Refining the Valuation of Seasoned Equity Offerings by Creating a More Accurate Replicating Portfolio Using Machine Learning Techniques

Student: Andrey Aksyutkin BAE’24

Adviser: Olga Kuzmina

Motivation

- Why is it important to know how SEO affects company returns?
  - "SEO are important for firms that want to finance growth opportunities" (Veld et al., 2018)
- Zero or negative effects of SEO on the long-term stock performance
  - Market timing theory. Managers try to time the market and issue equity when valuations are highest. (Baker and Wurgler, 2002)
  - "The decreased leverage associated with an equity issue lowers the sensitivity of the stock price to inflation shocks, and the extra shares outstanding make the stock more liquid which lowers expected returns" (Eckbo et al., 2000)
  - Managers, who are better informed than shareholders, try to sell overpriced equity (Myers and Majluf, 1984)
- Synthetic controls are used to produce controls with a small number of existing elements (Abadie, 2015)
- **My research question: evaluating the effect of SEO on stock returns using synthetic control on the newest available data.**

Literature review

- The effects of SEO. Negative abnormal returns (mostly, significantly different from 0)
  - -4.8% p(t) = 0.000 (Eckbo, 2000) by equity market value
  - -4.6% (Ritter, 2003) by closest market capitalization
  - -0.9% (Ritter, 2003) by the same size decile, closest B/M ratio
  - -18% p(t) = 0.000 (Huang, 2014) by the same SIC code, same size-decile, same B/M quantile, closest 6-month stock return before SEO
- Applying machine learning methods to evaluate various corporate events and beyond
  - Building synthetic control for M&A deals evaluation (Chava and Reguly, 2024 wp)
  - Machine learning can predict company takeovers (M&A) (Geha, 2021)
  - AI-based hedgers overperform direct replication of derivatives in the incomplete market setting and transaction costs (Cannelli et al., 2023)
- Synthetic controls in Econometrics for countries and states
  - Synthetic Basque Country using a combination of other Spanish regions (Abadie, 2003)
  - Synthetic per-capita cigarettes sales to compare with real CA (Abadie, 2010)
  - Synthetic West Germany before Reunification using 16 other countries (Abadie 2015)

Data

- Source: Refinitiv Workspace
- Features: exchange ticker, GICS Industry name, monthly returns 2016-2023, market capitalization on the end of each year, B/M ratio, number of issued common stocks from annual financial reports.
- All publicly traded (NYSE or Nasdaq) companies with capitalization >1Bln$
- Number of firms which did not make SEO in 2021-2023 with complete data: 1385
- Number of firms which made SEO in 2021: 109
- Number of firms which made SEO in 2022: 78
- Number of firms which made SEO in 2023: 76
- 74 GICS Industry presented

Methodology (Lasso regression)

Synthetic control portfolio with Lasso selected weights (inspired by Abadie, 2015)

SEO abnormal return

$$\widehat{R}^{abnormal}_{SEO,t} = R_{SEO,t} - \sum_{k=1}^{n} w_k^* \cdot R_{non-SEO,t}$$

$t \in T_{test}$

For each SEO-made company:

1. Making a dataset of only non-SEO companies with the same industry as SEO has
2. Finding best regularization coefficient using validation set
3. Minimizing (force 0 intercept) optimization formula using 5 years pre-SEO monthly returns from the dataset
4. Memorizing best weights (w*) from the minimized Lasso optimization formula
5. Forecasting SEO returns multiplying w* by returns of non-SEO dataset

