

s35v3 Pruned Network Momentum

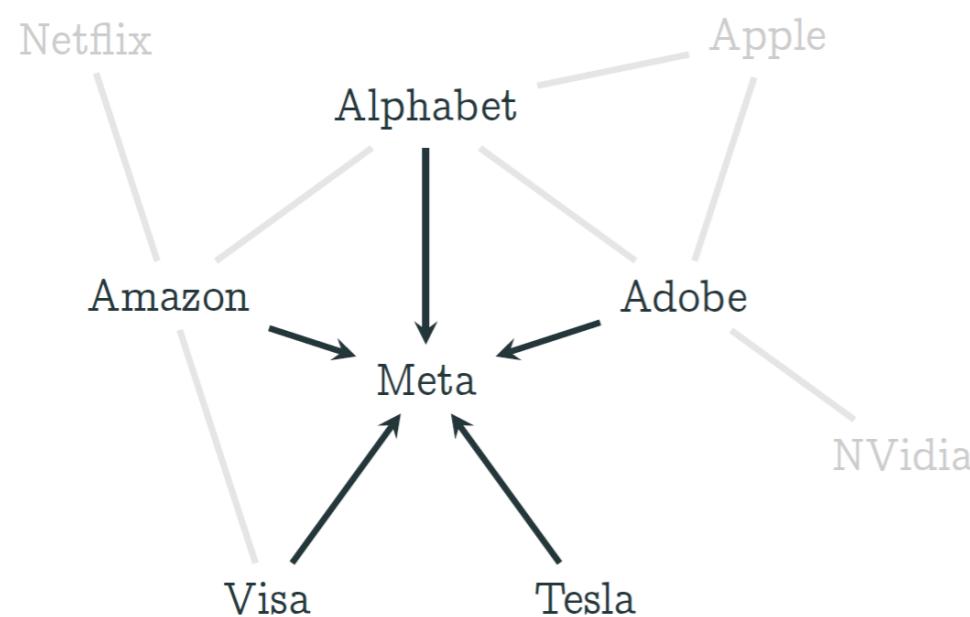
FA·Network

(Gautier Marti, Sebastien Andler, Yao Lei Xu, Patrick Yam, Stacy Pu)

Stacy Pu on May 24 2024

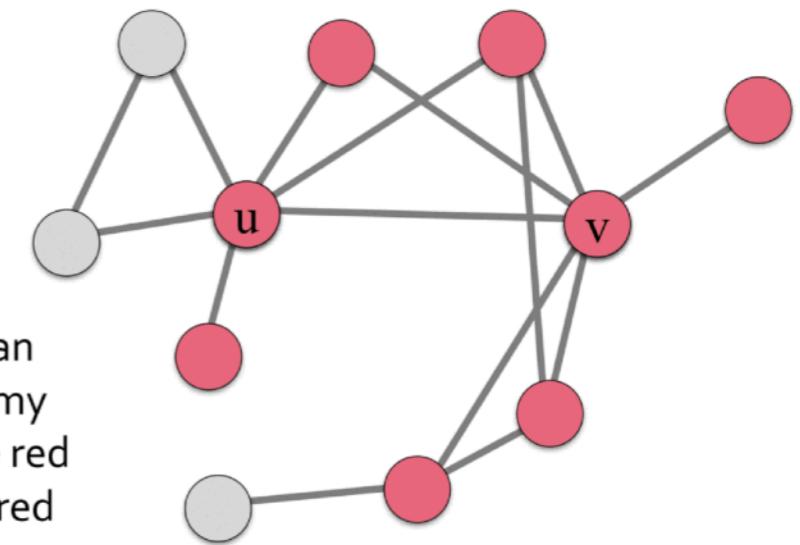
Motivation

- Network momentum works in a way propagating the **historical momentum** to similar peers.
 - The underlying rationale: the price of a target stock continues to rise if its peers have shown a strong performance in the recent past.
- To identify peers, we have access to many fundamental networks provided by data vendors.
 - Sources like Bloomberg/FactSet offer insights into supply chains, competitors, co-analysts, and co-mentions.
 - Can we rely entirely on them? Are there any inaccuracies or disturbances?
- Any noise in a node could potentially trigger a **cascading effect** that contaminates signal propagation throughout the entire network.



$$S = \{u, v\}$$

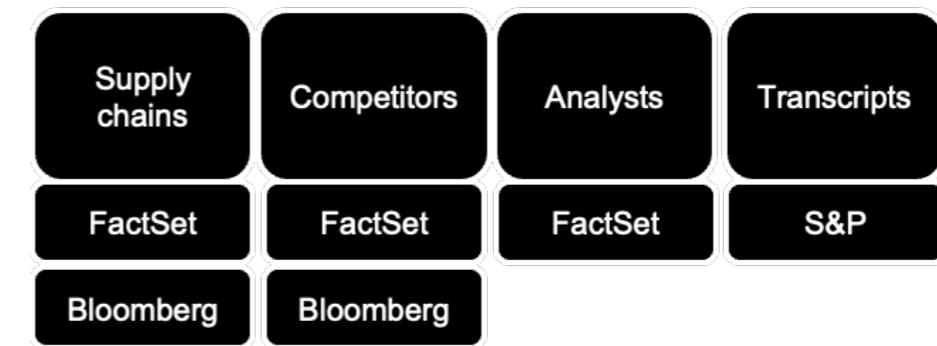
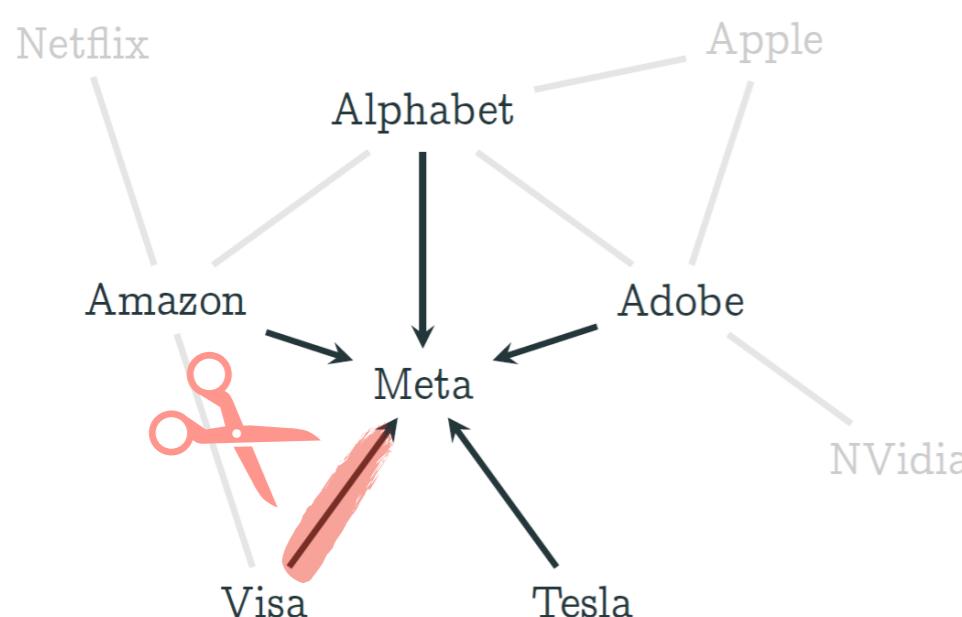
If more than
q=50% of my
friends are red
I'll also be red



Network Effects And Cascading Behaviour

Method

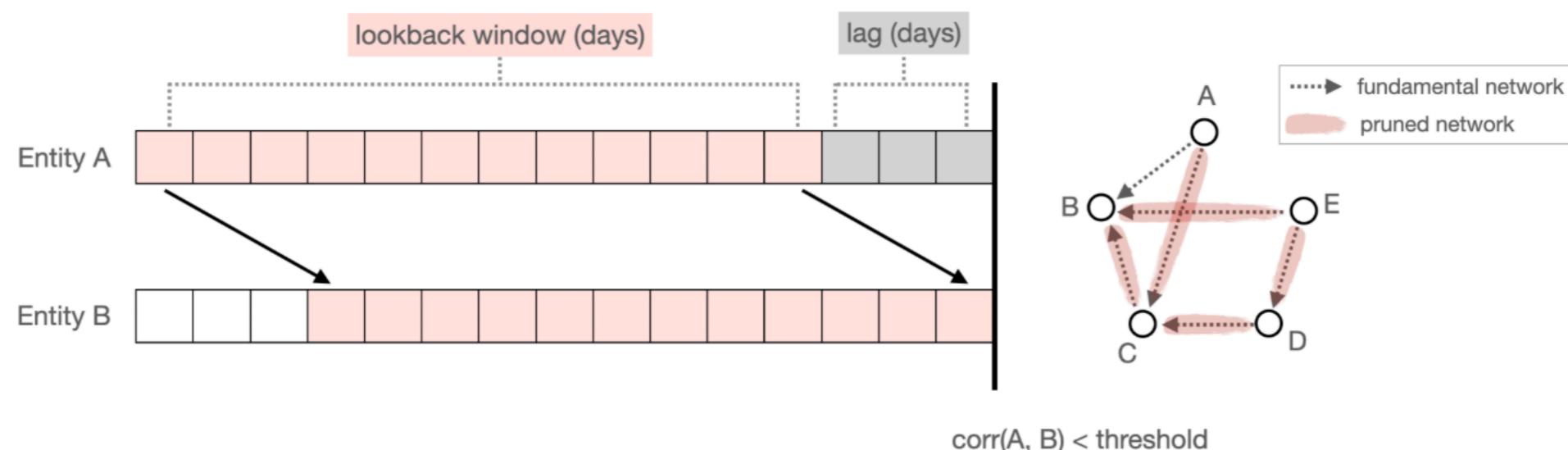
- There is **statistical evidence** to determine if one fundamental edge indeed incorporates lagged flow information.
 - We can analyze the lead-lag correlation among source-target pairs using fundamental network and market data.
- Pairs with a correlation below a certain threshold are excluded, resulting in a **pruned network**.
- Historical momentum is then propagated along these pruned networks, generating a new signal - **pruned network momentum**.



from_id	to_id	start_date	end_date	weight
BKCRJ1M-R	FHBMT4-R	2015-08-28	2016-03-11	<null>
S7R61T-R	FHBMT4-R	2015-01-19	2015-03-12	<null>
BCHCPG-R	G5M19V-R	2005-04-28	2005-11-24	<null>
JJ8KCS-R	G5M19V-R	2009-10-19	2010-07-09	<null>

