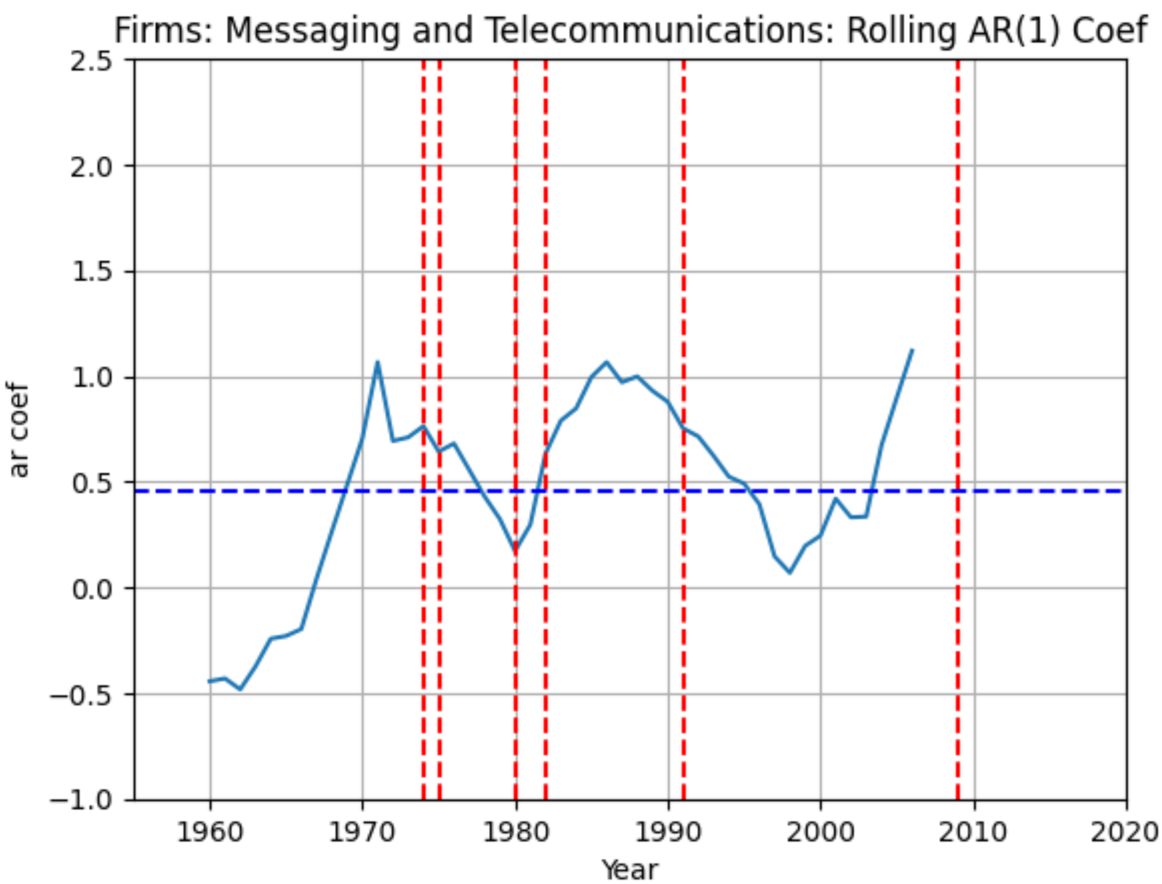

Firms: Information Technology: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