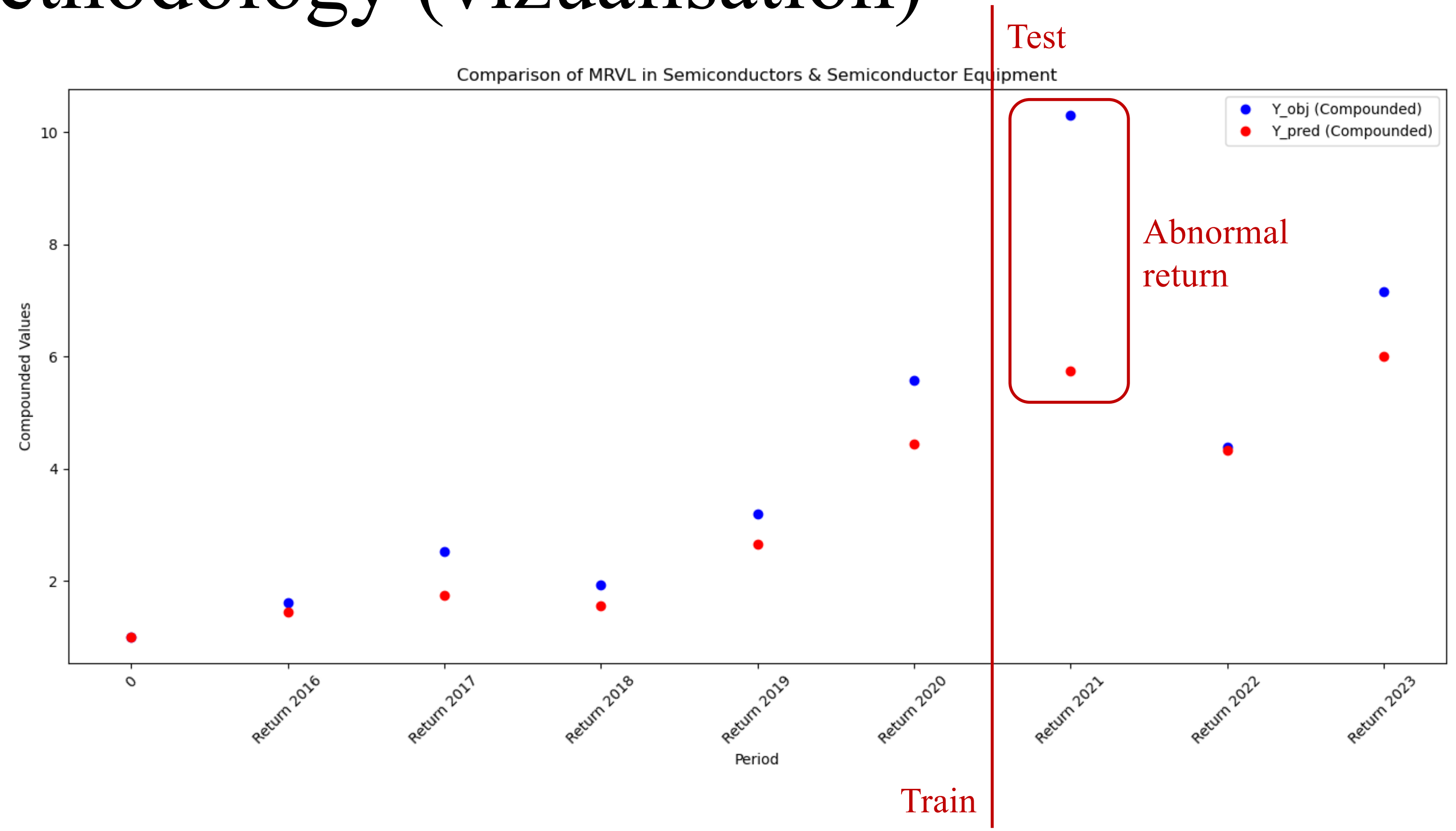
Optimization formula

Optimal weights $\vec{w}^*$

$$\min_{w \in R^{60}_{\geq 0}} \left\{ \sum_{t \in T_{train}} \left(R_{SEO,t} - \sum_{i=1}^{n} w_i \cdot R_{non-SEO,t}\right)^2 + \lambda \cdot \sum_{j=1}^{60} |w_j| \right\}$$

$t \in T_{train}$

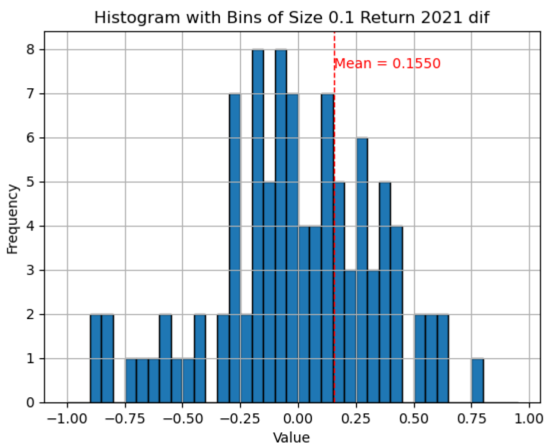
Methodology (vizualisation)



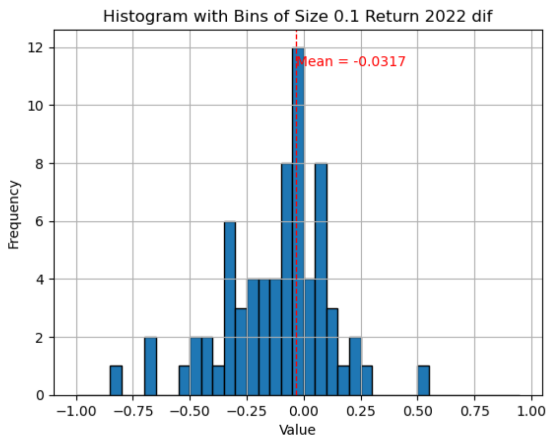
Results

- SEO aggregated mean abnormal return over synthetic portfolio was 15%* in 2021, -3% in 2022, 30%** in 2023 (abnormal return on common stock for SEO made respectively in 2021, 2022, 2023)

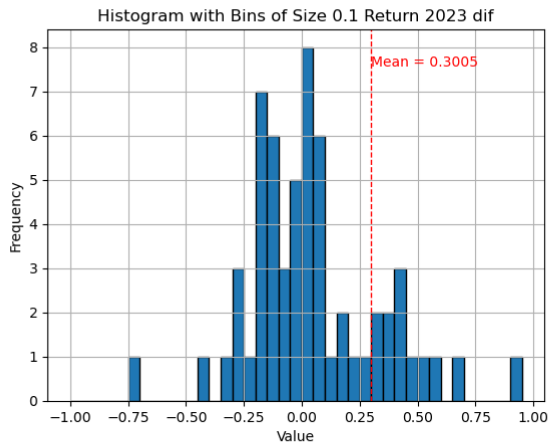
Distribution of abnormal SEO returns for each year (in pp annualized)



t-statistic = 1.903 p-value = 0.0598



t-statistic = -0.557 p-value = 0.5788

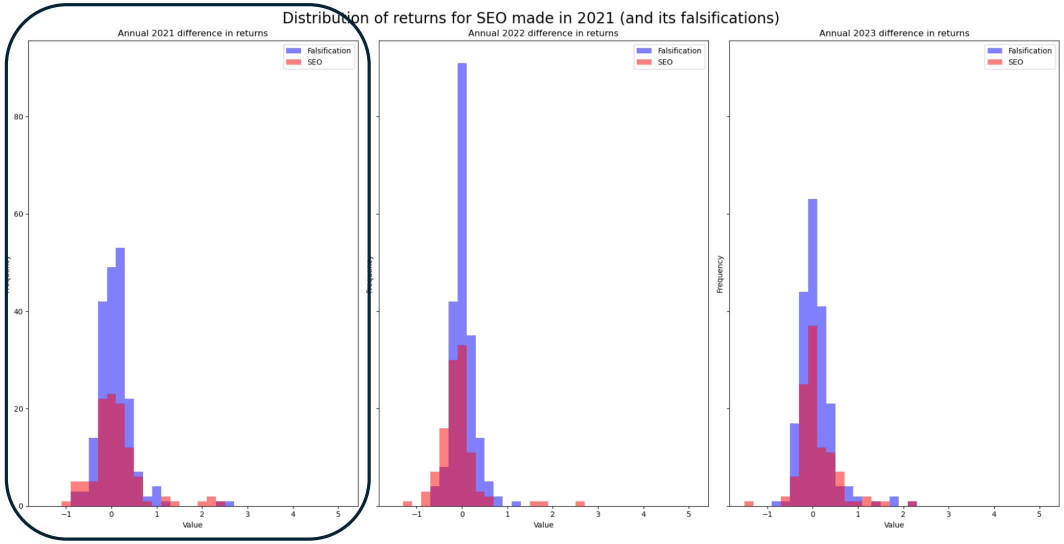


t-statistic = 2.588 p-value = 0.0119

Aggregated from the plots above: mean = 0.13539720978564185 t- statistic=2.707, p-value=0.0072, df=252