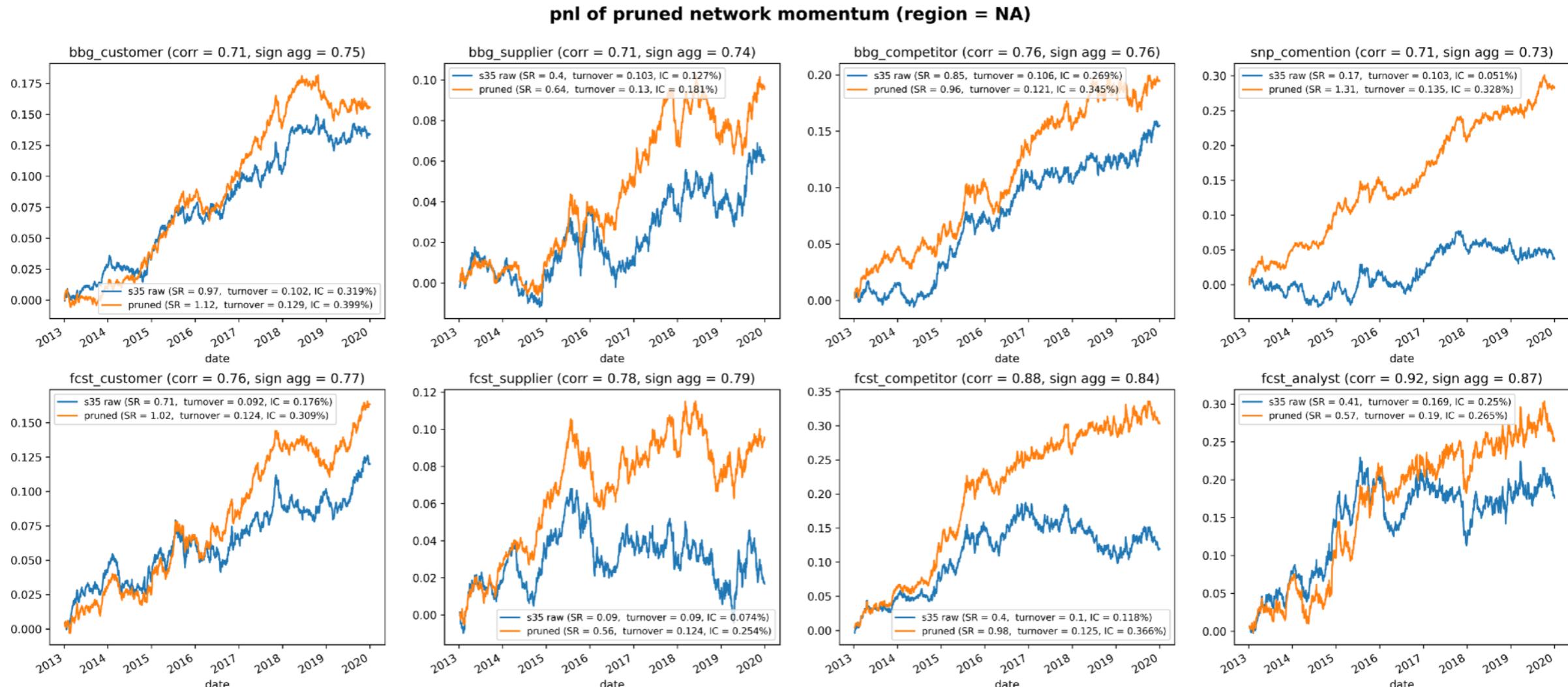
Method

- There is **statistical evidence** to determine if one fundamental edge indeed incorporates lagged flow information.
 - We can analyze the lead-lag correlation among source-target pairs using fundamental network and market data.
- Pairs with a correlation below a certain threshold are excluded, resulting in a **pruned network**.
- Historical momentum is then propagated along these pruned networks, generating a new signal - **pruned network momentum**.



Performance of Pruned Network Momentum

- Generally, the pruned network momentum outperformed the original network momentum from s35.
 - The **snp comention network** has observed the most outstanding improvement, from SR = 0.17 to 1.31.



- The performance was calculated from `fa/fa-signal/signal-analytics`: A tool for fast portfolio construction (fastPC) and signal analysis
<https://qromatiq.codes/fa/fa-signal/signal-analytics>
- network momentum signals → uniform ranking score and scaled into [-1,1] → fastPC (mean-variance optimisation with market neutral constraints w.r.t barra risk factors) → alpha

Understanding Performance Gains

- pp.13 • **Consistent performance across regions:** Whether there is consistent performance improvements across different fundamental networks and regions (NA/EU/AP).
- pp.24 • **Impact of number of edges removed:** Investigating the effect of number of edges removed from various parameter settings on signal performance.
- pp.26 • **Entity connections:** Analyzing the change in entity connections before and after pruning.
- pp.27 • **Influence of Q-universe filter:** Investigating the impact on signal performance when restricting source-target edge pairs to entities in the **GIGA** universe.
- pp.28 • **Similarity to GICS:** Exploring whether pruned networks exhibit greater similarity to GICS, potentially affecting alpha values.
- pp.29 • **Negative correlations:** What if we only keep the edges with negative correlations, would this be significantly worse than s35 network momentum?
- pp.30 • **Types of removed edges:** Qualitatively identifying the types of edges removed from the network.
- pp.35 • **Causal graph**

Squad Outputs & Next Step

1. Immediate product

Implementing it as an enhanced signals to replace s35v1.

2. Long-term output

Exploring pruning's potential as a submodule in FA/networks, systematically applied to all networks and base signals.

This technique could generate new signals for broader network strategies.

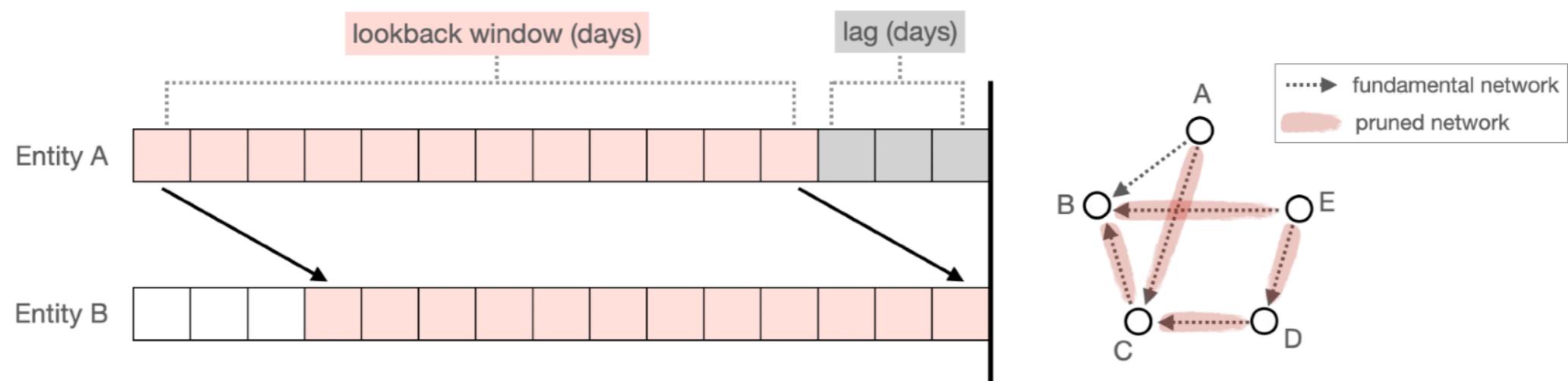
3. ML extensions

Considering re-weighting, ML-based pruning, and edge imputation as natural extensions. Exploring correlations as a new edge feature that could enrich GNN projects.

More Technical Details

Settings and Parameters

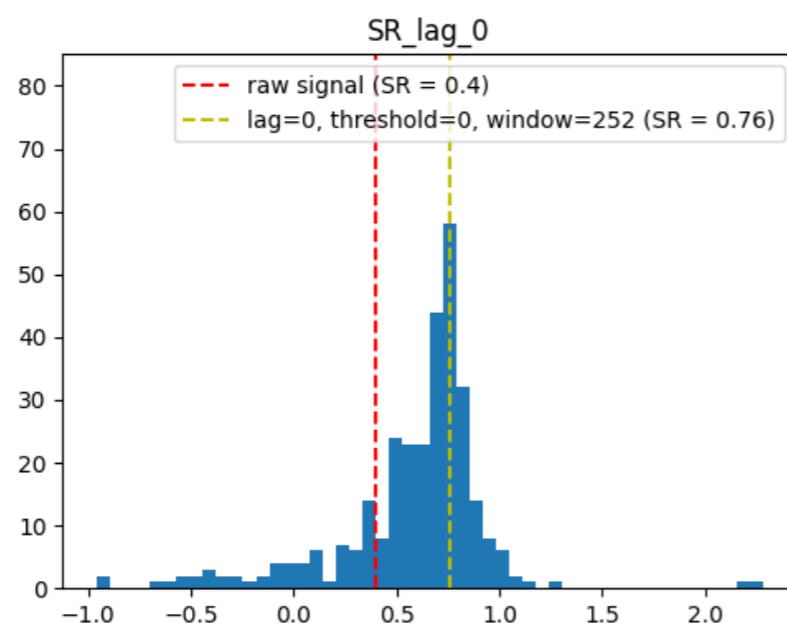
- Returns to calculate the correlations
 - **raw return** (from `qfin.factset_market_data`)
 - residual returns (from `fa.feature_store`)
 - **gemtl_country_residual_return** = raw - global_market_contrib_factors - country_contrib_factors
 - **gemtl_industry_residual_return** = gemtl_country_residual_return - industry_contrib_factors
 - **specific_return** = gemtl_industry_residual_return - barra_risk_factor
- Lookback windows (`lookback_window = [252, 504, 756, 1008, 1260]`)
- Lead-lag windows (`lag = [0, 1, 7, 21, 63, 126, 252]`)
- Threshold to remove edges
 - Too big - information loss | Too small - noisy edges kept
 - `threshold = [0., 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2]`
 - Should we choose a threshold that removed `5%` edges?



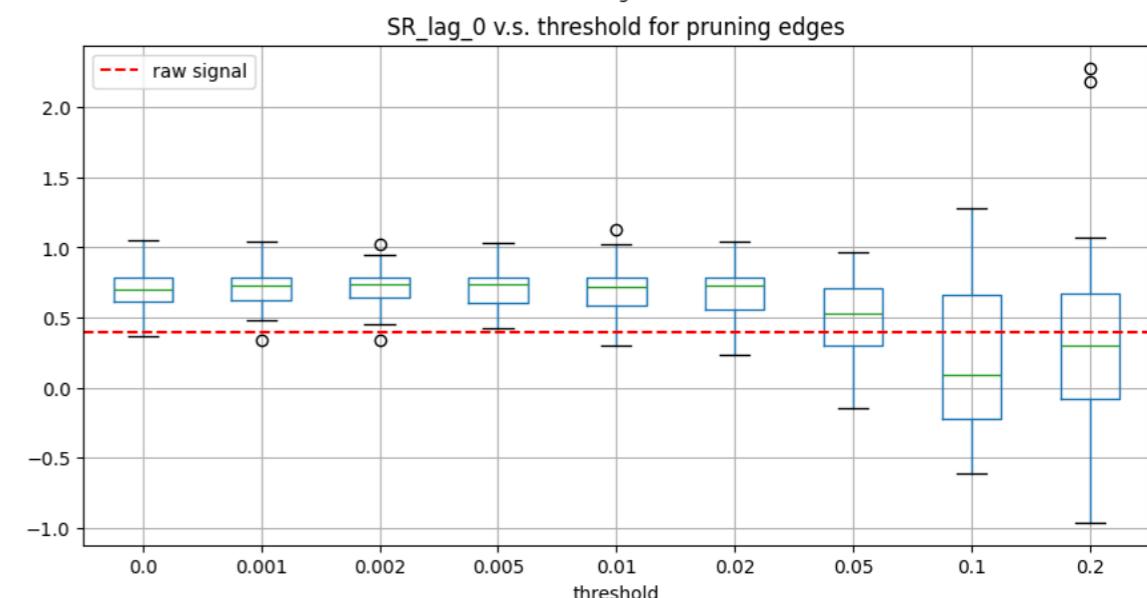
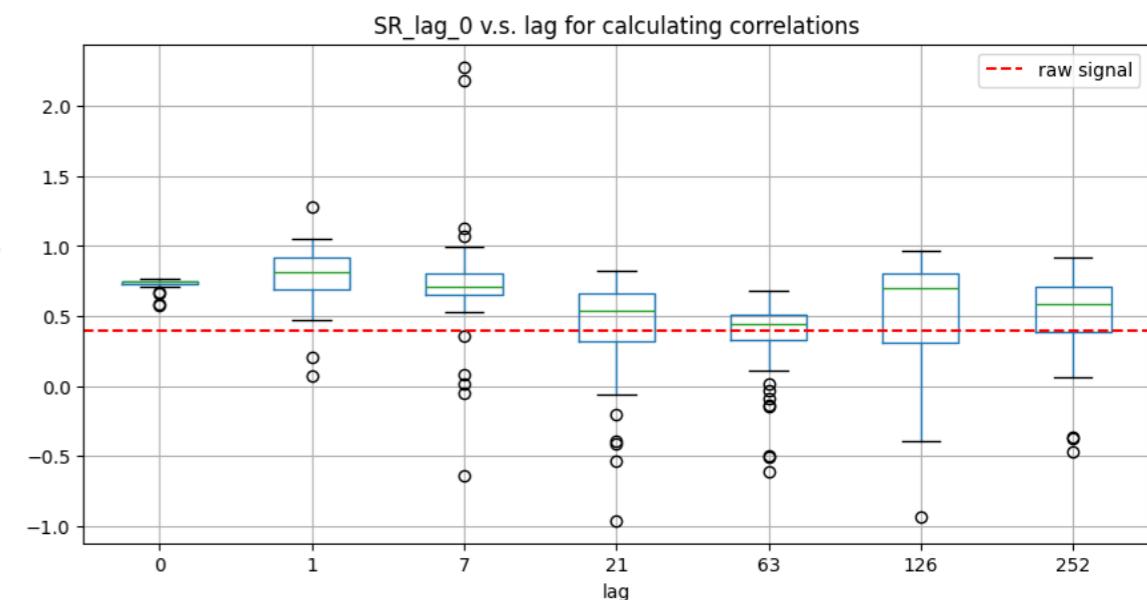
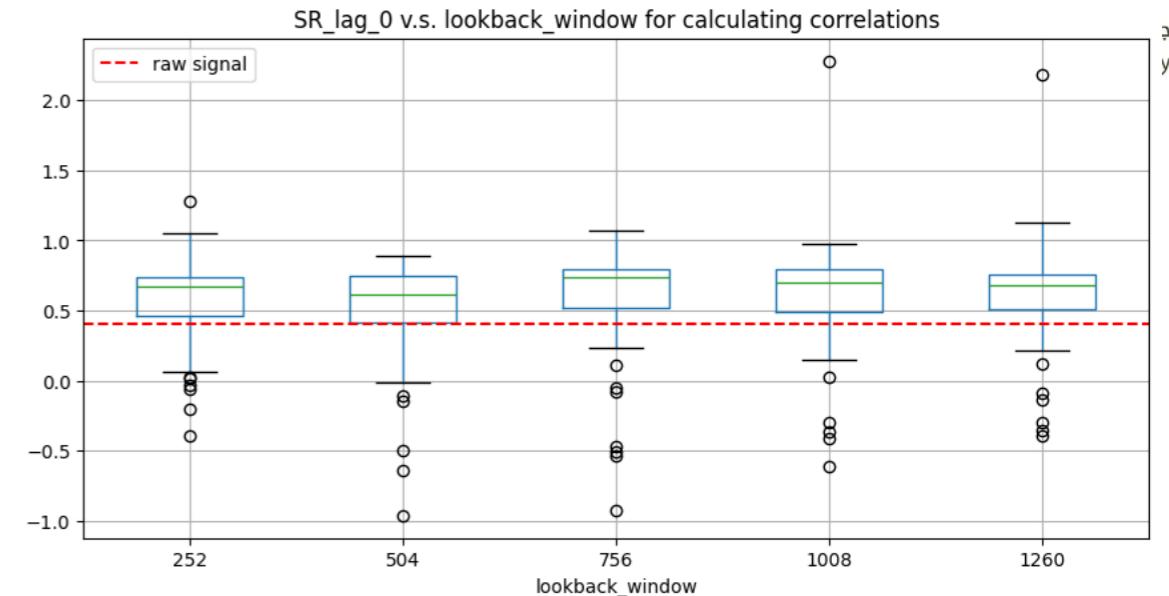
The Impact of Parameters

