

IQ from IP: Simplifying Search in Portfolio Choice

Online Appendix

Huaizhi Chen
University of Notre Dame

Lauren Cohen
Harvard Business School and NBER

Umit Gurun
University of Texas at Dallas and NBER

Dong Lou
London School of Economics and CEPR

Christopher Malloy
Harvard Business School and NBER

Appendix Table A1. Alternative Timing

This table reports calendar-time portfolio returns of various types of institutional holdings and trading. Panel A corresponds to tracked insider buys and Panel B corresponds to tracked insider sells. In terms of timing, we require that institutions both view the insider trading record (which takes place in quarter t) and trade the underlying stock in quarter $t+1$. We then analyze the equal-weighted returns to those trades in quarter $t+2$. Row 5 constructs a long-short portfolio, where the long portfolio includes all stocks of which the fund manager checks insider trades and also trade in the same direction, the short portfolio includes all other stocks that the fund manager buys/sells in the same quarter. In Row 8, the benchmark portfolio includes all stocks that are not checked by a fund manager but are traded by insiders in the same direction as the fund manager. Reported below are the quarterly Raw, DGTW, and 4 Factor adjusted returns of these aforementioned portfolios (as well as the long and short sides). T-statistics, reported in parenthesis, are based on Newey-West standard errors. 5% statistical significance is indicated in bold.

Panel A: Tracked Insider Buys	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Buying and Bought vs. Rest Bought	1.20% (1.05)	0.86% (0.76)	1.11% (1.01)
8) Checked Insider Buying and Bought vs. Not Checked and Bought	1.38% (1.30)	1.00% (0.90)	1.18% (1.09)
Panel B: Tracked Insider Sales	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Selling and Sold vs. Rest Sold	-1.22% (-2.22)	-1.16% (-2.16)	-1.42% (-2.67)
8) Checked Insider Selling and Sold vs. Not Checked and Sold	-1.07% (-2.05)	-1.10% (-2.08)	-1.26% (-2.42)

Appendix Table A2. Value-Weighted Portfolio Returns

This table reports calendar-time portfolio returns of various types of institutional holdings and trading. Panel A corresponds to tracked insider buys and Panel B corresponds to tracked insider sells. In terms of timing, we require that institutions both view the insider trading record and potentially trade the underlying stock in quarter t . We then analyze the holding-value weighted returns to those trades in quarter $t+1$. Row 5 constructs a long-short portfolio, where the long portfolio includes all stocks of which the fund manager checks insider trades and also trade in the same direction, the short portfolio includes all other stocks that the fund manager buys/sells in the same quarter. In Row 8, the benchmark portfolio includes all stocks that are not checked by a fund manager but are traded by insiders in the same direction as the fund manager. Reported below are the quarterly Raw, DGTW, and 4 Factor adjusted returns of these aforementioned portfolios (as well as the long and short sides). T-statistics, reported in parenthesis, are based on Newey-West standard errors. 5% statistical significance is indicated in bold.

Panel A: Tracked Insider Buys	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Buying and Bought vs. Rest Bought	2.18% (1.50)	2.27% (1.70)	1.75% (1.22)
8) Checked Insider Buying and Bought vs. Not Checked and Bought	2.63% (1.77)	2.71% (1.96)	2.18% (1.45)
Panel B: Tracked Insider Sales	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Selling and Sold vs. Rest Sold	-1.70% (-2.80)	-1.33% (-2.39)	-1.85% (-3.02)
8) Checked Insider Selling and Sold vs. Not Checked and Sold	-1.68% (-2.77)	-1.35% (-2.42)	-1.84% (-3.00)