Falsification

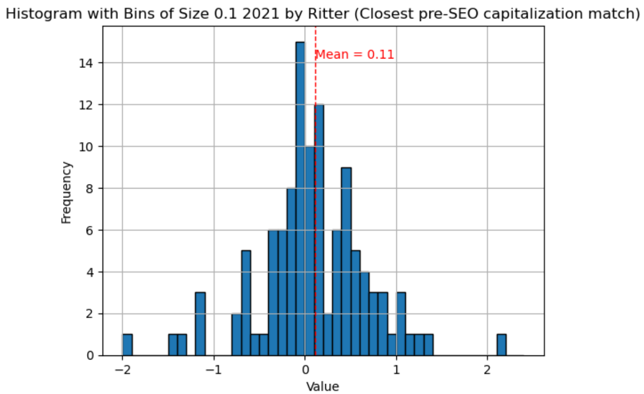
- How confident are these results? Alternative randomized inference from Young (2019)
  - Assign treatment (SEO) randomly through all companies which did not make SEO in years 2021-2023
  - Apply the same optimization as before
  - Check how predictions were made in the following SEO year (abnormal returns differ from 0).



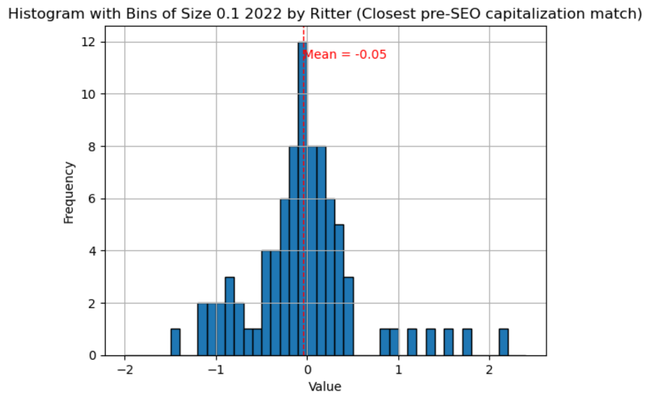
2021 Falsification mean = 0.0921, t-statistic = 3.250, p-value = 0.0014
2021 Real mean = 0.1550, t-statistic = 1.903 p-value = 0.0598

Comparison to a previous method

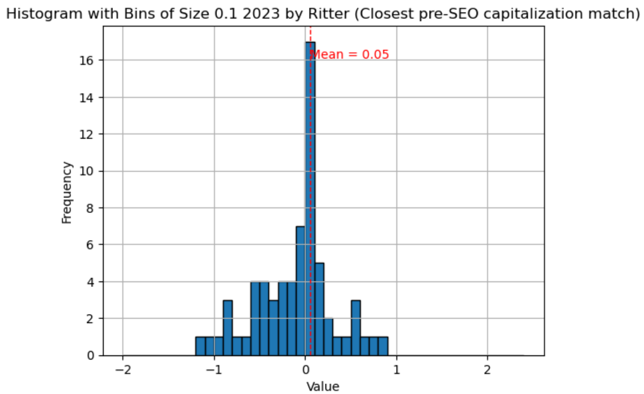
Matching by the closest capitalization year before SEO. Ritter (2003)



mean = 0.11
t-statistic = 1.331
p-value = 0.1859

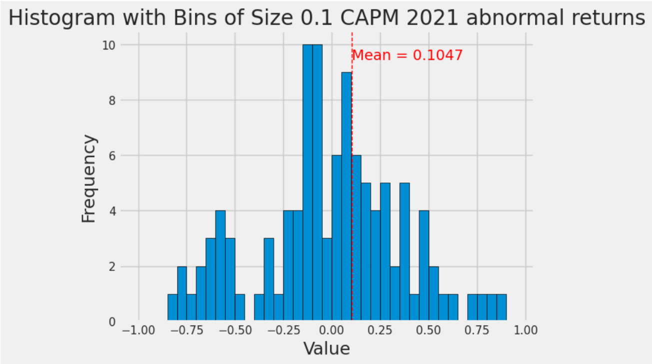


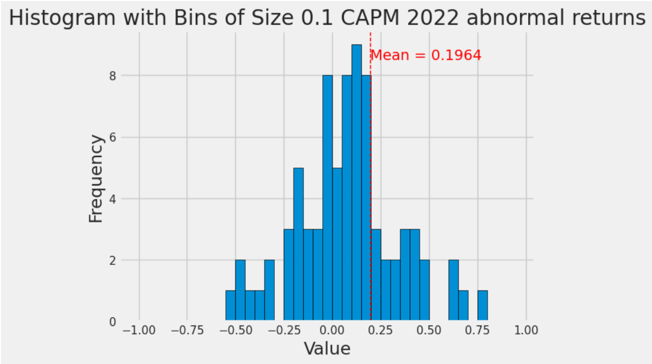
mean = -0.05
t-statistic = -0.706
p-value = 0.4820