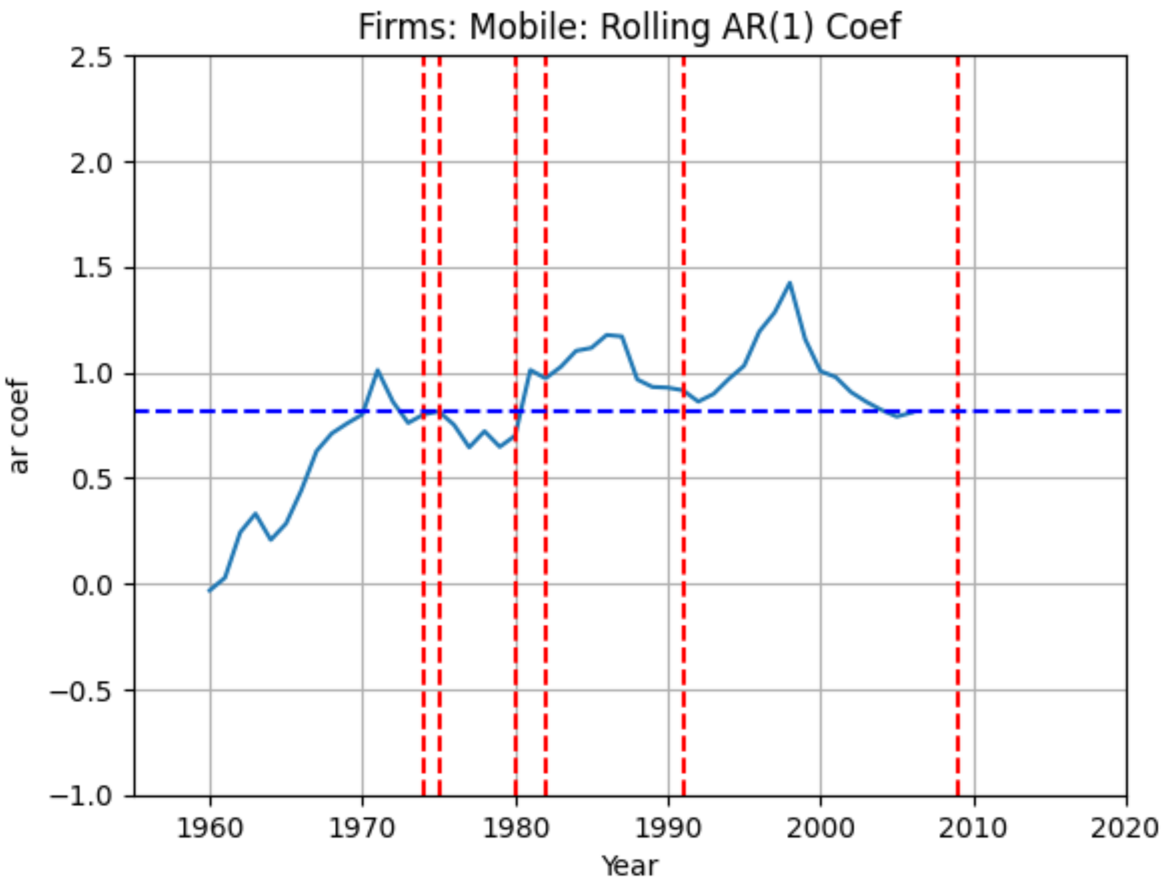

Firms: Internet Services: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

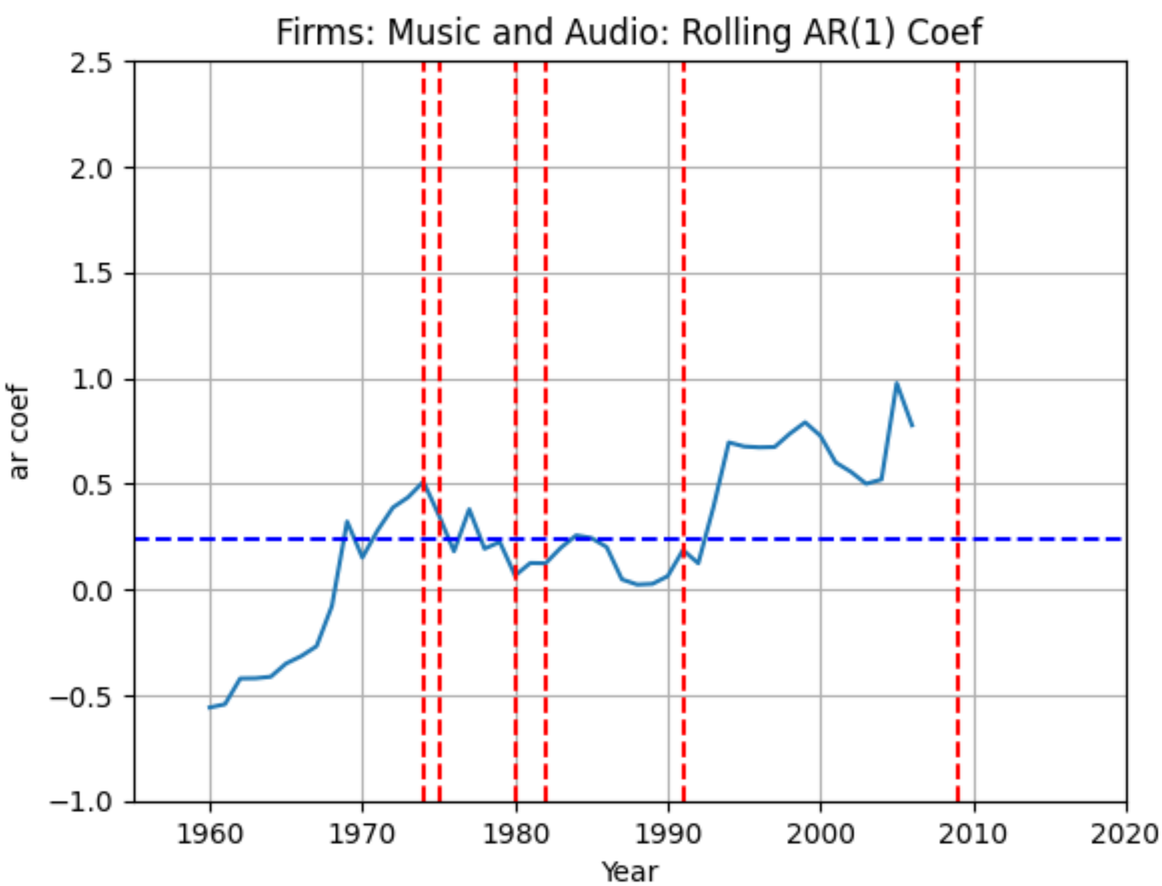

Firms: Lending and Investments: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

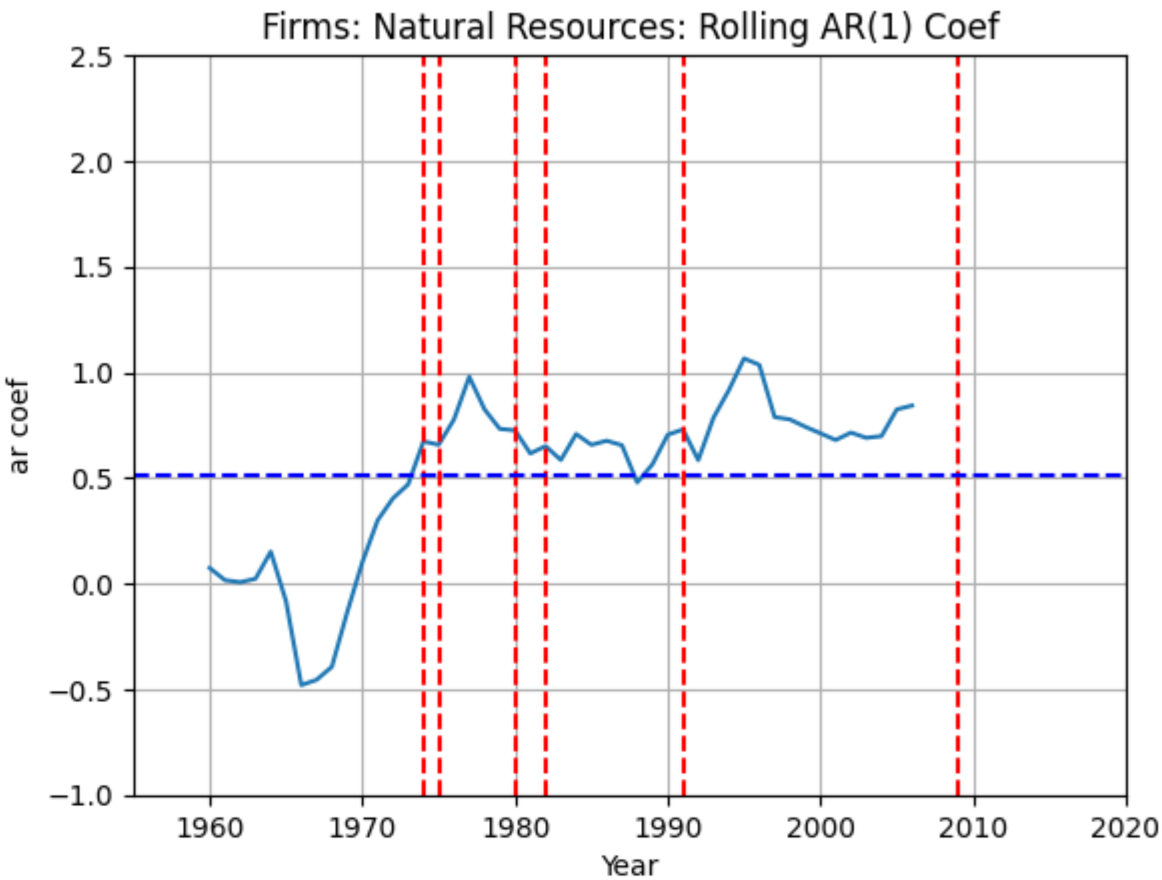

Firms: Manufacturing: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

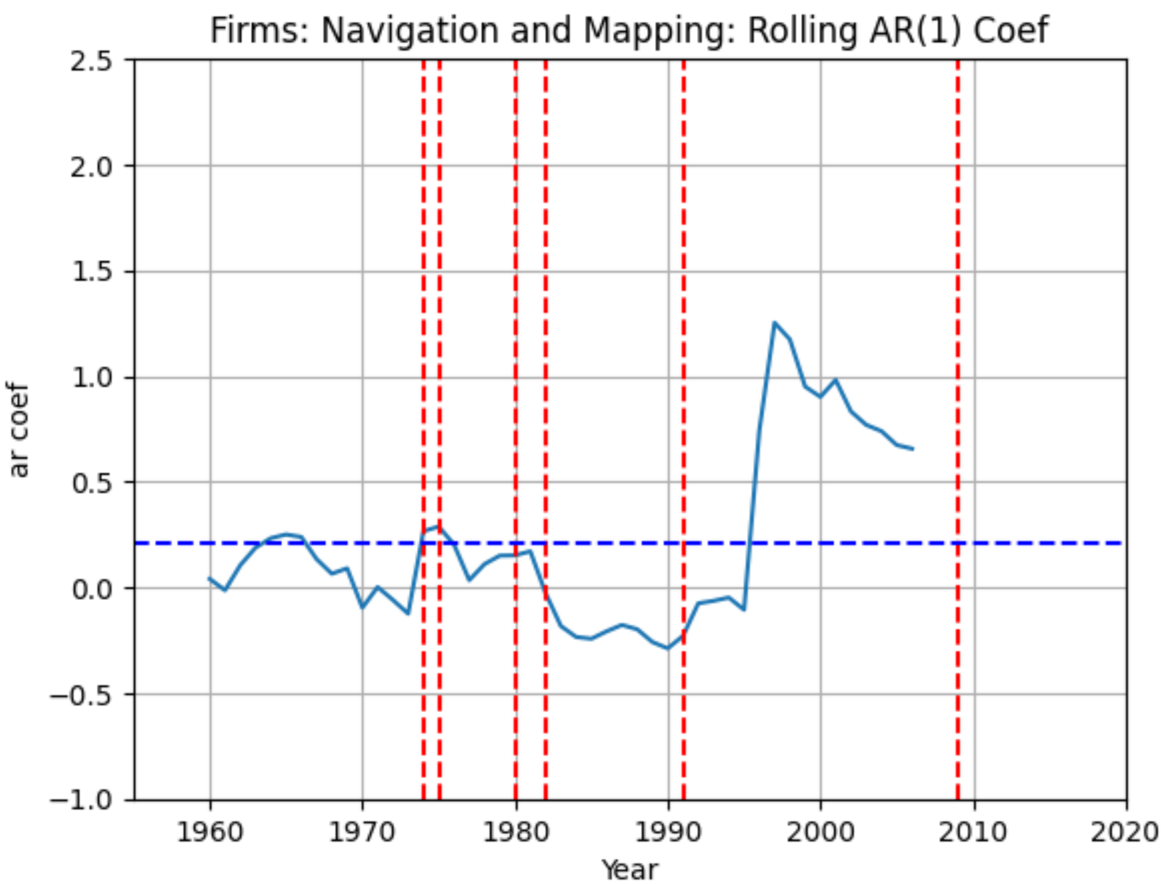

Firms: Media and Entertainment: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