- Settings
 - Fundamental network: `BBG Competitor Network`
 - Raw signals from `s6 - qfa.staging_store where signal_id = 10`)
 - Raw returns to calculate correlations (from `qfin.factset_market_data`)
 - `SR_lag_0` (Sharpe Ratio of targeting future 1d return)

- Findings
 - Removing pairs with a synchronous correlation < 0 improves the SR.
 - The majority of parameter settings are expected to result in a higher SR compared to the raw signals.
 - The number of days (`lookback_window`) used for calculating correlations does not significantly impact the outcome.
 - Removing too many edges (by setting a high `threshold > 0.01`) results in a lower SR.



10



More Results & Findings

Results & Findings

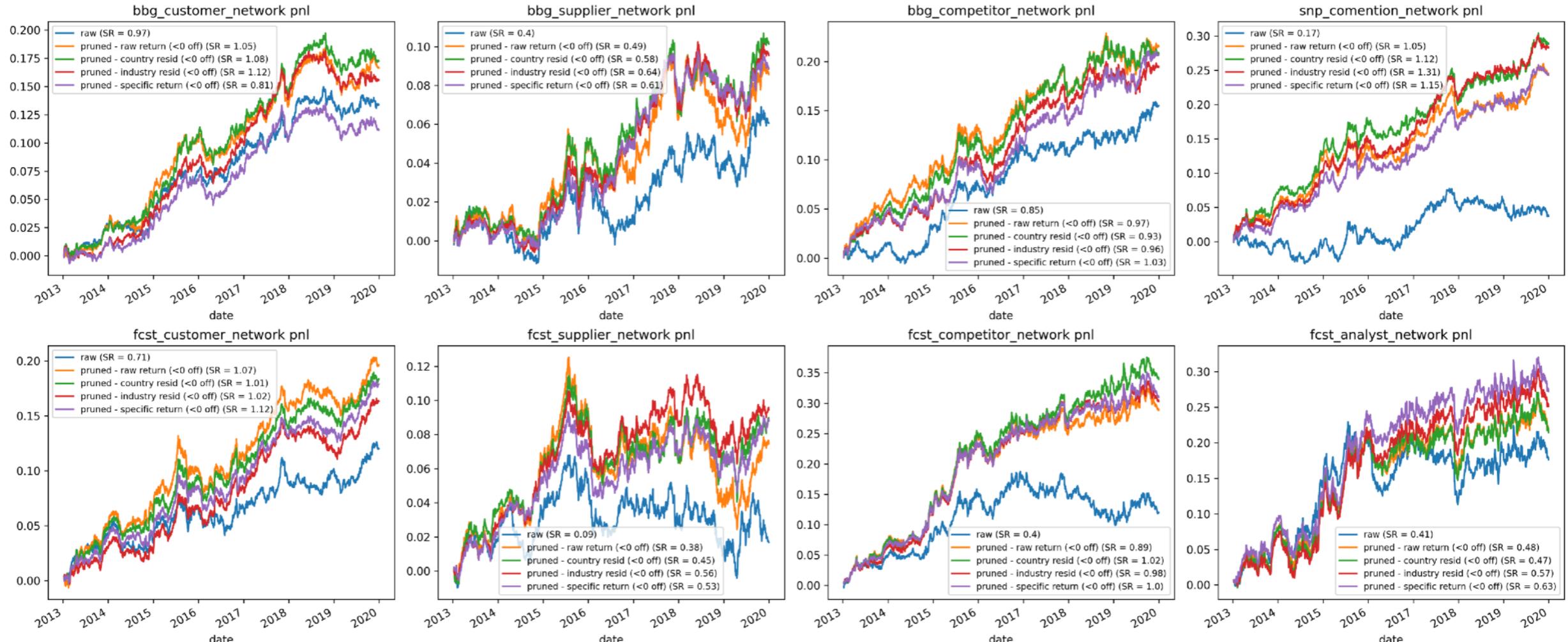
Research Questions

- Is the PnL performance consistent across different time / fundamental networks / regions (NA/EU/AP)?
- Which returns are most suitable for calculating correlations? Do they result in different PnL outcomes?

Research Questions

- Is the PnL performance consistent across different time / fundamental networks / regions (NA/EU/AP)?
- Which returns are most suitable for calculating correlations? Do they result in different PnL outcomes?

pnl of pruned network momentum (region = NA)



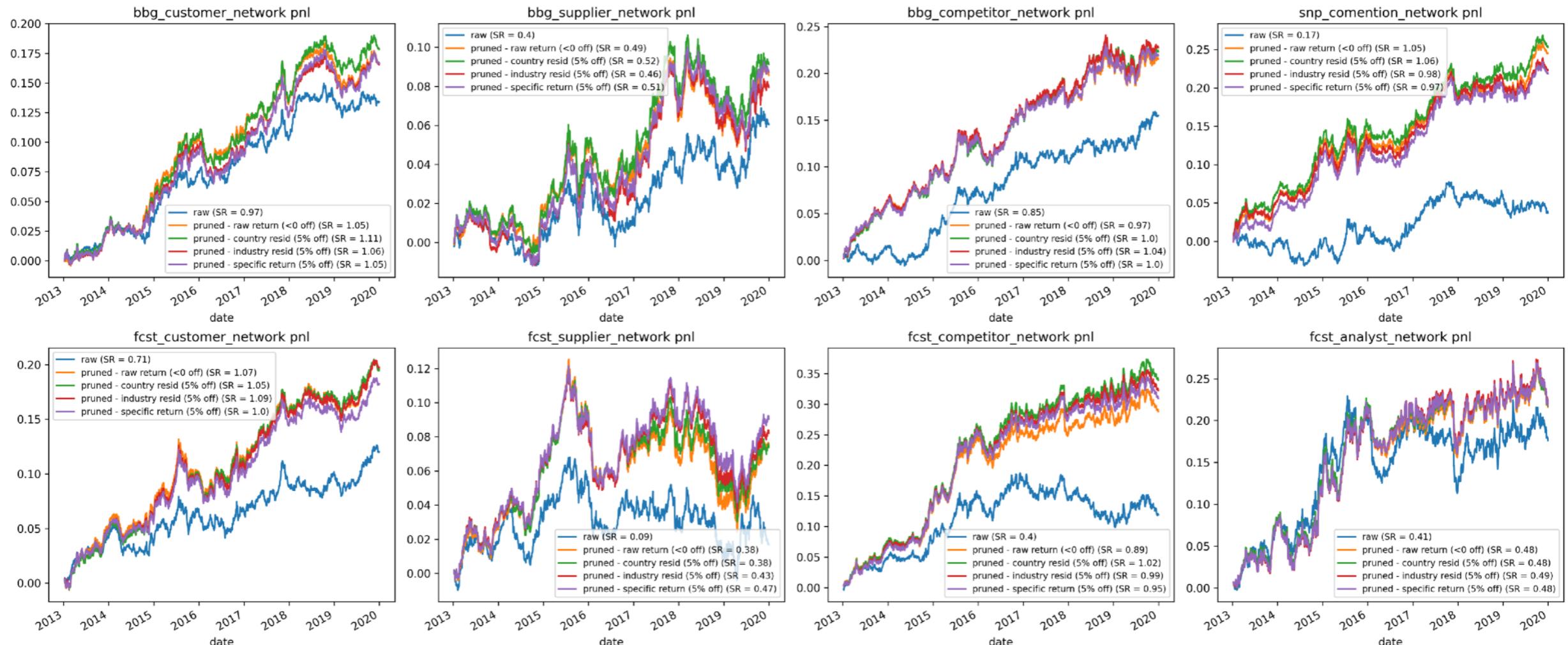
- The choice of `return types` to calculate the correlations and `thresholds` determined the number of connections removed, which further influences the pnl.

o In AP and EU, `removing 5% edges` futher outperformed `removing edges with a correlation < 0` in residual returns.

In NA, `removing edges with a correlation < 0` was better than `removing 5% edges`.