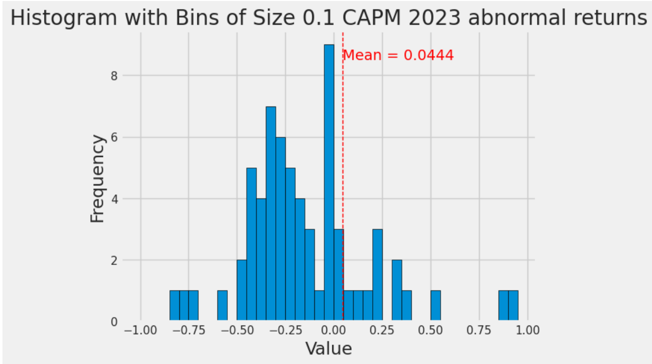


mean = 0.05
t-statistic = 0.508
p-value = 0.6132

Comparison to CAPM







mean = 0.1047
t-statistic = 1.469
p-value = 0.1445

mean = 0.1964
t-statistic = 3.512
p-value = 0.0007

mean = 0.0444
t-statistic = 0.481
p-value = 0.6323

Conclusion

- SEO does not have a significant effect (on 1% level) on the stock returns 1 year after issuance
  - Found using novel synthetic control method
  - Alternative inference procedure confirms this result
  - In line with the results of the previous method on the same data (Ritter, 2003)
  - Partly in line with the results of CAPM on the same data
  - Not in line with the effects in the previous literature on the previous data
- Although SEO effects are not significantly different from zero, we can try to speculate about the signs: positive in 2021 and 2023, negative in 2022
  - Crypto mining and AI-related companies have the momentum (more on the next slide)
- Future research is required to understand why the results may be different in 2021-2023.

Q&A

- Novel result? (2 comment)

Applied new approach in making synthetic control for SEO.

- Do weights up to 1? (4 comment)

No, because they do not have to. Weights may be any positive numbers.

- Why the sign changes? (7 comment)

Abnormal returns are not significantly different from zero so we can only speculate about the signs.
In 2020 and 2021 all three highest abnormal returns had “Software” GICS Industry companies. They are all crypto mining companies. For them it was essential to raise new capital through SEO to finance increasing operational expenses. On the other hand, stock prices of these companies and full returns grew up very quickly driven by incredibly fast-growing crypto-assets prices. Other “Software” companies were not connected directly to crypto, hence they could not grow so fast. This effect was captured by abnormal returns. It would not exist if GICS Industry counted crypto projects as a separate industry.

- Is the model wrong? (6, 8 comments)

Mean average errors on the training datasets for all years (except 2020, read the previous answer) have mean close to zero (less than 10%) while on average full returns equal 31.5% annually. In-month errors on the training data compensate each other, mean average error is less than 2pp. On the training period falsification and real distributions are close and do not have outliers.

Appendix

Study	Horizon, weighting[b]	Sample size	Sample period	Mean buy-and-hold return		Annualized difference
				SEOs	Matching	
USA data						
Mitchell & Stafford[c]	3 yr (EW)	4439	1961–1993	34.8%	45.0%	–2.7%
Eckbo, Masulis & Norli[d]	5 yr (EW)	3315	1964–1995	44.3%	67.5%	–4.8%
Jegadeesh[e]	5 yr (EW)	2992	1970–1993	59.4%	93.6%	–4.9%
Spiess & Affleck-Graves[f]	5 yr (EW)	1247	1975–1989	55.7%	98.1%	–6.1%
Brav, Geczy & Gompers[g]	5 yr (EW)	3775	1975–1992	57.6%	83.9%	–3.9%
Mitchell & Stafford[c]	3 yr (VW)	4439	1961–1993	41.1%	45.3%	–1.1%
Eckbo, Masulis & Norli[d]	5 yr (VW)	3315	1964–1995	51.6%	62.2%	–2.2%
Brav, Geczy & Gompers[g]	5 yr (VW)	3775	1975–1992	72.5%	97.5%	–3.4%

Figure 2: Previous aggregated results by Ritter (2003)