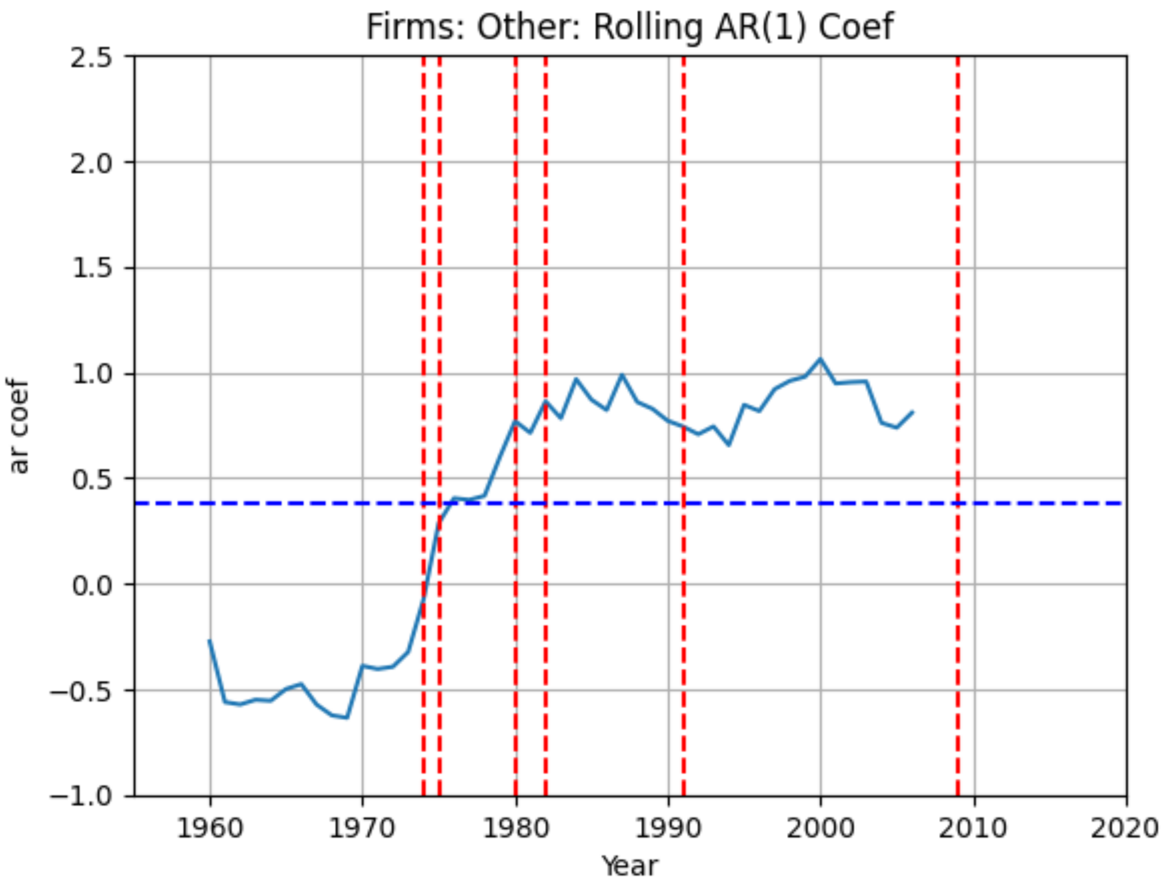

Firms: Messaging and Telecommunications: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