Companies of AP and EU have **fewer** connections than those of NA.

pnl of pruned network momentum (region = NA)



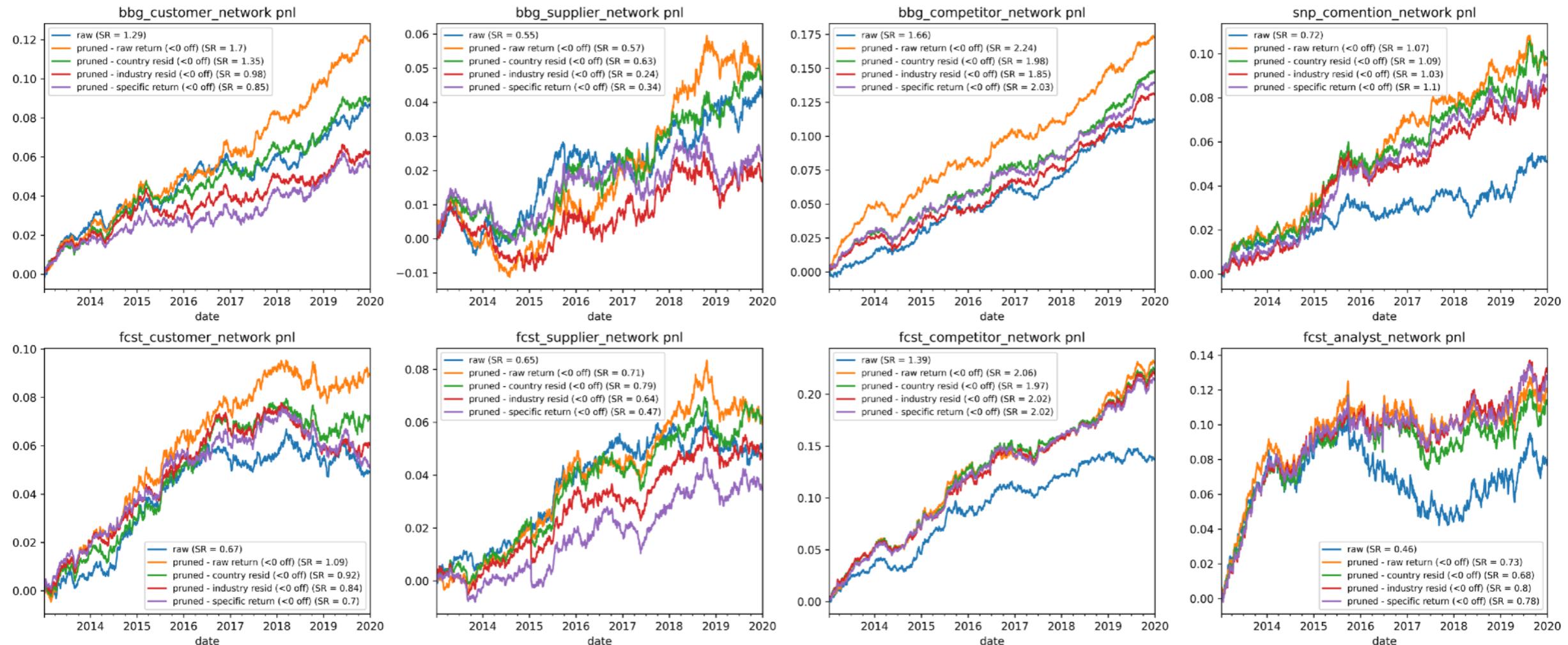
- The choice of `return types` to calculate the correlations and `thresholds` determined the number of connections removed, which further influences the pnl.

o In AP and EU, `removing 5% edges` futher outperformed `removing edges with a correlation < 0` in residual returns.

In NA, `removing edges with a correlation < 0` was better than `removing 5% edges`.

Companies of AP and EU have **fewer** connections than those of NA.

pnl of pruned network momentum (region = EU)



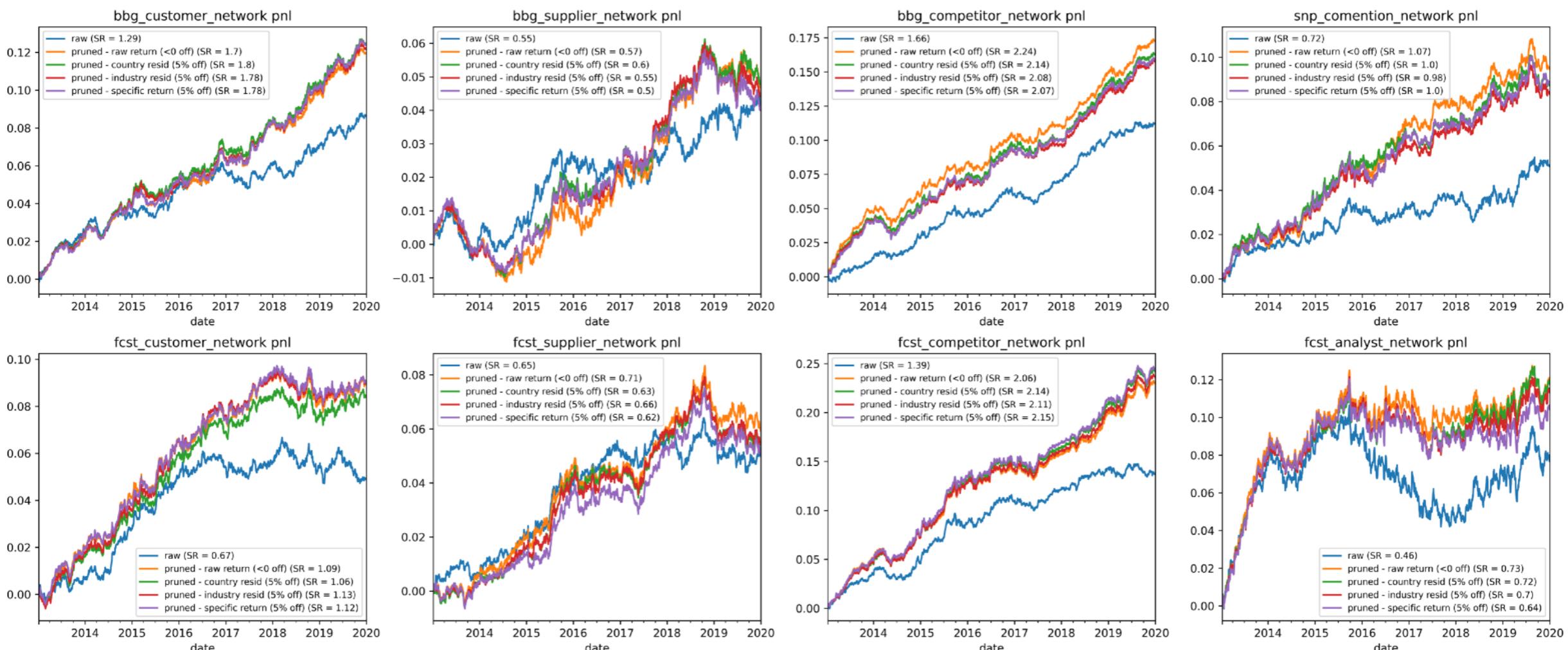
- The choice of `return types` to calculate the correlations and `thresholds` determined the number of connections removed, which further influences the pnl.

◦ In AP and EU, `removing 5% edges` further outperformed `removing edges with a correlation < 0` in residual returns.

In NA, `removing edges with a correlation < 0` was better than `removing 5% edges`.

Companies of AP and EU have **fewer** connections than those of NA.

pnl of pruned network momentum (region = EU)

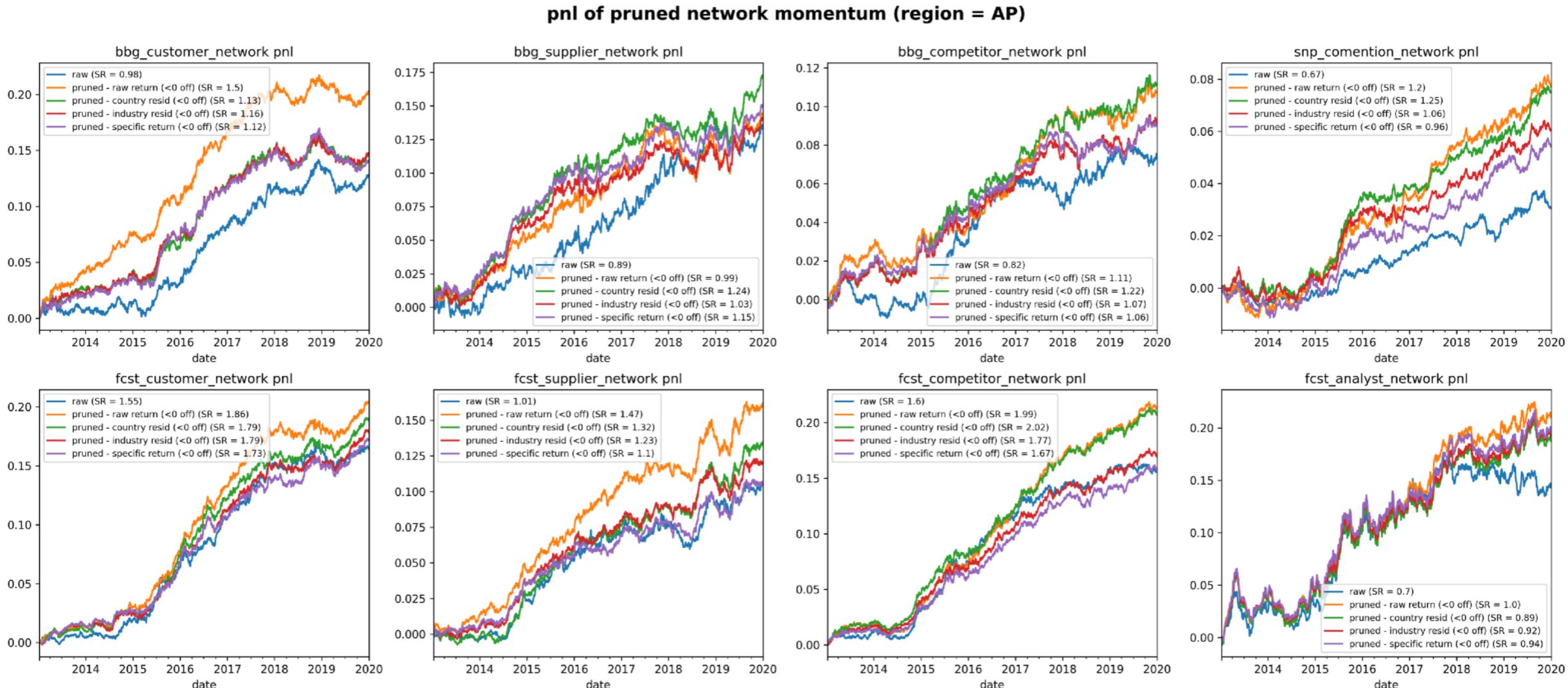


- The choice of `return types` to calculate the correlations and `thresholds` determined the number of connections removed, which further influences the pnl.

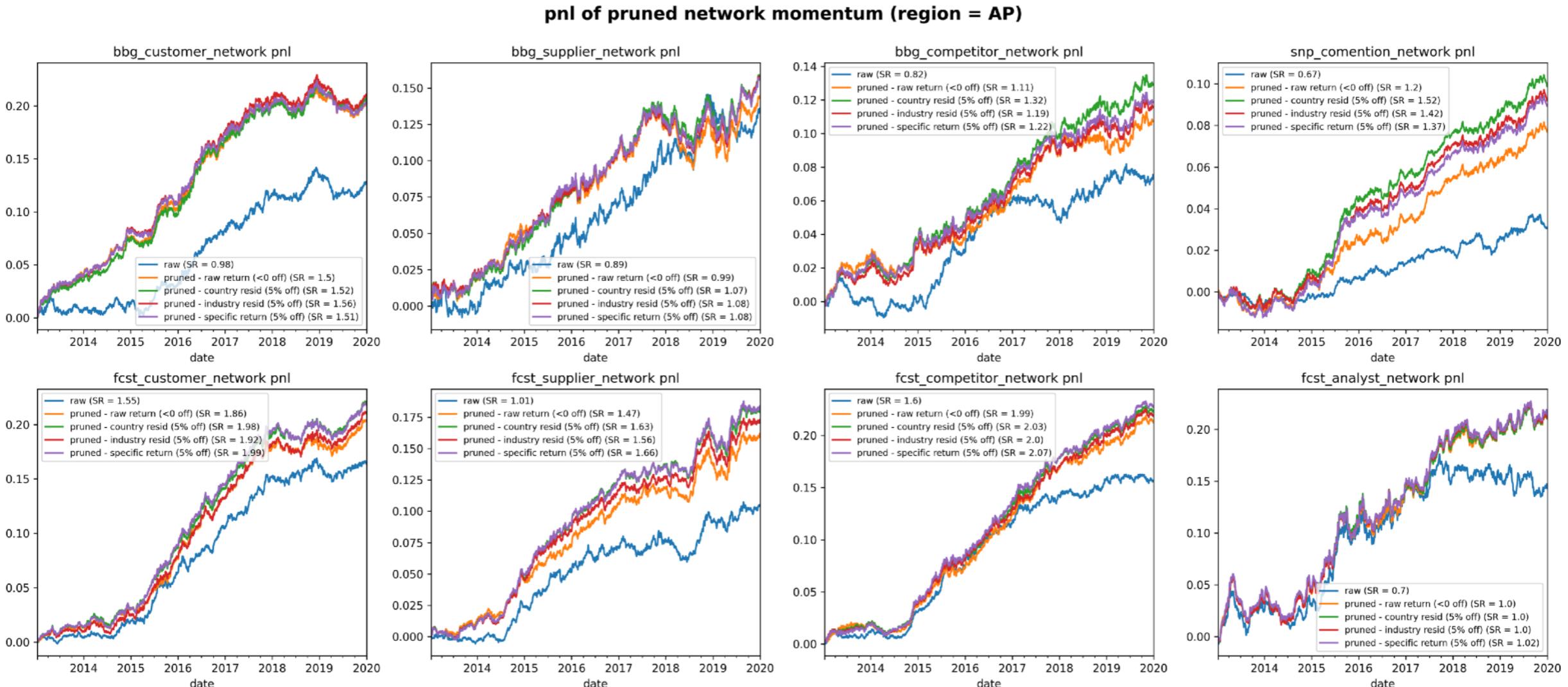
o In AP and EU, `removing 5% edges` futher outperformed `removing edges with a correlation < 0` in residual returns.