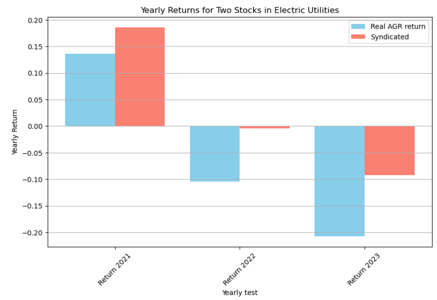
Appendix

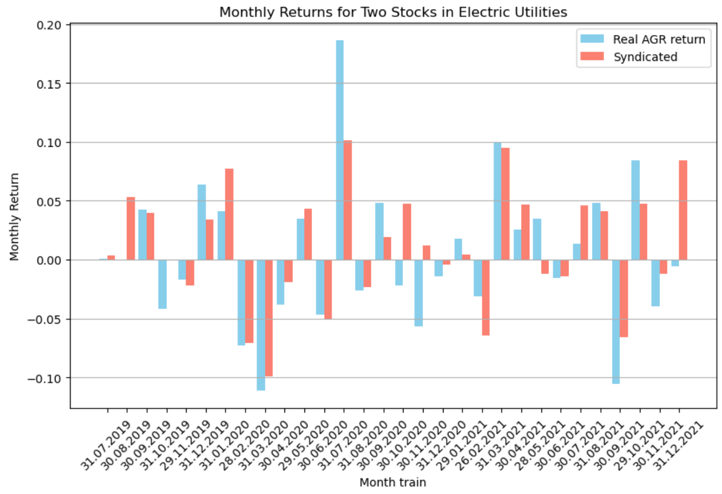
Refinitiv Workspace Data sample

016	29.07.2016	31.08.2016	30.09.2016	...	GICS Industry Name	Exchange Ticker	Difference in 4 years	Difference 2021-2020	Difference 2022-2021	Difference 2023-2022	Return 2020	Return 2021	Return 2022	Return 2023
029	-0.005563	0.113292	-0.048718	...	Mortgage Real Estate Investment Trusts (REITs)	ABR	0.530309	0.228777	0.17751	0.057649	0.117601	0.394183	-0.20693	0.30169
237	0.05141	-0.053825	0.041788	...	Retail REITs	ADC	0.674723	0.187663	0.264965	0.114734	0.021652	0.112609	0.034807	-0.070914
801	0.126942	-0.013331	0.024642	...	Semiconductors & Semiconductor Equipment	ADI	0.343118	0.421792	-0.030523	-0.025593	0.252636	0.209752	-0.049222	0.233547
767	0.319328	-0.272611	0.269702	...	Biotechnology	ADMA	1.154974	0.86662	0.132795	0.019142	-0.504447	-0.276923	1.751773	0.164948
579	-0.019974	-0.078423	0.007524	...	Electric Utilities	AGR	0.252031	0.251404	0.000145	0.000356	-0.071828	0.136403	-0.104832	-0.207511
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
813	-0.046967	0.035557	0.005919	...	Commercial Services & Supplies	VSEC	0.425316	0.151209	0.007068	0.229414	0.038687	0.595446	-0.223486	0.388626
552	-0.043893	0.225549	0.032826	...	Oil, Gas & Consumable Fuels	VTLE	1.946179	0.420489	-0.018296	1.112712	-0.643116	2.052284	-0.144853	-0.115325
462	0.630486	-0.291465	-0.033316	...	Metals & Mining	X	0.248157	0.220059	0.010608	0.012292	0.557146	0.42443	0.061217	0.956975
586	0.127951	0.208643	0.014104	...	Ground Transportation	XPO	0.137255	0.127451	0.0	0.008696	0.481481	0.1165	-0.276352	1.63112
342	-0.156742	-0.061546	0.111111	...	Energy Equipment & Services	XPRO	1.97343	1.876591	0.009236	0.024205	-0.438525	-0.127129	0.263415	-0.121897

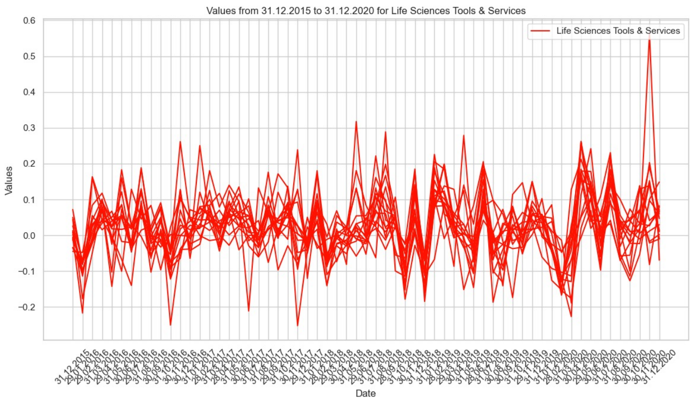
Result for each SEO-made stock

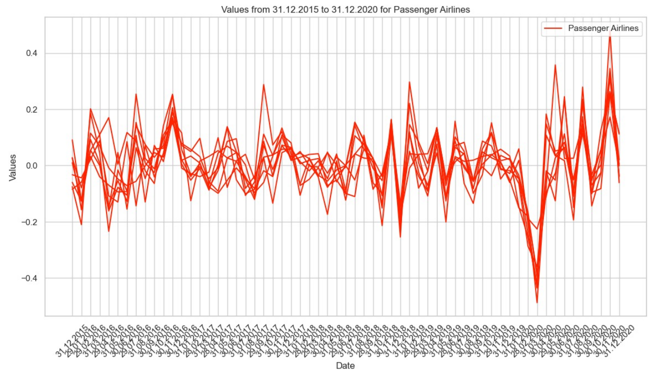
```
Mean Squared Error: 0.8998319783356502
Number of used stocks: 6
ARWR with weight: 0.1819394761303203
BMRN with weight: 0.9584714386133741
BPMC with weight: 0.25237618576004806
CPRX with weight: 0.21162185316681736
IRWD with weight: 0.03224016956597153
TGTX with weight: 0.01015219013861003
```

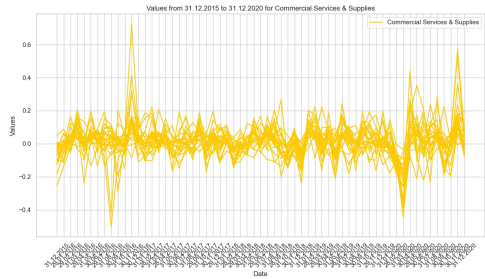


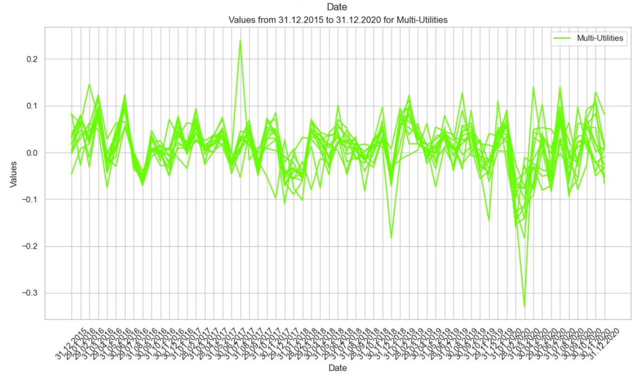


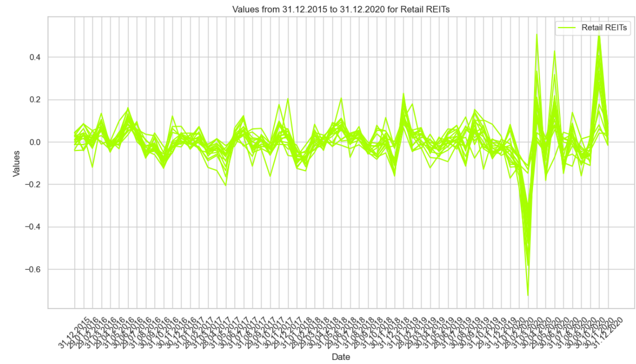
Appendix

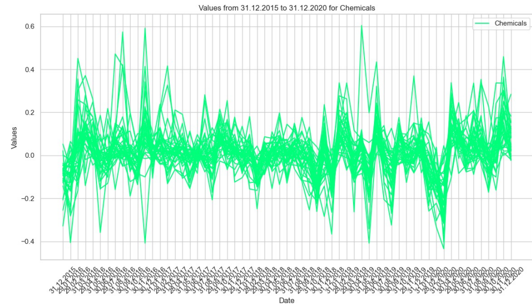

Values from 31.12.2015 to 31.12.2020 for Life Sciences Tools & Services
Life Sciences Tools & Services
Values
Date