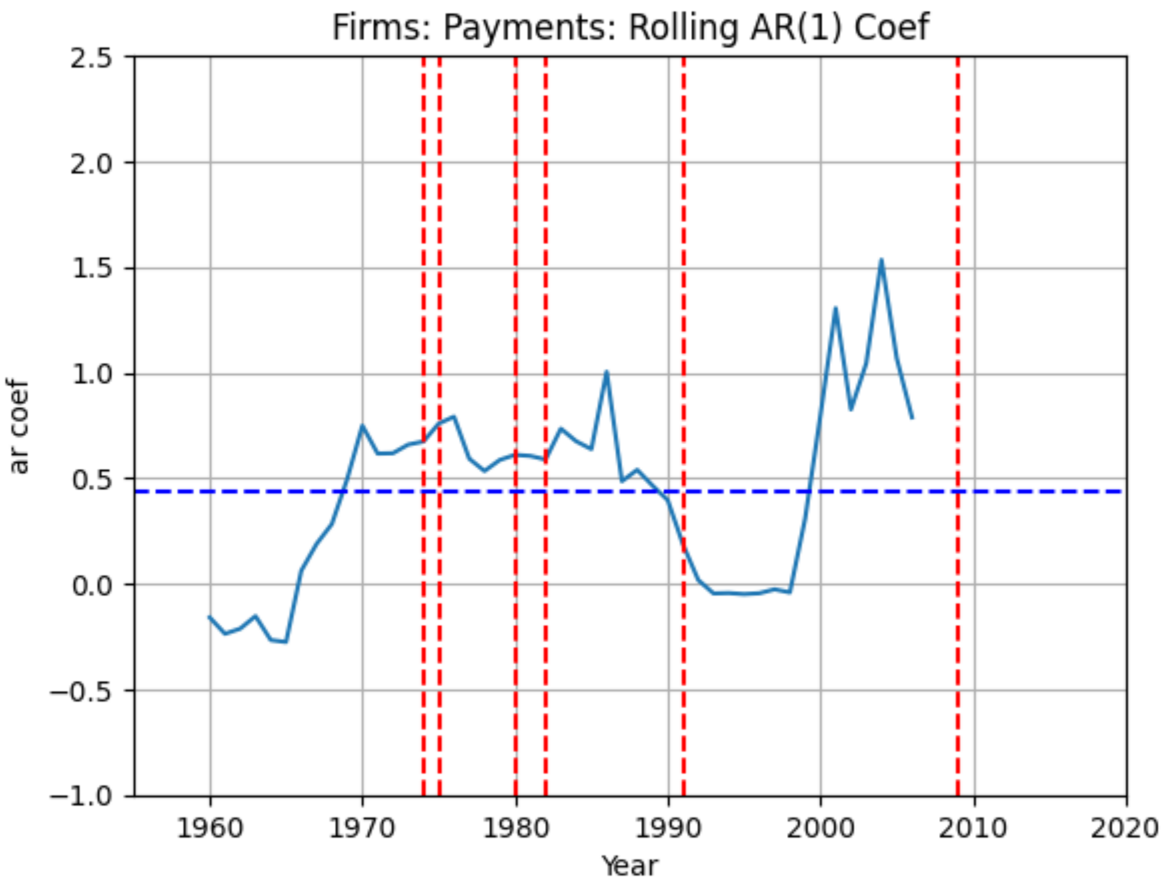

Firms: Mobile: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

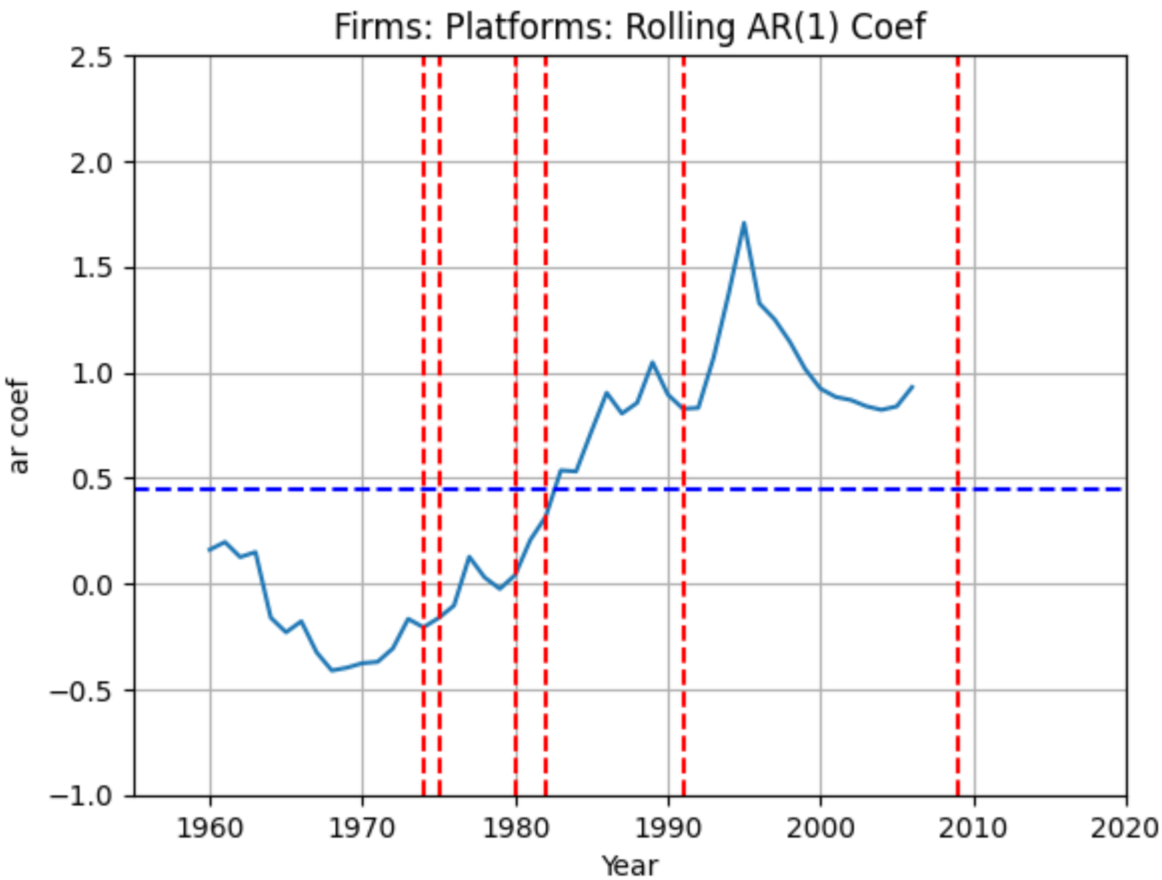

Firms: Music and Audio: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

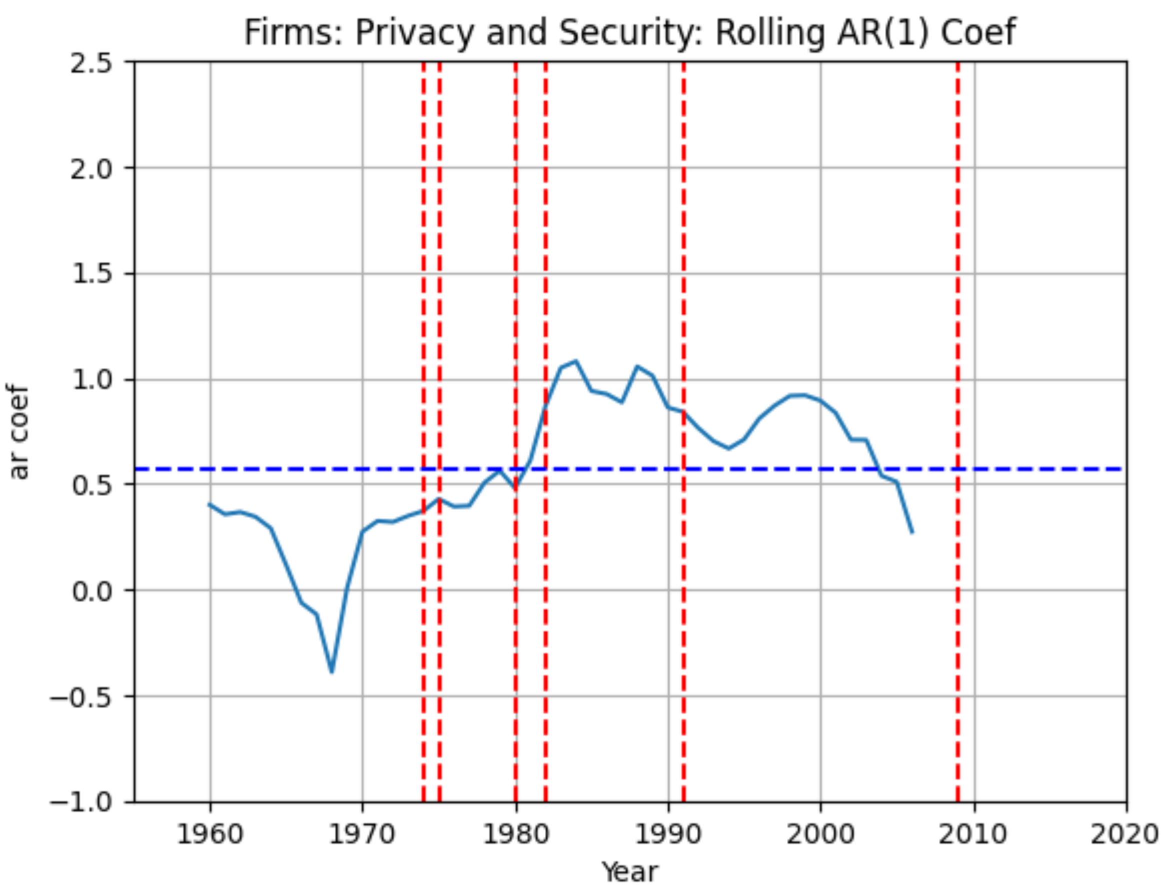

Firms: Natural Resources: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

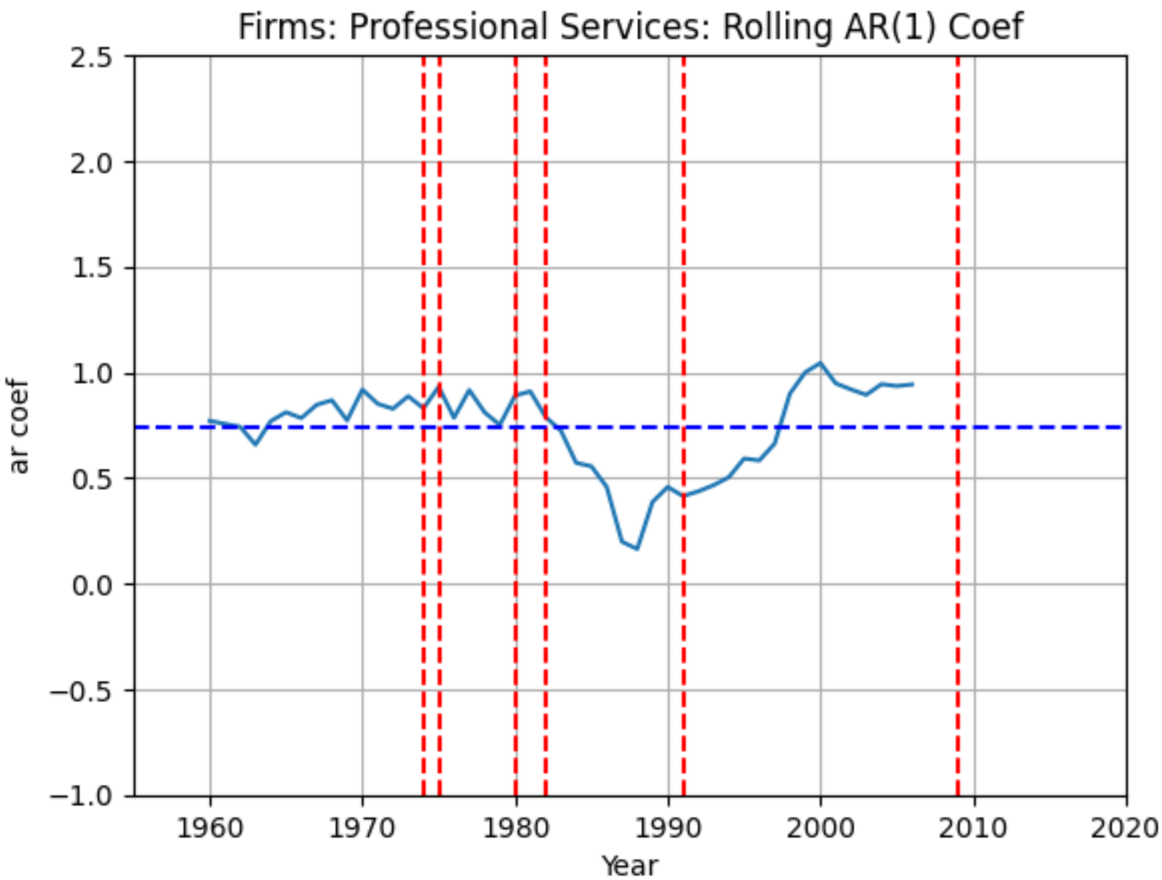

Firms: Navigation and Mapping: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