In NA, `removing edges with a correlation < 0` was better than `removing 5% edges`.

Companies of AP and EU have **fewer** connections than those of NA.



- The choice of `return types` to calculate the correlations and `thresholds` determined the number of connections removed, which further influences the pnl.
 - In AP and EU, `removing 5% edges` futher outperformed `removing edges with a correlation < 0` in residual returns.
 - In NA, `removing edges with a correlation < 0` was better than `removing 5% edges`.
 - Companies of AP and EU have **fewer** connections than those of NA.



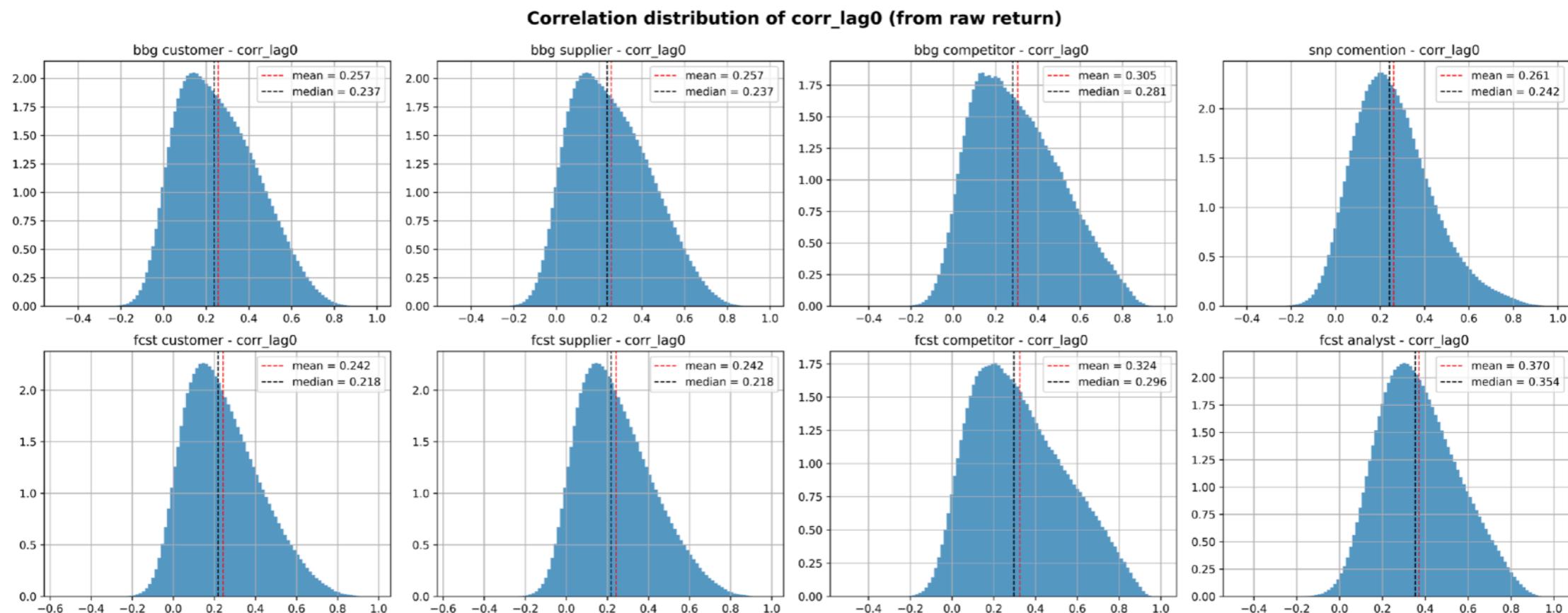
Results & Findings

Research Questions

- What are the correlation distributions of different fundamental networks?
- How many edges removed/kept if we set `threshold=0` ?
- How many connections (aka. node degree) an entity has before/after pruning?

Robustness Analysis: Threshold v.s. Sparsity

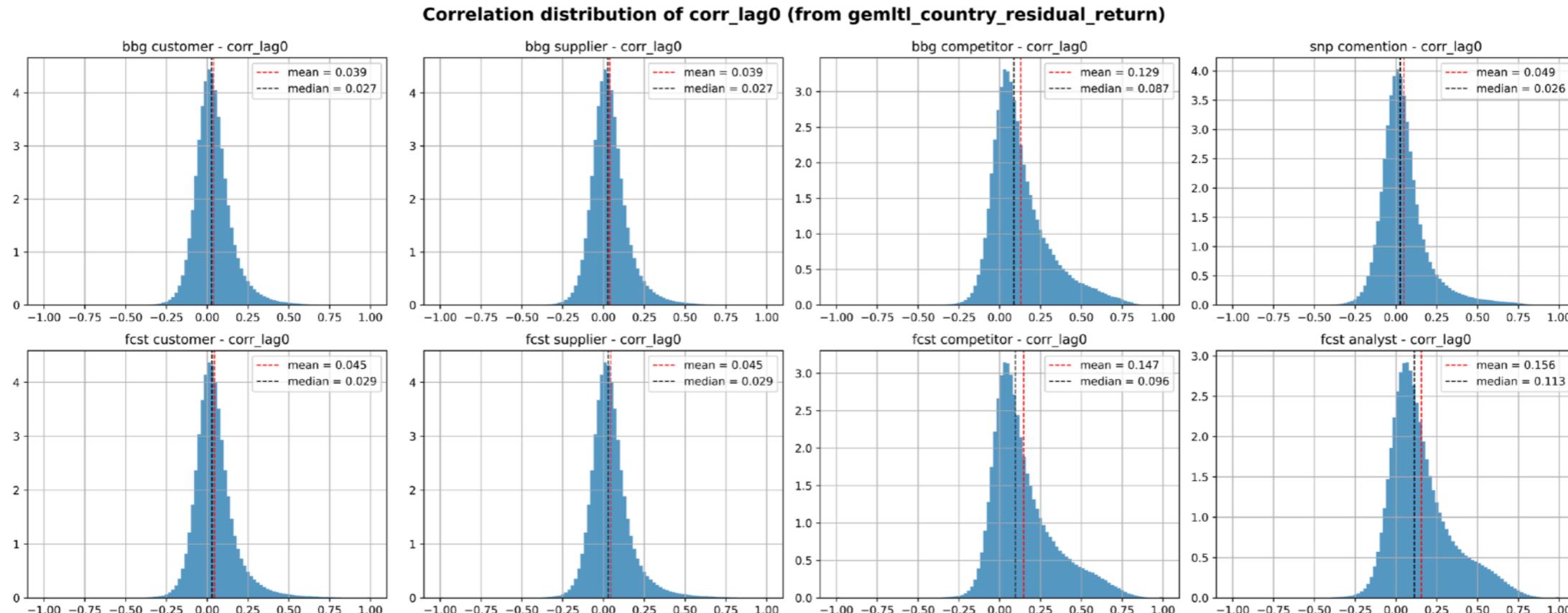
- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - How many edges removed/kept if we set `threshold=0` ?
 - How many connections (aka. node degree) an entity has before/after pruning?



- Findings
 - The synchronous correlations of raw return between fundamental peers averaged around 0.24 to 0.37.
 - The **fcst/bbg supply chain** and **snp comention network** has the lowest correlations.
 - The **fcst co-analyst network** has the highest correlations.

Robustness Analysis: Threshold v.s. Sparsity

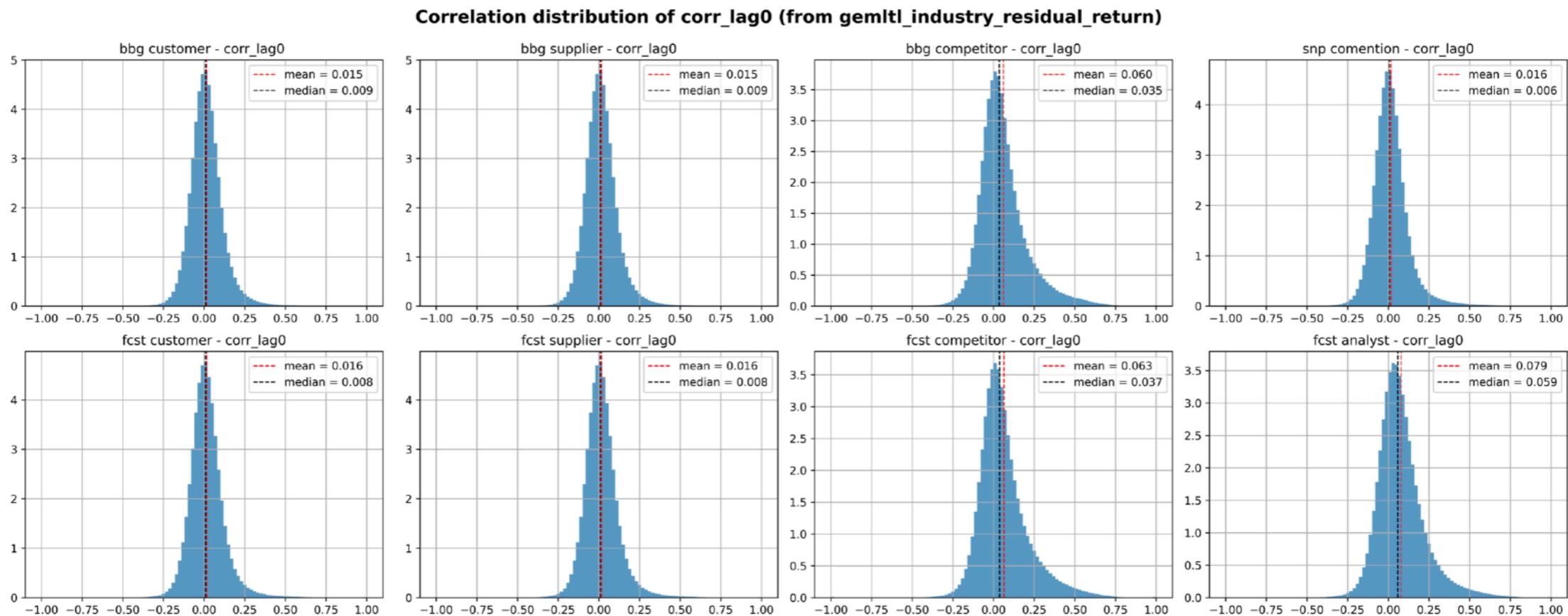
- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - How many edges removed/kept if we set `threshold=0` ?
 - How many connections (aka. node degree) an entity has before/after pruning?



- Findings
 - The synchronous correlations of raw return between fundamental peers averaged around 0.24 to 0.37.
 - The **fcst/bbg supply chain** and **snр comention network** has the lowest correlations.
 - The **fcst co-analyst network** has the highest correlations.