Values from 31.12.2015 to 31.12.2020 for Passenger Airlines
Passenger Airlines
Values
Date


Values from 31.12.2015 to 31.12.2020 for Commercial Services & Supplies
Commercial Services & Supplies
Values
Date


Values from 31.12.2015 to 31.12.2020 for Multi-Utilities
Multi-Utilities
Values
Date


Values from 31.12.2015 to 31.12.2020 for Retail REITs
Retail REITs
Values
Date

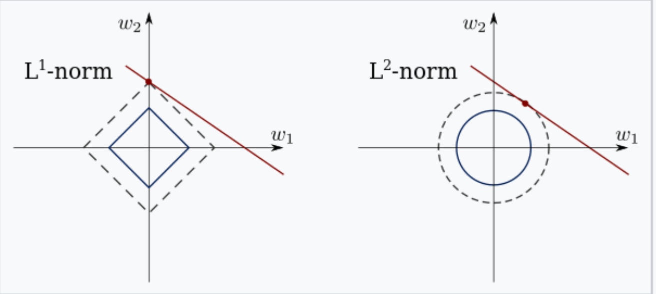

Values from 31.12.2015 to 31.12.2020 for Chemicals
Chemicals
Values
Date

Appendix

Least squares [edit]

Consider a sample consisting of *N* cases, each of which consists of *p* covariates and a single outcome. Let y_i be the outcome and $x_i := (x_1, x_2, \ldots, x_p)_i^T$ be the covariate vector for the *i* th case. Then the objective of lasso is to solve

$$\min_{\beta_0, \beta} \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - x_i^T \beta \right)^2 \right\} \text{ subject to } \sum_{j=1}^{p} |\beta_j| \leq t.$$



Forms of the constraint regions for lasso and ridge regression.

Source: Wikipedia

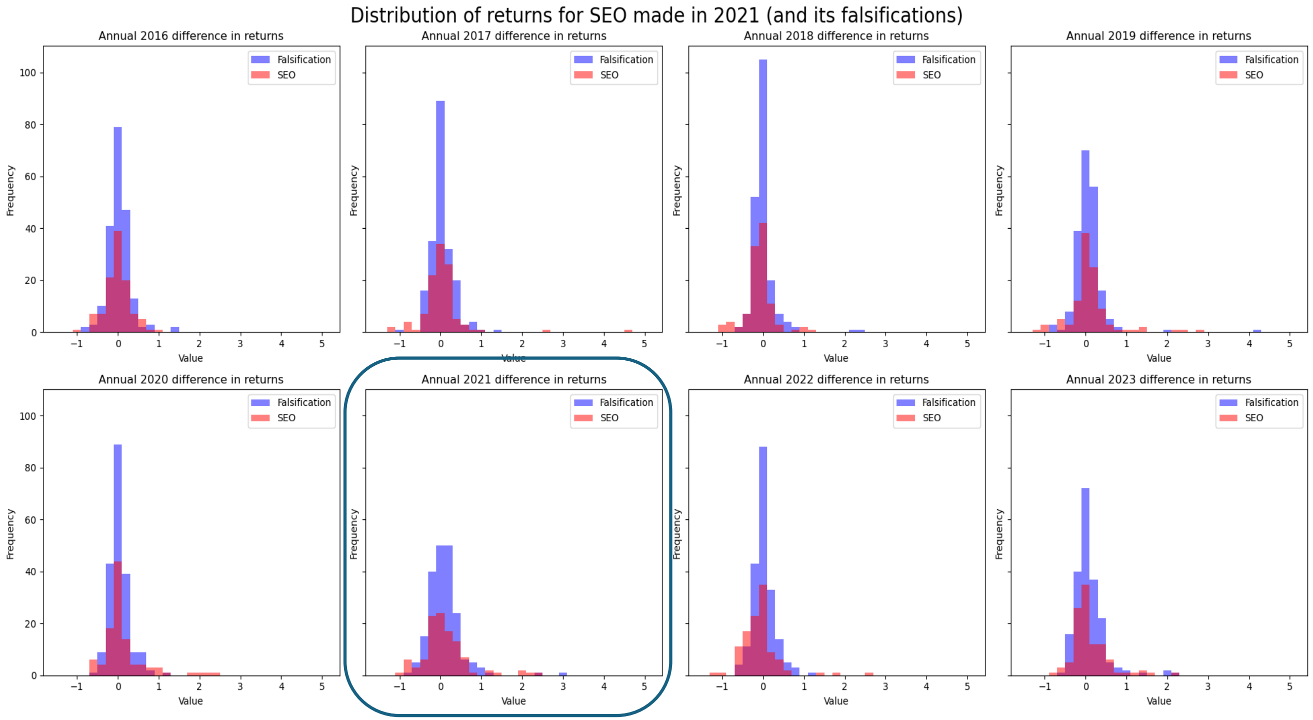
in the so-called Lagrangian form

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}$$

where the exact relationship between t and λ is data dependent.