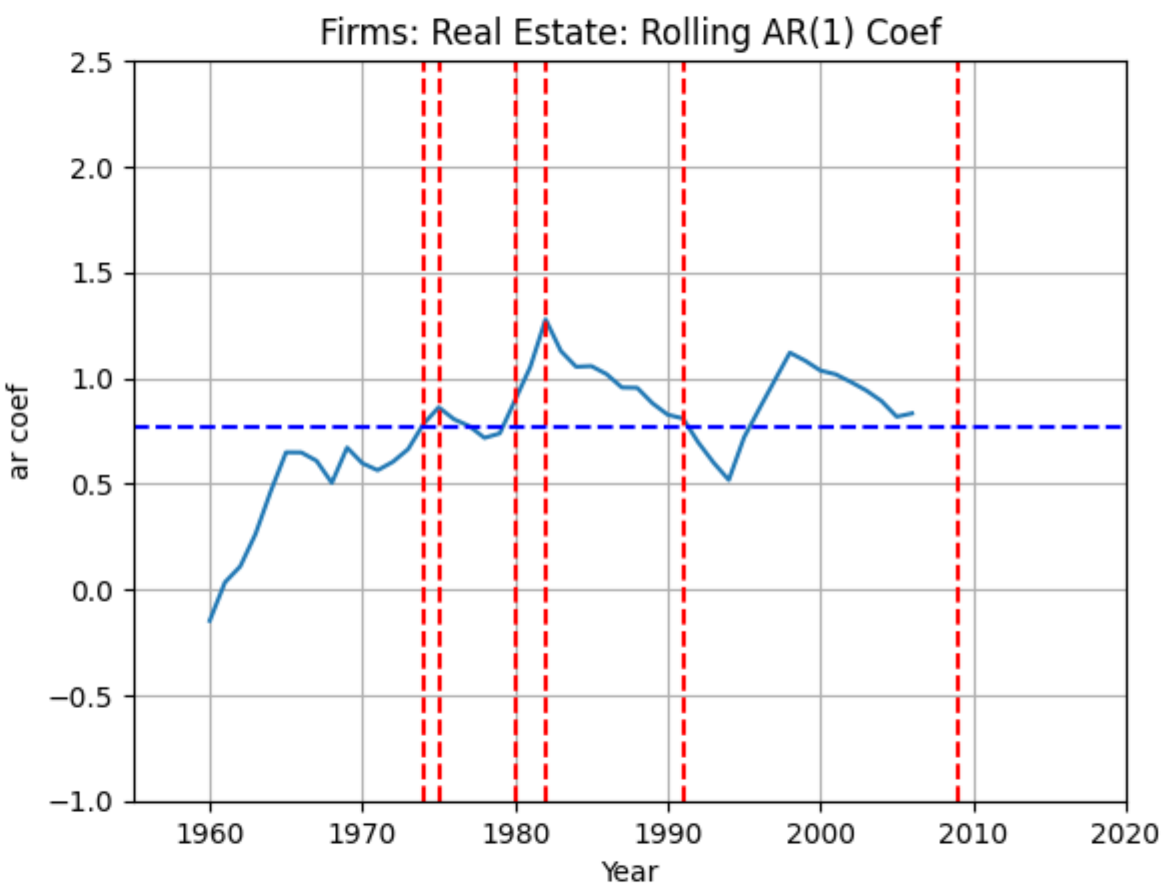

Firms: Other: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

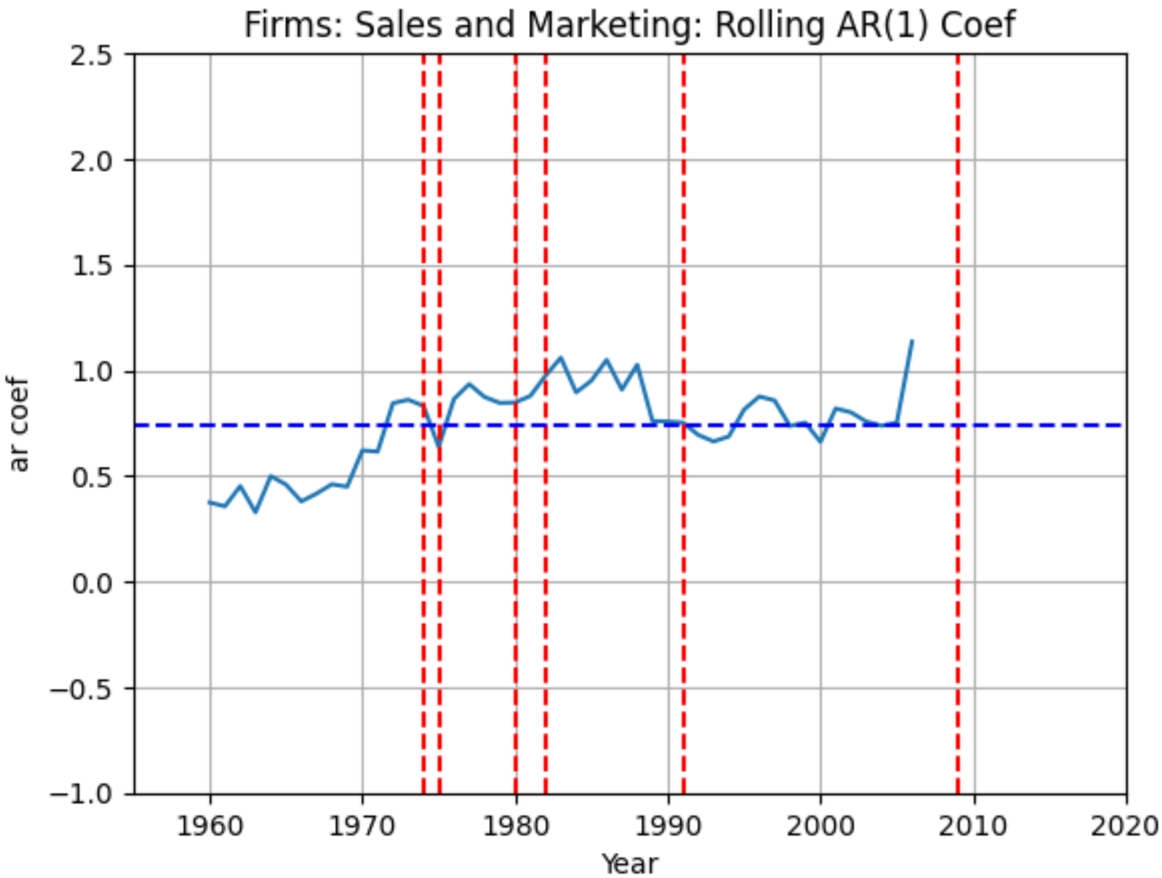

Firms: Payments: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year


Firms: Platforms: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

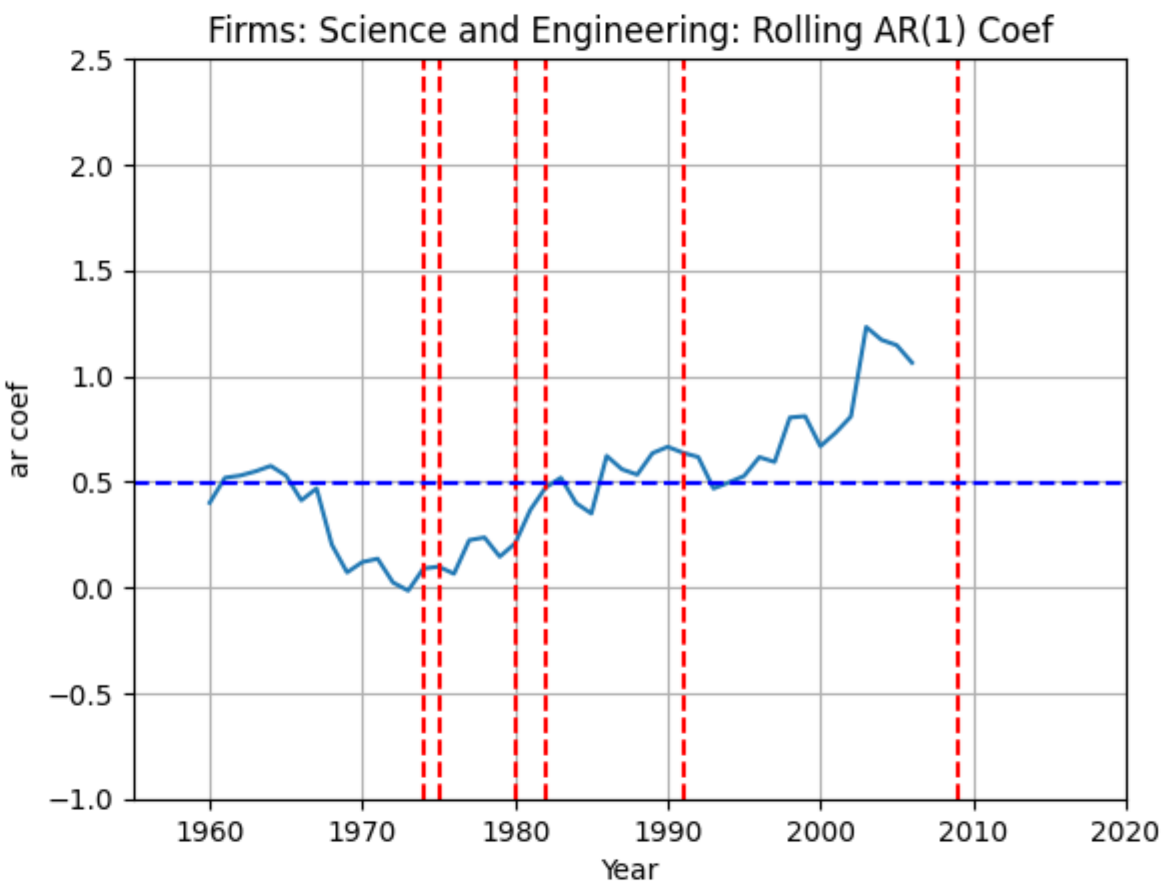

Firms: Privacy and Security: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Professional Services: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