Robustness Analysis: Threshold v.s. Sparsity

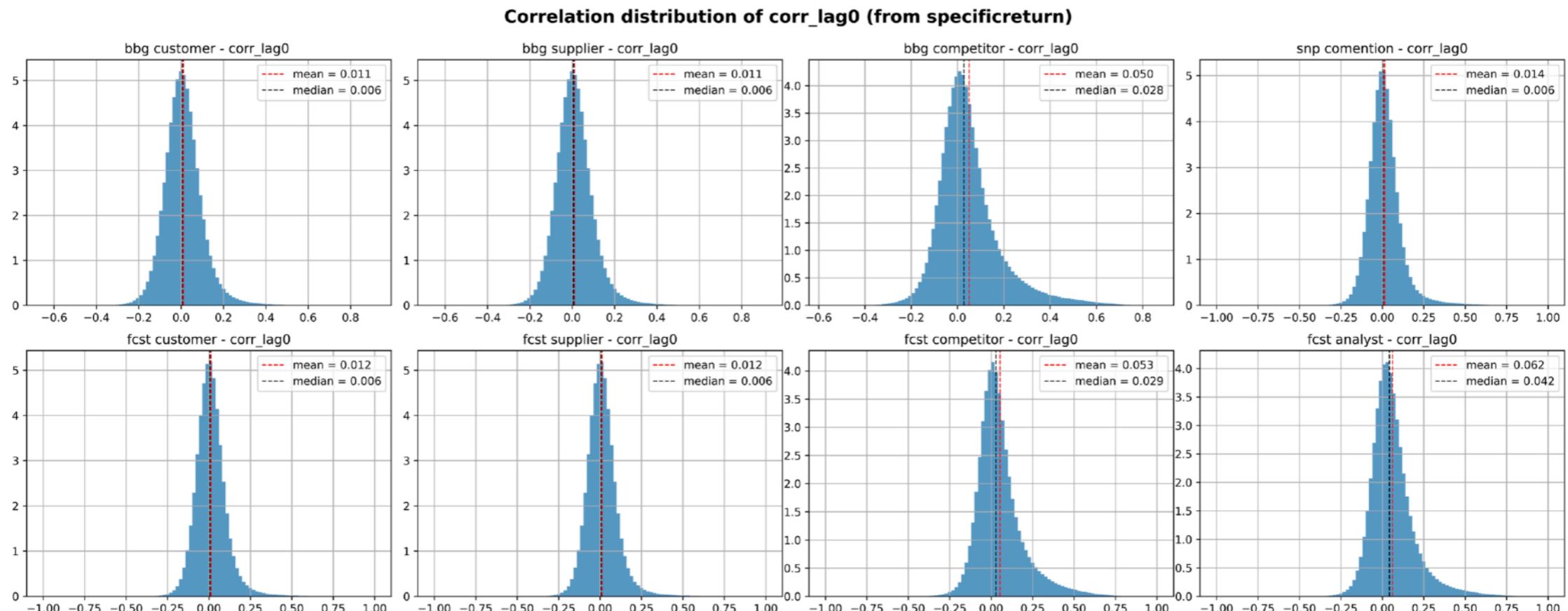
- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - How many edges removed/kept if we set `threshold=0` ?
 - How many connections (aka. node degree) an entity has before/after pruning?



- Findings
 - The synchronous correlations of raw return between fundamental peers averaged around 0.24 to 0.37.
 - The **fcst/bbg supply chain** and **snp comention network** has the lowest correlations.
 - The **fcst co-analyst network** has the highest correlations.

Robustness Analysis: Threshold v.s. Sparsity

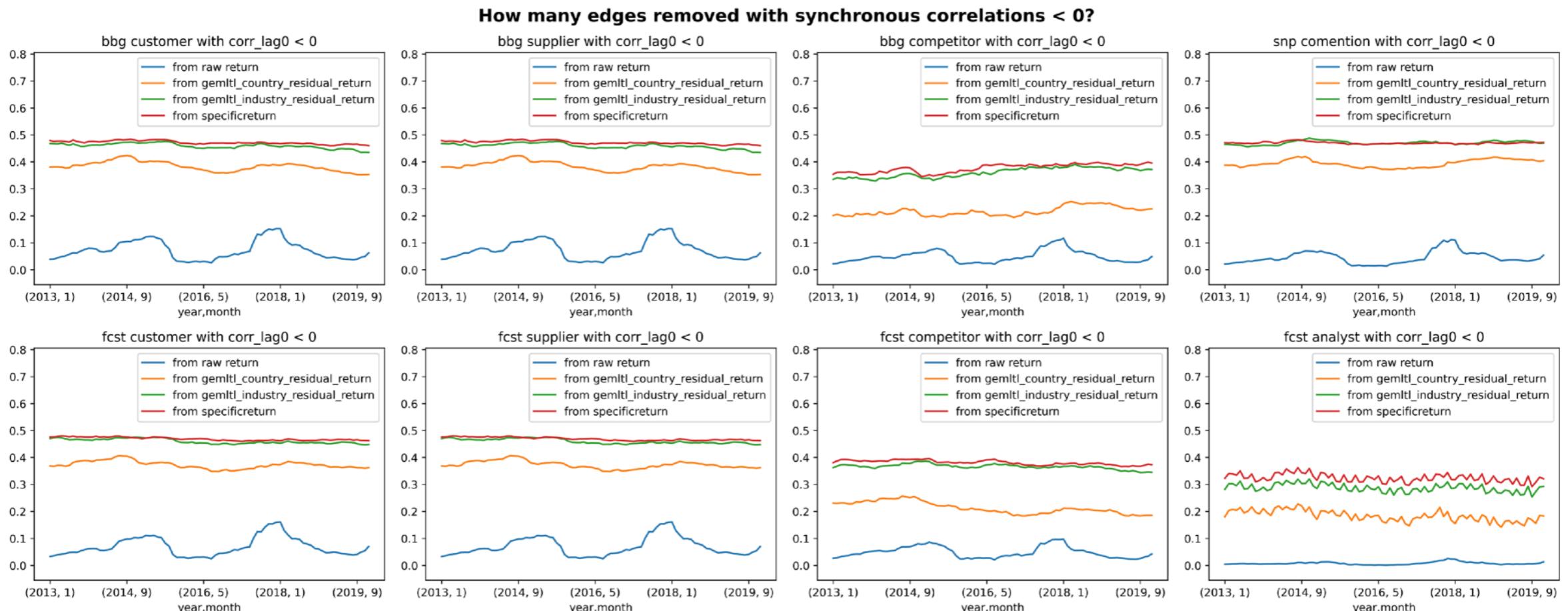
- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - How many edges removed/kept if we set `threshold=0` ?
 - How many connections (aka. node degree) an entity has before/after pruning?



- Findings
 - The synchronous correlations of raw return between fundamental peers averaged around 0.24 to 0.37.
 - The **fcst/bbg supply chain** and **snp comention network** has the lowest correlations.
 - The **fcst co-analyst network** has the highest correlations.

Robustness Analysis: Threshold v.s. Sparsity

- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - **How many edges removed/kept if we set `threshold=0` ?**
 - How many connections (aka. node degree) an entity has before/after pruning?

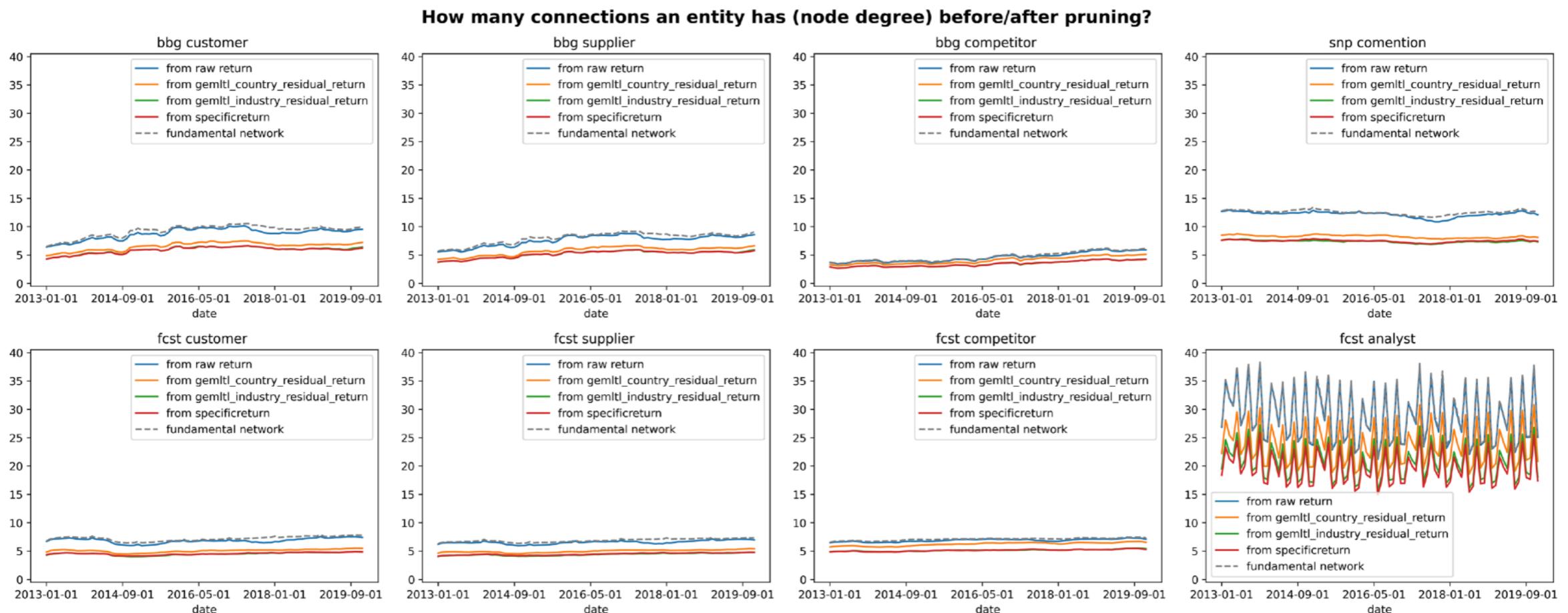


- Findings
 - How many edges removed v.s. Return types
 - For raw returns, 5% - 10% edges were removed (except for fcst analyst, which is 1%)
 - For country residual returns, 20% - 40% edges were removed
 - For industry residual returns/specifc returns, 30% - 50% eedges were removed
 - In terms of the number of edge removed,

bbg/fcst suupply chain \geq snp cometion > bbg/fcst competitor > fcst analyst

Robustness Analysis: Threshold v.s. Sparsity

- Research Questions
 - What are the correlation distributions of different fundamental networks?
 - How many edges removed/kept if we set `threshold=0` ?
 - **How many connections (aka. node degree) an entity has before/after pruning with synchronous corr < 0?**
 - How many connections (aka. node degree) an entity in different regions has?



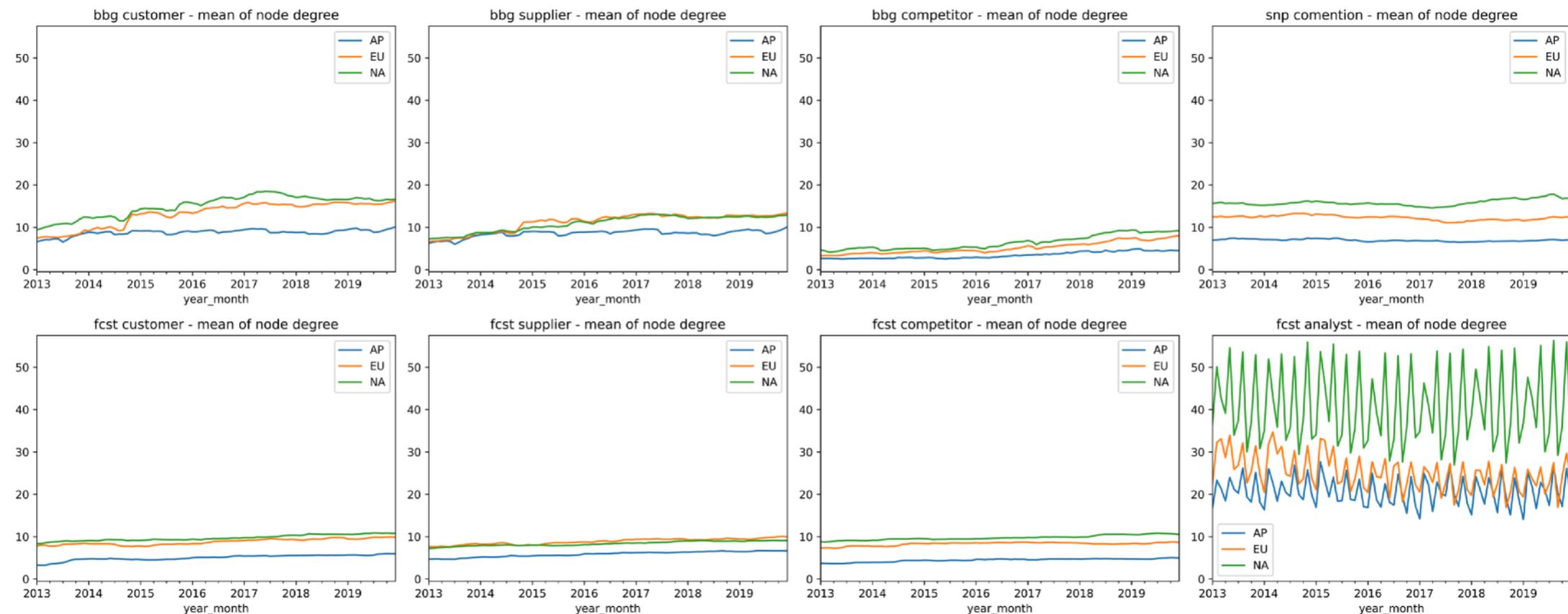
- Findings:
 - Removing edges with `corr < 0` from raw returns does not change the node degree to a large extent.
 - In terms of the average node degree:
 - fcst analyst >> snp comention > bbg/fcst supply chain > fcst competitor > bbg competitor
 - In terms of the average node degree of companies in different regions:
 - NA >= EU > AP

Robustness Analysis: Threshold v.s. Sparsity

- Research Questions

- What are the correlation distributions of different fundamental networks?
- How many edges removed/kept if we set `threshold=0` ?
- How many connections (aka. node degree) an entity has before/after pruning with synchronous corr < 0?