Appendix

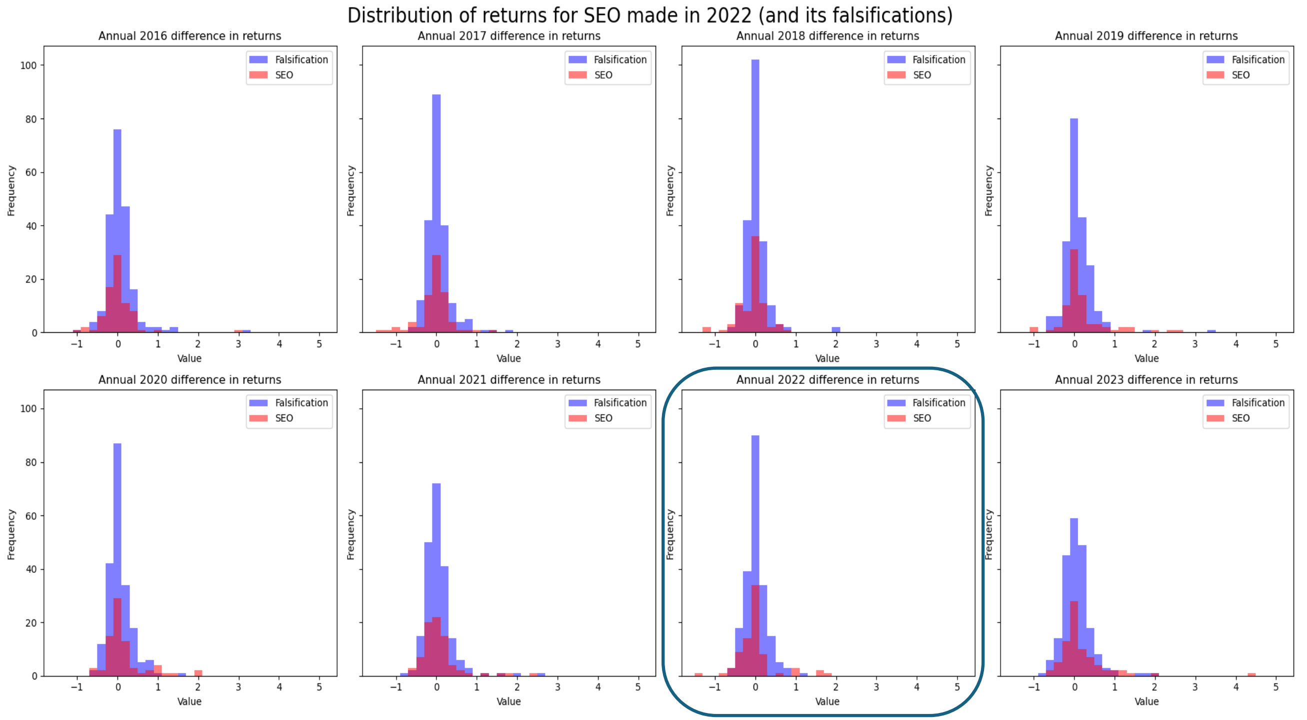
Falsification comparison (2021)



mean = 0.1550, t-statistic = 1.903 p-value = 0.0598

Appendix

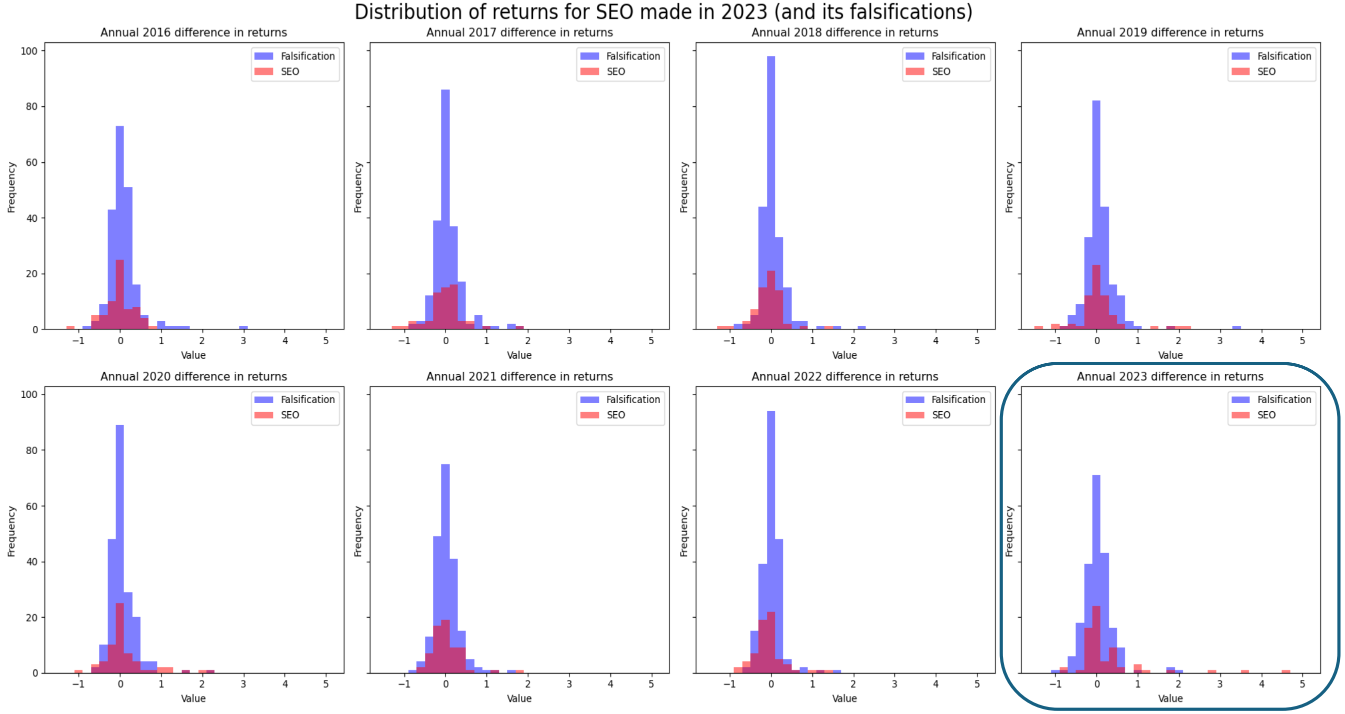
Falsification comparison (2022)



mean = -0.0281 t-statistic = -0.284 p-value = 0.7909

Appendix

Falsification comparison (2023)