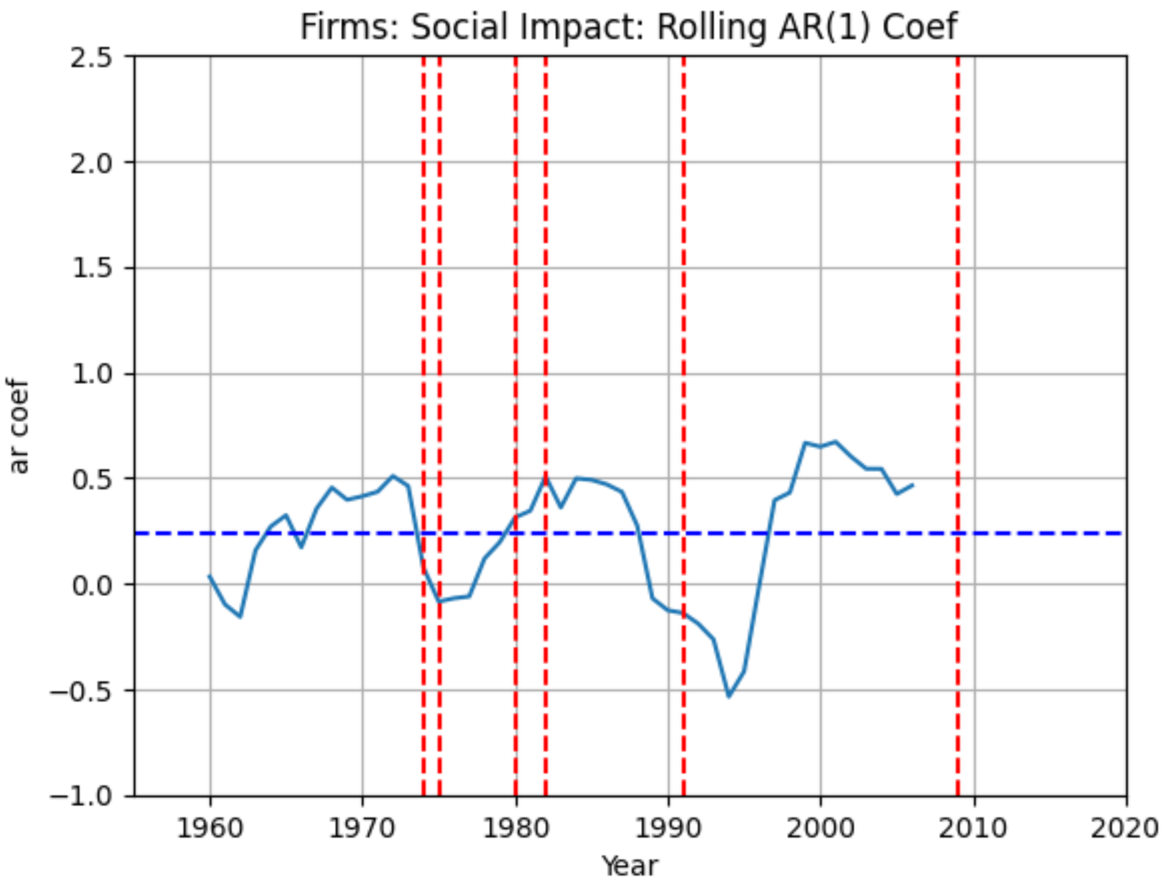

Firms: Real Estate: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

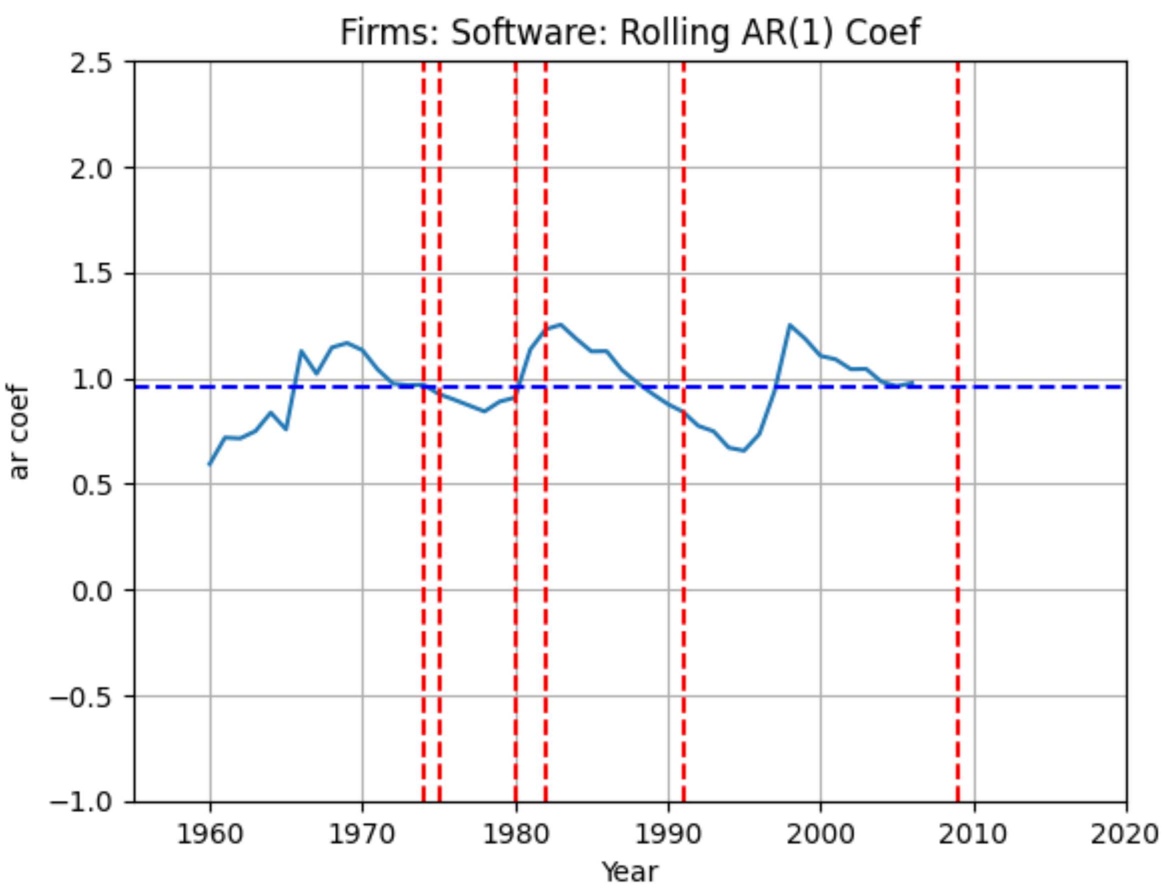


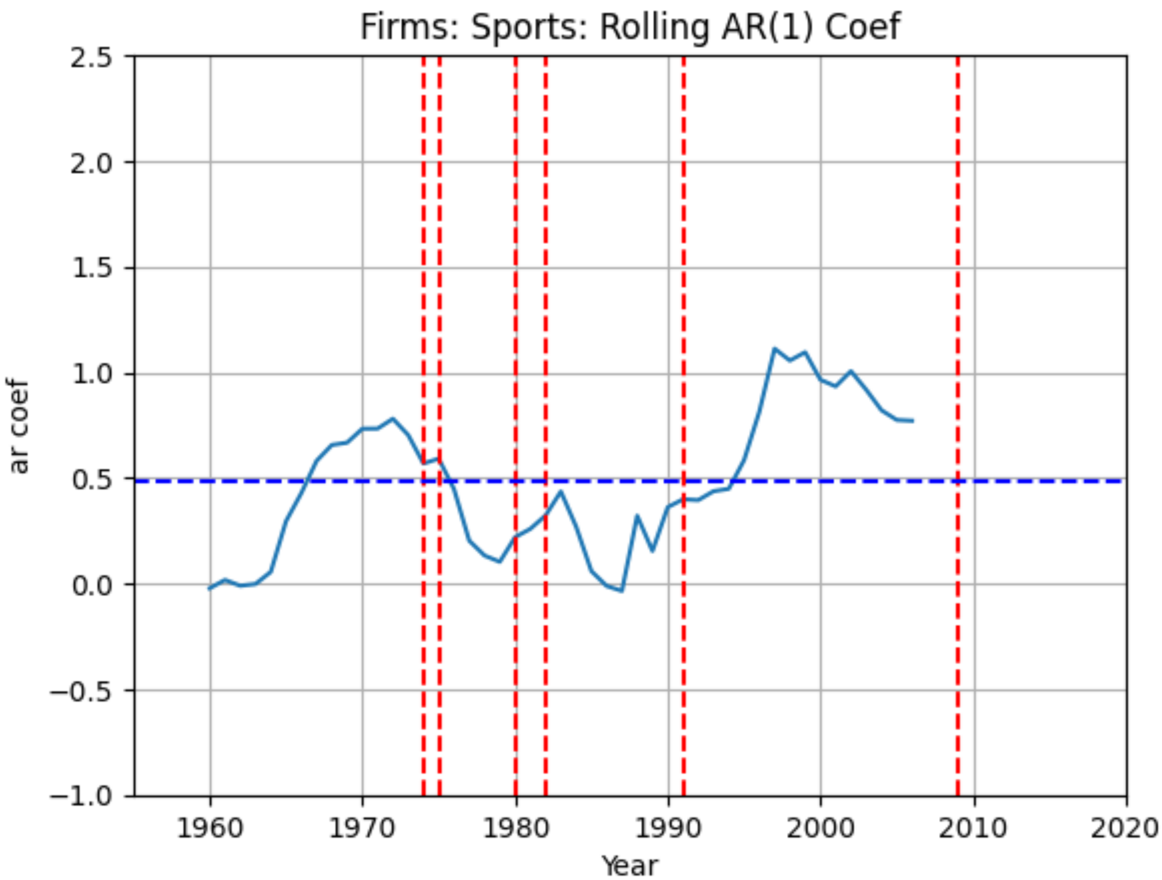