- **How many connections (aka. node degree) an entity in different regions has?**



- Findings:

- Removing edges with `corr < 0` from raw returns does not change the node degree to a large extent.
- In terms of the average node degree:
 - fcst analyst >>> snc comention > bbg/fcst supply chain > fcst competitor > bbg competitor
- In terms of the average node degree of companies in different regions:
 - NA >= EU > AP

Robustness Analysis: Universe Filter

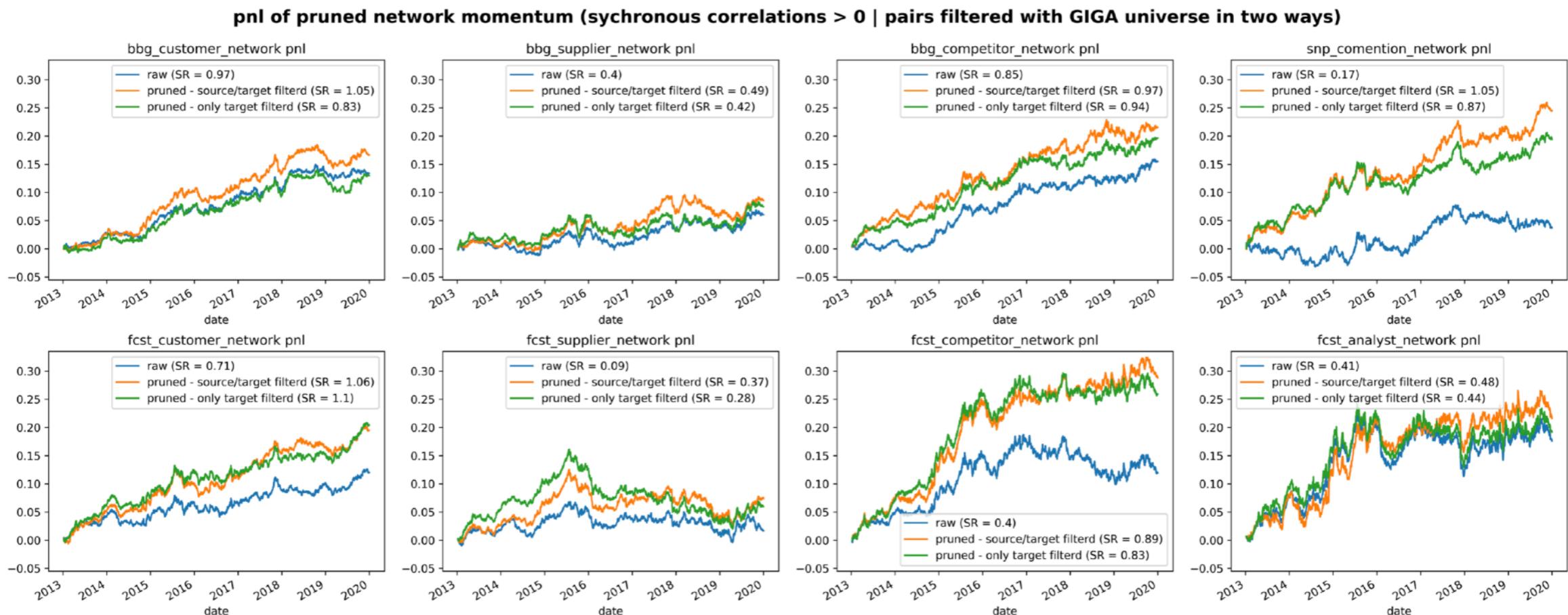
- In order to reduce computational complexity, two filters are adopted before calculating the correlations of peers.

For an edge `source -> target` appearing in a fundamental network:

- We remove it if both `source` and `target` are not in the Q-universe. (**source/target filtered**)
- We remove it if `target` is not in the Q-universe. (**only target filtered**)

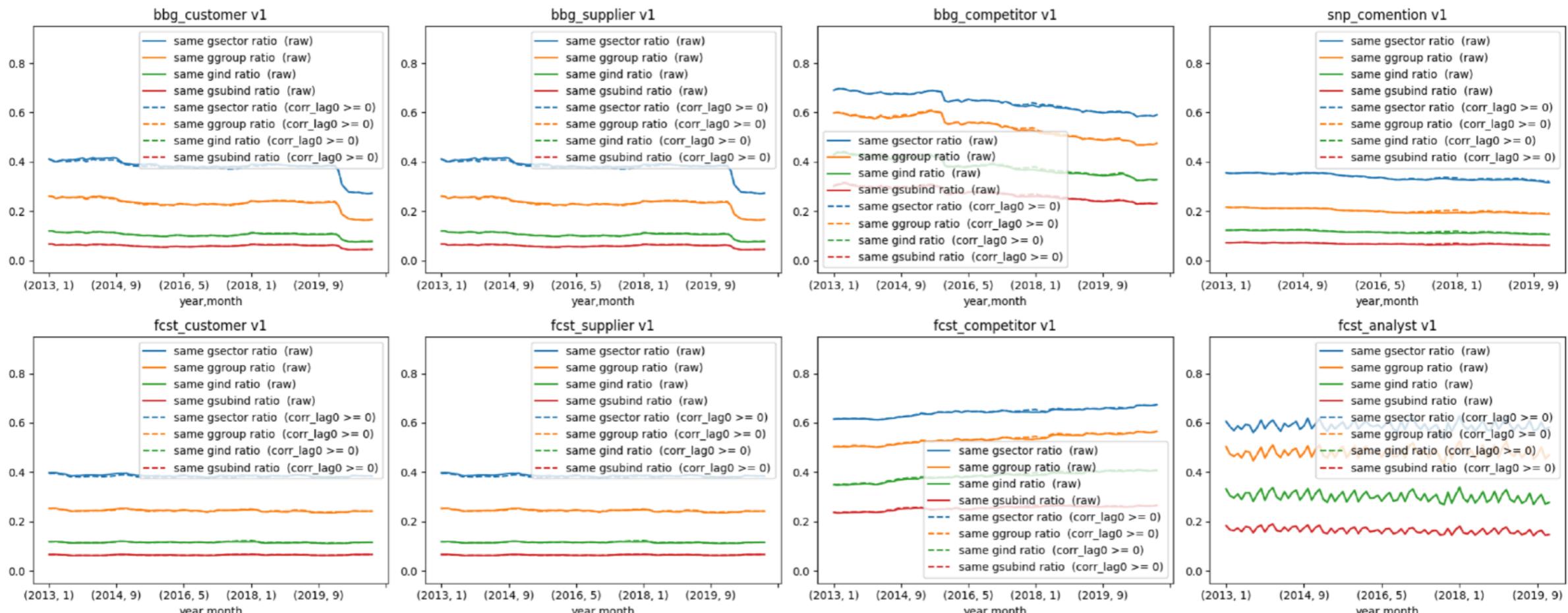
When we propagate momentum from `source` to `target`, we are not going to trade those `targets` that are not in the Q-universe.

- (**source/target filtered**) generally had a better Sharpe Ratio in the backtest (except for the fcst customer network).



Robustness Analysis: GICS

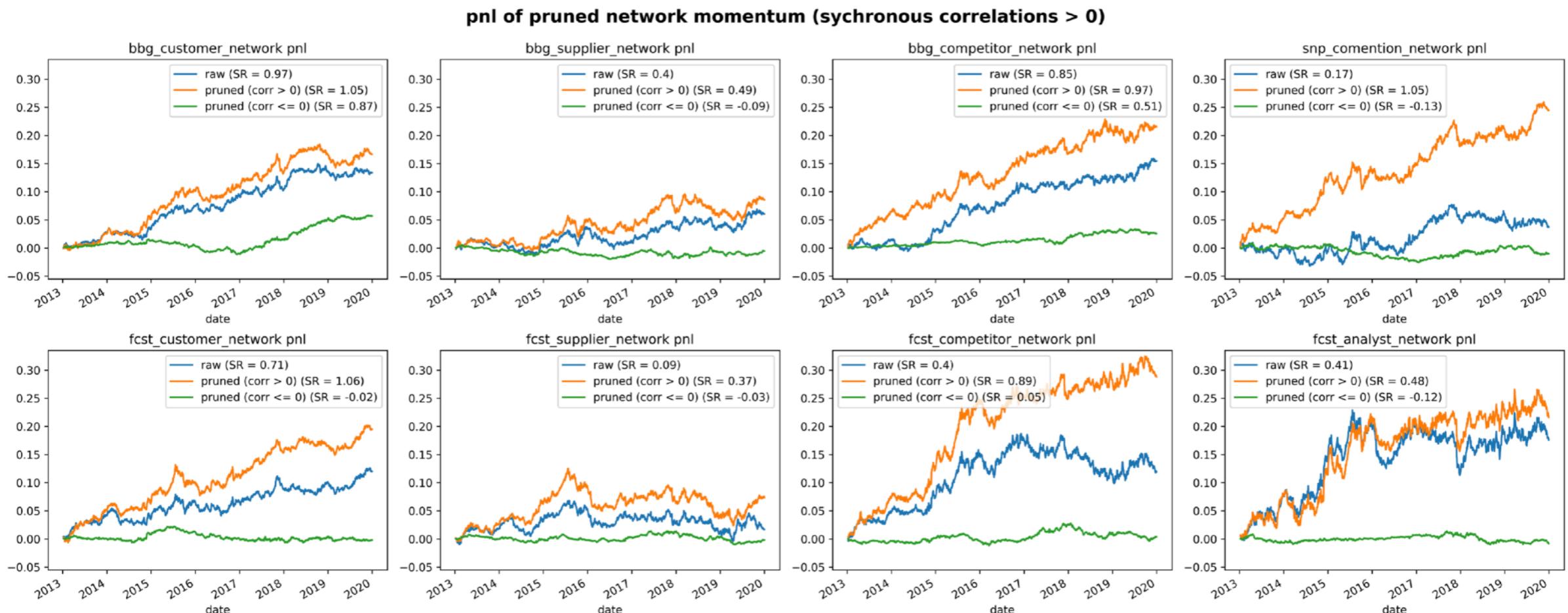
- Are the pruned networks more similar to GICS?
- **No**, the original network and pruned network have similar ratio of gics edges to non-gics edges.



Robustness Analysis: Negative Edges

- What if we only take negative edges - would it be significantly worse than raw network momentum signal?

| Yes.



Qualitative Analysis

Qualitative Analysis: Exploring Removed Edges

- The **snp comention network** shows the most significant improvement in the pruned network momentum signal compared to the s35 network momentum, SR from 0.17 to 1.31.

We are interested in **what are those edges removed** in the snp comention network?