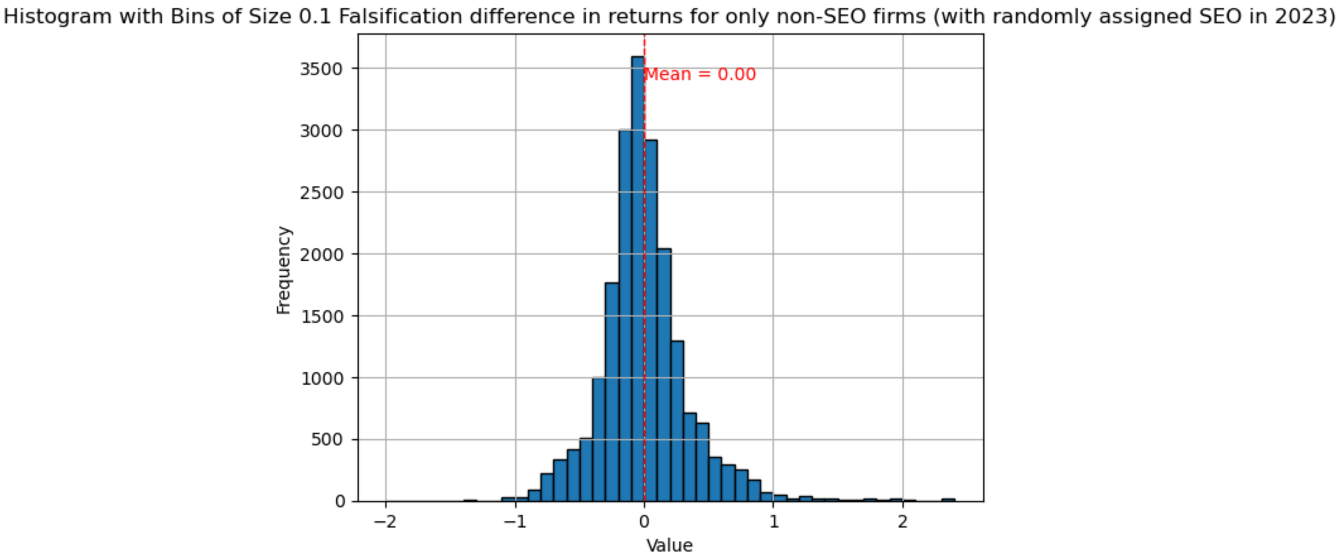


mean = 0.0284 t-statistic = 1.180 p-value = 0.2392

Appendix

Falsification test

- Monte-Carlo simulation
- With 10 rounds of random treatment assignment we get 0 as a mean average error



References

Abadie, Alberto, and Javier Gardeazabal. "The economic costs of conflict: A case study of the Basque Country." *American economic review* 93, no. 1 (2003): 113-132.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program." *Journal of the American statistical Association* 105, no. 490 (2010): 493-505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Comparative politics and the synthetic control method." *American Journal of Political Science* 59, no. 2 (2015): 495-510.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. "Synthetic difference-in-differences." *American Economic Review* 111, no. 12 (2021): 4088-4118.

Baker, Malcolm P., and Jeffrey Wurgler. "Why are dividends disappearing? An empirical analysis." (2002).

Cannelli, Loris, Giuseppe Nuti, Marzio Sala, and Oleg Szehr. "Hedging using reinforcement learning: Contextual k-armed bandit versus Q-learning." *The Journal of Finance and Data Science* 9 (2023): 100101.

Carpenter, Robert E., and Bruce C. Petersen. “Capital Market Imperfections, High- Tech Investment, and New Equity Financing.” *The Economic Journal* 112, no. 477 (2002): F54–72.

References

Chava, Sudheer and Reguly, Agoston. “The Long-Run Stock Market Performance of Mergers and Acquisitions” (2024). Georgia Tech Scheller College of Business Research Paper No. 4742401

Eckbo, B. Espen, and Ronald W. Masulis. ”Seasoned Equity Offerings: A Survey.” In *Finance, Handbooks in Operation Research and Management Science*, Vol. 9, edited by R.A. Jarrow and V.M.W.T. Ziemba, 1017-1072. Chapter 31. Amsterdam: North-Holland, 1995.

Eckbo, Espen, Ronald Masulis, and Øyvind Norli. “Seasoned Public Offerings: Resolution of the ‘New Issues Puzzle’.” *Journal of Financial Economics* 56 (2000): 251-291.

Geha, Georges. "Use of modern machine learning techniques to predict the occurrence and outcome of corporate takeover events." PhD diss., Massachusetts Institute of Technology, 2021.

Huang, Chia-Wei, Po-Hsin Ho, Chih-Yung Lin, and Ju-Fang Yen. ”Firm Age, Idiosyncratic Risk, and Long-Run SEO Underperformance.” *International Review of Economics & Finance* 34 (2014): 246-266.

Imbens, Guido W. "Matching methods in practice: Three examples." *Journal of Human Resources* 50, no. 2 (2015): 373-419.

References

Myers, Stewart C., and Nicholas S. Majluf. "Corporate financing and investment decisions when firms have information that investors do not have." *Journal of financial economics* 13, no. 2 (1984): 187-221.

Veld, Chris, Patrick Verwijmeren, and Yuriy Zabolotnyuk. "Wealth effects of seasoned equity offerings: A meta-analysis." *International Review of Finance* 20, no. 1 (2020): 77-131.

Ritter, Jay R. ”Investment Banking and Securities Issuance.” In *Handbook of the Economics of Finance*, edited by George M. Constantinides, Milton Harris, and Ren´e M. Stulz, 255-306. Amsterdam: Elsevier, 2003.

Young, Alwyn. “Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results.” *Quarterly Journal of Economics* 134, no. 2 (2019): 557–598.