Firms: Sales and Marketing: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

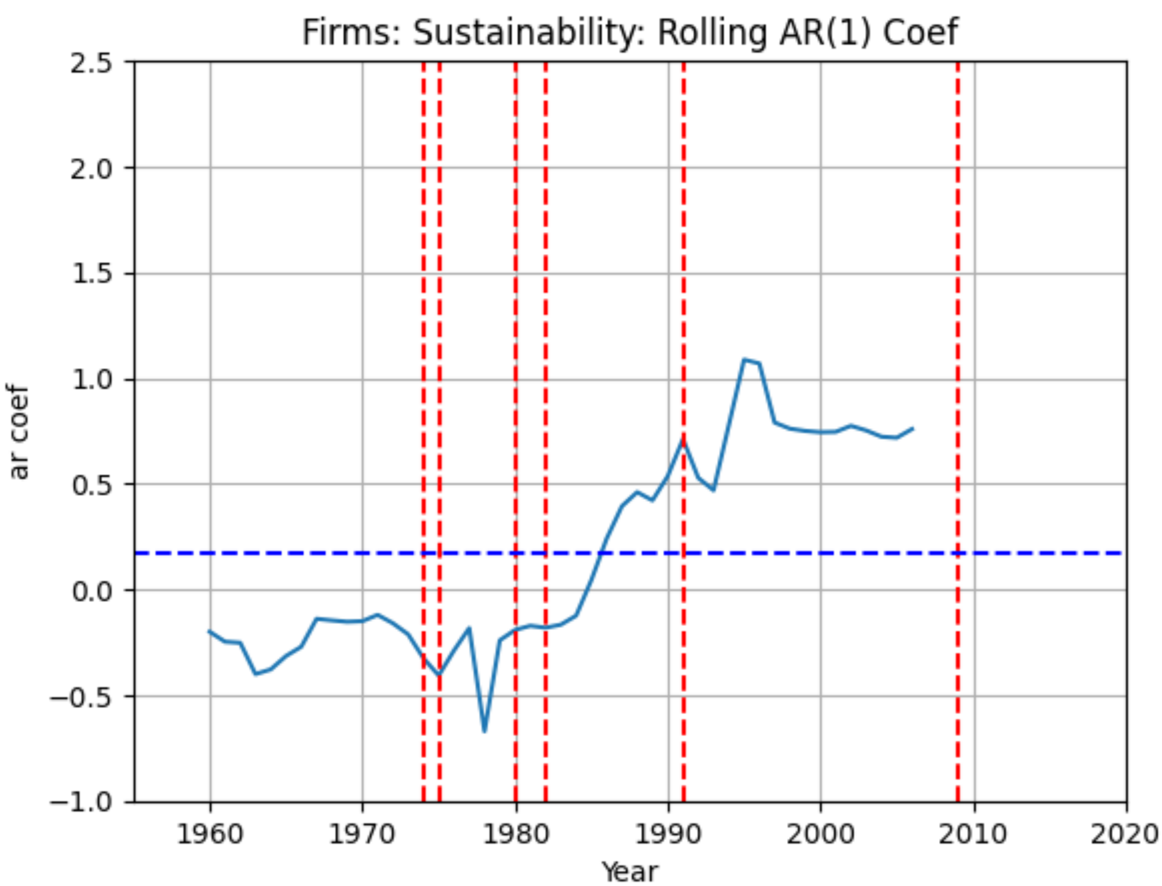


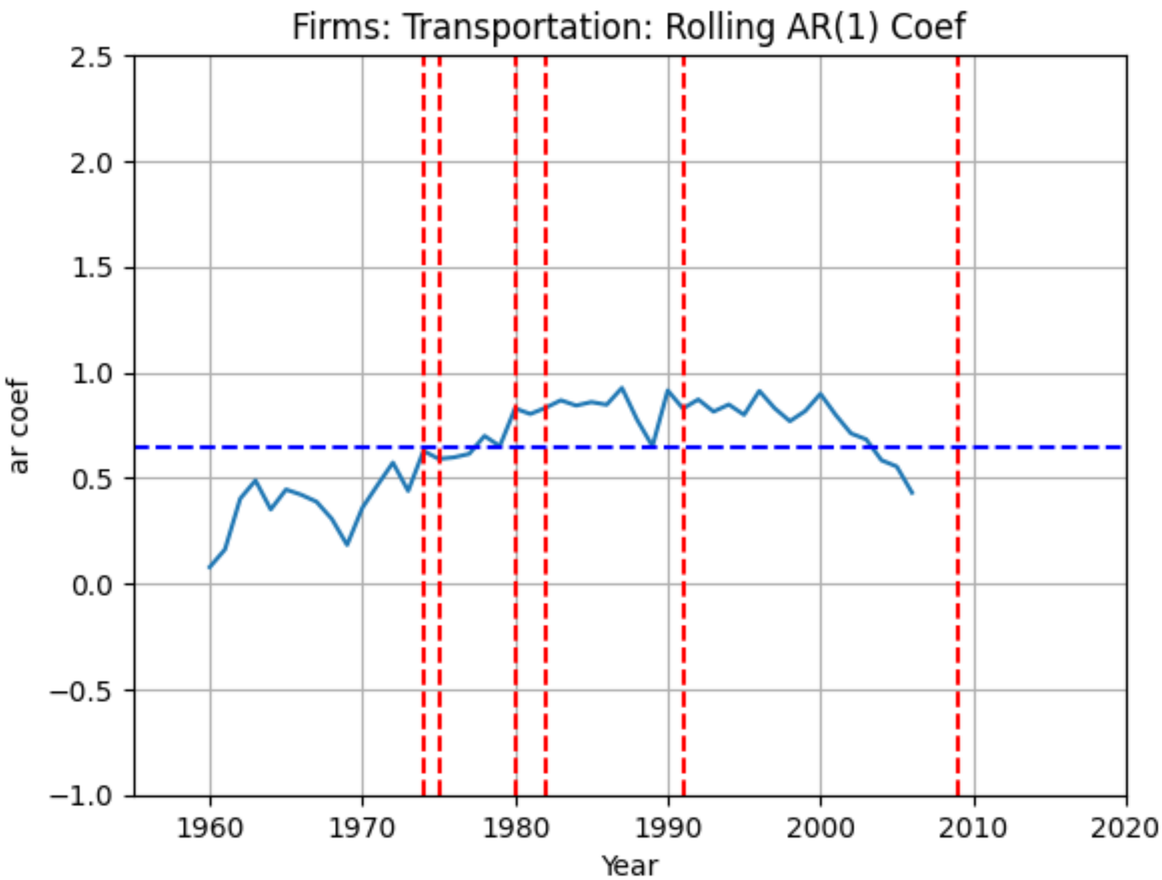
Firms: Science and Engineering: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year