- There are following types of edges removed

- Competitors
- Comentioned, but semantically negative
- Not very meaningful comention
- No comention found through online searches

	date	source_qid	target_qid	corr_lag0	source	target
0	2018-05-20	R96YBR-R	RMFR2L-R	-0.503623	Woodside Petroleum Ltd	TechnipFMC Plc
1	2018-05-20	RMFR2L-R	R96YBR-R	-0.503623	TechnipFMC Plc	Woodside Petroleum Ltd
2	2018-05-20	JPXNB3-R	GM702V-R	-0.494657	XL Group Ltd	AXA SA
3	2018-05-20	GM702V-R	JPXNB3-R	-0.494657	AXA SA	XL Group Ltd
4	2018-05-20	HKR6LR-R	ND2X37-R	-0.404209	Shell PLC	Daseke, Inc.
5	2018-05-20	ND2X37-R	HKR6LR-R	-0.404209	Daseke, Inc.	Shell PLC
6	2018-05-20	FT8FQW-R	DTKPJP-R	-0.383674	Verizon Communications Inc.	American Tower Corporation
7	2018-05-20	DTKPJP-R	FT8FQW-R	-0.383674	American Tower Corporation	Verizon Communications Inc.
8	2018-05-20	FMFR39-R	P7L1L7-R	-0.377780	NCR Corporation	Shyft Group, Inc.
9	2018-05-20	P7L1L7-R	FMFR39-R	-0.377780	Shyft Group, Inc.	NCR Corporation
10	2018-05-20	LSNJJXJ-R	CSFJSZ-R	-0.376745	easyJet plc	American Airlines Group, Inc.
11	2018-05-20	CSFJSZ-R	LSNJJXJ-R	-0.376745	American Airlines Group, Inc.	easyJet plc
12	2018-05-20	BS0GWP-R	M0H6M5-R	-0.375656	Tokyo Electron Ltd.	Applied Materials, Inc.
13	2018-05-20	M0H6M5-R	BS0GWP-R	-0.375656	Applied Materials, Inc.	Tokyo Electron Ltd.

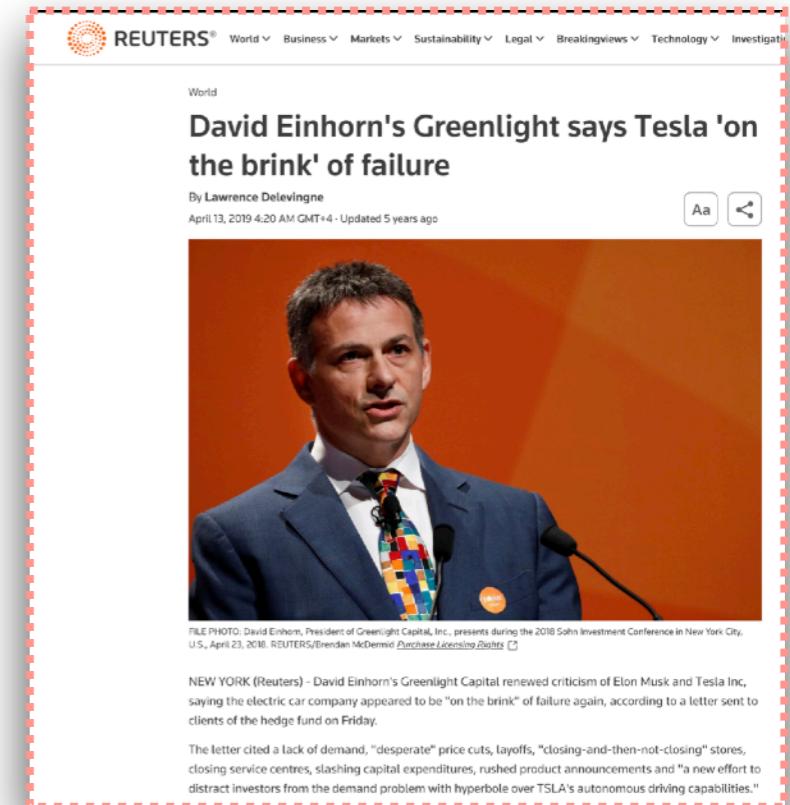
 **ChatGPT**
On May 20, 2008, AXA SA, a French multinational insurance firm, announced its acquisition of XL Group Ltd, a global insurance and reinsurance company based in Bermuda. The acquisition was a significant move by AXA to expand its presence in the property and casualty insurance market, particularly in the United States.

Qualitative Analysis: Exploring Removed Edges

- The **snp comention network** shows the most significant improvement in the pruned network momentum signal compared to the s35 network momentum, SR from 0.17 to 1.31.

We are interested in **what are those edges removed** in the snp comention network?

- There are following types of edges removed
 - Competitors
 - Comentioned, but semantically negative
 - Not very meaningful comention
 - No comention found through online searches



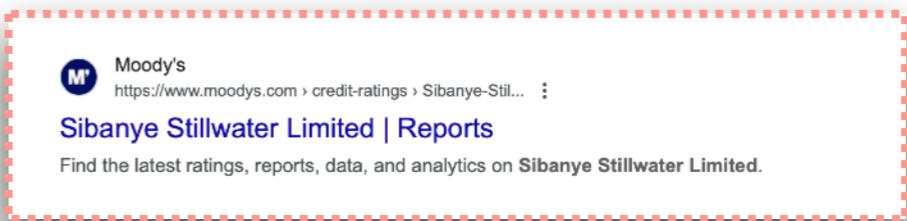
source_qid	target_qid	corr_lag0_max	corr_lag0_mean	start_date	end_date	source	target
0 MCNYYL-R	W5QDQD-R	-0.084150	-0.170365	2019-01-01	2019-12-31	Amazon.com, Inc.	Japan Metropolitan Fund Investment Corporation
1 T0SQZH-R	D68983-R	-0.059092	-0.140310	2019-01-01	2019-12-31	General Motors Company	CMS Energy Corporation
2 T0SQZH-R	FYT21Z-R	-0.034038	-0.132558	2019-01-01	2019-12-31	General Motors Company	B2Gold Corp.
3 CYB6Y5-R	RBG33K-R	-0.024658	-0.126435	2019-01-01	2019-12-31	Coca-Cola Company	Coca-Cola Amatil Limited
4 NXHLP0-R	DQ36JT-R	-0.025608	-0.123531	2019-01-01	2019-12-31	Dominion Energy Inc	CIE Automotive, S.A.
5 Q3VCTK-R	JDQ0K1-R	-0.045289	-0.121036	2019-01-01	2019-12-31	Computershare Limited	AutoZone, Inc.
6 LSK8PG-R	HQ4DBK-R	-0.055186	-0.110936	2019-01-01	2019-12-31	Lawson, Inc.	Oracle Corporation
7 Q2YN1N-R	C4BHC3-R	-0.000788	-0.105029	2019-01-01	2019-12-31	Tesla, Inc.	Greenlight Capital Re, Ltd. Class A
8 RK8Z6D-R	LKVSR-Z-R	-0.017686	-0.101580	2019-01-01	2019-12-31	Entergy Corporation	McDermott International, Inc.

Qualitative Analysis: Exploring Removed Edges

- The **snp comention network** shows the most significant improvement in the pruned network momentum signal compared to the s35 network momentum, SR from 0.17 to 1.31.

We are interested in **what are those edges removed** in the.snp comention network?

- There are following types of edges removed
 - Competitors
 - Comentioned, but semantically negative
 - Not very meaningful comention
 - No comention found through online searches



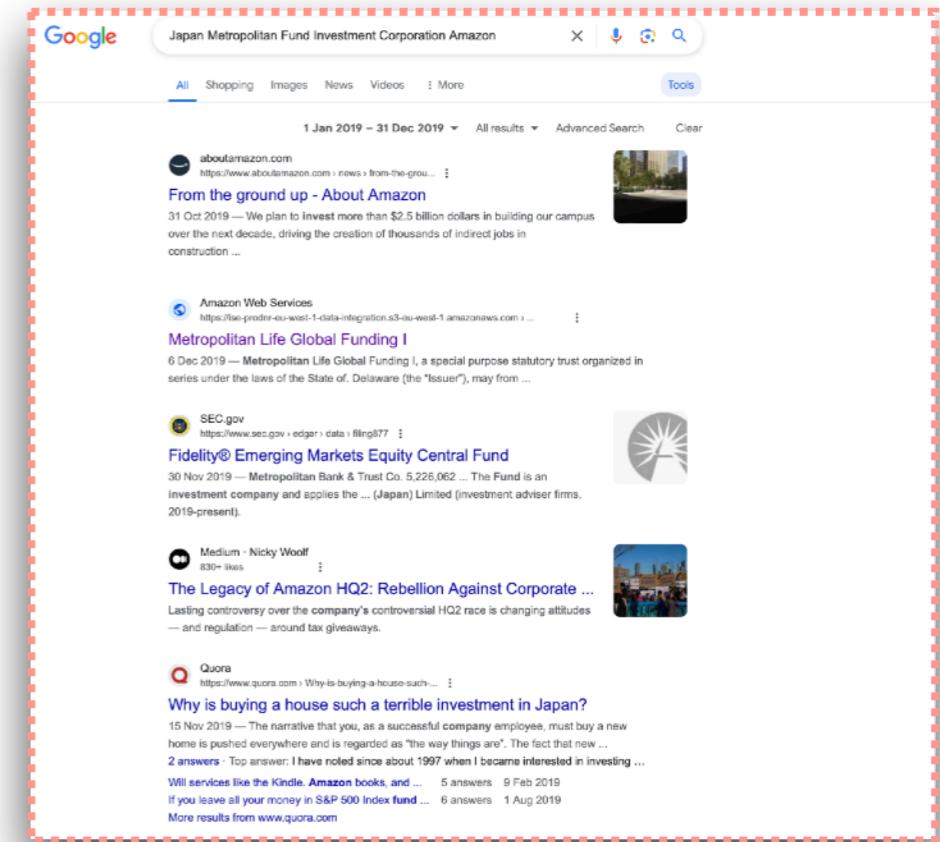
25	J2N4RC-R	N46BJL-R	-0.013585	-0.047632	2019-01-01	2019-12-31	National Grid plc	Alarm.com Holdings, Inc.
26	M61G3T-R	RXHN9P-R	-0.010784	-0.044213	2019-01-01	2019-12-31	ANA Holdings Inc.	Boeing Company
27	MDYC2K-R	G50SJ-R	-0.003618	-0.044143	2019-01-01	2019-12-31	Moody's Corporation	Sibanye Stillwater Limited
28	CVGGV1-R	N27LVR-R	-0.000310	-0.035253	2019-01-01	2019-12-31	Peptidream Inc.	Biohaven Pharmaceutical Holding Company Ltd.

Qualitative Analysis: Exploring Removed Edges

- The **snp comention network** shows the most significant improvement in the pruned network momentum signal compared to the s35 network momentum, SR from 0.17 to 1.31.

We are interested in **what are those edges removed** in the snp comention network?

- There are following types of edges removed
 - Competitors
 - Comentioned, but semantically negative
 - Not very meaningful comention
 - No comention found through online searches



source_qid	target_qid	corr_lag0_max	corr_lag0_mean	start_date	end_date	source	target	
0	MCNYYL-R	W5QDQD-R	-0.084150	-0.170365	2019-01-01	2019-12-31	Amazon.com, Inc.	Japan Metropolitan Fund Investment Corporation
1	T0SQZH-R	D68983-R	-0.059092	-0.140310	2019-01-01	2019-12-31	General Motors Company	CMS Energy Corporation
2	T0SQZH-R	FYT21Z-R	-0.034038	-0.132558	2019-01-01	2019-12-31	General Motors Company	B2Gold Corp.
3	CYB6Y5-R	RBG33K-R	-0.024658	-0.126435	2019-01-01	2019-12-31	Coca-Cola Company	Coca-Cola Amatil Limited
4	NXHLP0-R	DQ36JT-R	-0.025608	-0.123531	2019-01-01	2019-12-31	Dominion Energy Inc	CIE Automotive, S.A.
5	Q3VCTK-R	JDQ0K1-R	-0.045289	-0.121036	2019-01-01	2019-12-31	Computershare Limited	AutoZone, Inc.
6	LSK8PG-R	HQ4DBK-R	-0.055186	-0.110936	2019-01-01	2019-12-31	Lawson, Inc.	Oracle Corporation
7	Q2YN1N-R	C4BHC3-R	-0.000788	-0.105029	2019-01-01	2019-12-31	Tesla, Inc.	Greenlight Capital Re, Ltd. Class A
8	DK07ZD-R	UWV7ZD-R	-0.017000	-0.101500	2019-01-01	2019-12-31	ExxonMobil	M&P Global

Causal graph based on s35