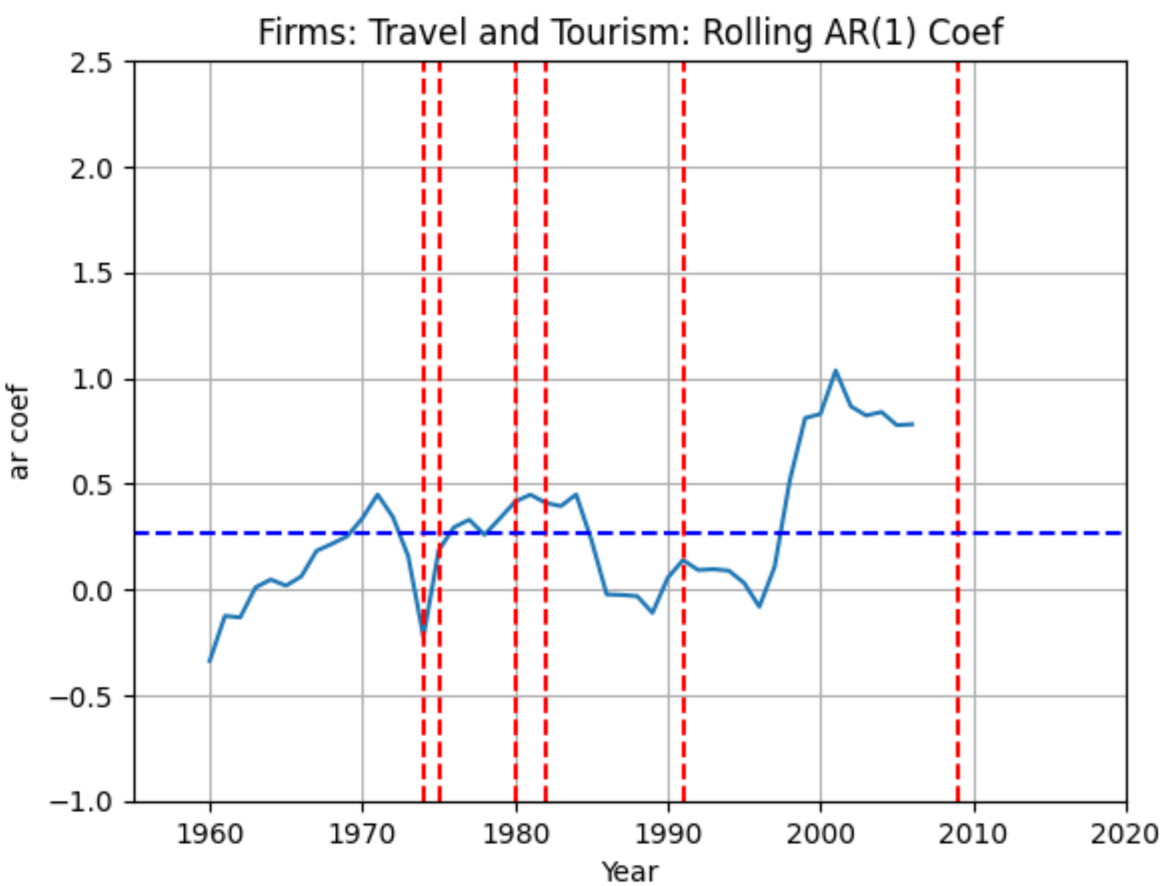

Firms: Social Impact: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020

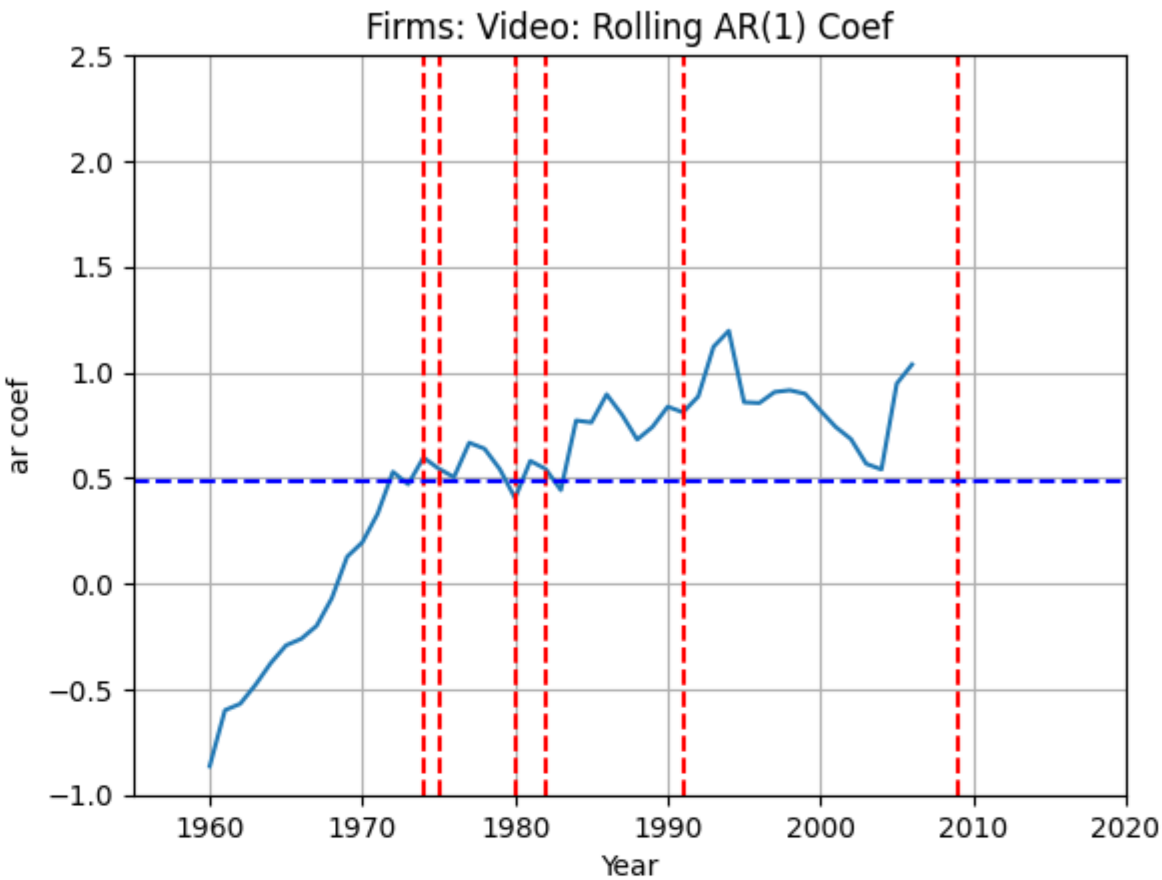

Firms: Software: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Sports: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Sustainability: Rolling AR(1) Coef
ar coef
Year
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
1960
1970
1980
1990
2000
2010
2020


Firms: Transportation: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year


Firms: Travel and Tourism: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year


Firms: Video: Rolling AR(1) Coef
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
ar coef
1960
1970
1980
1990
2000
2010
2020
Year