Appendix Table A3: Trade-Size-Weighted Portfolio Returns

This table reports calendar-time portfolio returns of various types of institutional holdings and trading. Panel A corresponds to tracked insider buys and Panel B corresponds to tracked insider sells. In terms of timing, we require that institutions both view the insider trading record and trade the underlying stock in quarter t . We then analyze the trade-size weighted returns to those trades in quarter $t+1$. A stock's trade-size weight is the increase (decrease) in portfolio weight of the tracked buy (the tracked sale) relative to the stock's final weight (initial weight). Row 5 constructs a long-short portfolio, where the long portfolio includes all stocks of which the fund manager checks insider trades and also trade in the same direction, the short portfolio includes all other stocks that the fund manager buys/sells in the same quarter. In Row 8, the benchmark portfolio includes all stocks that are not checked by a fund manager but are traded by insiders in the same direction as the fund manager. Reported below are the quarterly Raw, DGTW, and 4 Factor adjusted returns of these aforementioned portfolios (as well as the long and short sides). T-statistics, reported in parenthesis, are based on Newey-West standard errors. 5% statistical significance is indicated in bold.

Panel A: Tracked Insider Buys	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Buying and Bought vs. Rest Bought	3.92% (2.09)	3.95% (2.22)	4.05% (2.28)
8) Checked Insider Buying and Bought vs. Not Checked and Bought	3.75% (2.00)	3.65% (2.04)	3.63% (1.96)

Panel B: Tracked Insider Sales	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Selling and Sold vs. Rest Sold	-2.89% (-3.37)	-2.64% (-3.29)	-3.11% (-3.81)
8) Checked Insider Selling and Sold vs. Not Checked and Sold	-3.31% (-3.78)	-2.96% (-3.38)	-3.61% (-4.25)

Appendix Table A4: Exclude the Post-2013 Period from Our Sample

This table reports calendar-time portfolio returns of various types of institutional holdings and trading excluding the post 2013 sample period. Panel A corresponds to tracked insider buys and Panel B corresponds to tracked insider sells. In terms of timing, we require that institutions both view the insider trading record and trade the underlying stock in quarter t , and analyze the equal-weighted returns to those trades in quarter $t+1$. Row 5 constructs a long-short portfolio, where the long portfolio includes all stocks of which the fund manager checks insider trades and also trade in the same direction, the short portfolio includes all other stocks that the fund manager buys/sells in the same quarter. In Row 8, the benchmark portfolio includes all stocks that are not checked by a fund manager but are traded by insiders in the same direction as the fund manager. Reported below are the quarterly Raw, DGTW, and 4 Factor adjusted returns of these aforementioned portfolios (as well as the long and short sides). T-statistics, reported in parenthesis, are based on Newey-West standard errors. 5% statistical significance is indicated in bold.

Panel A: Tracked Insider Buys	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Buying and Bought vs. Rest Bought	1.58% (1.02)	1.27% (0.92)	1.47% (1.00)
8) Checked Insider Buying and Bought vs. Not Checked and Bought	1.70% (1.17)	1.48% (1.09)	1.64% (1.11)
Panel B: Tracked Insider Sales	Excess Returns	DGTW	4-Factor Alpha
5) Checked Insider Selling and Sold vs. Rest Sold	-1.88% (-3.02)	-1.51% (-2.44)	-1.85% (-2.96)
8) Checked Insider Selling and Sold vs. Not Checked and Sold	-1.82% (-3.02)	-1.52% (-2.46)	-1.81% (-2.93)

Appendix Table A5: Description of the Deciphering Procedure

The deciphering process decrypts the IP addresses from the SEC Edgar log files in order to match them to a commercial IP geolocation/organizational database. The IP addresses in the dataset are partially anonymized using a static cipher. A standard IP address has the form $###.###.###.###$. Each octet $###$ separated by periods has a numerical value between 0 and 255. An IP address in the Edgar data has the form $###.###.###.&&&$. The first 9 digits of the Edgar server IP correspond to the actual IP that visited the website (024.145.236). The last 3 digits correspond to a hidden octet (*jcf*). That is, each row of the data corresponds to a certain IP address (24.145.236.jcf) viewing a specific filing coded by an accession number (0000891020-04-000160) at a specific time and date (12:00 am on April 31, 2004).

The static ciphers in the Edgar log are 3-digit alphabetical codes that *uniquely* represent octets between 0 and 255 throughout the data. It is very apparent that this cipher is static in nature.

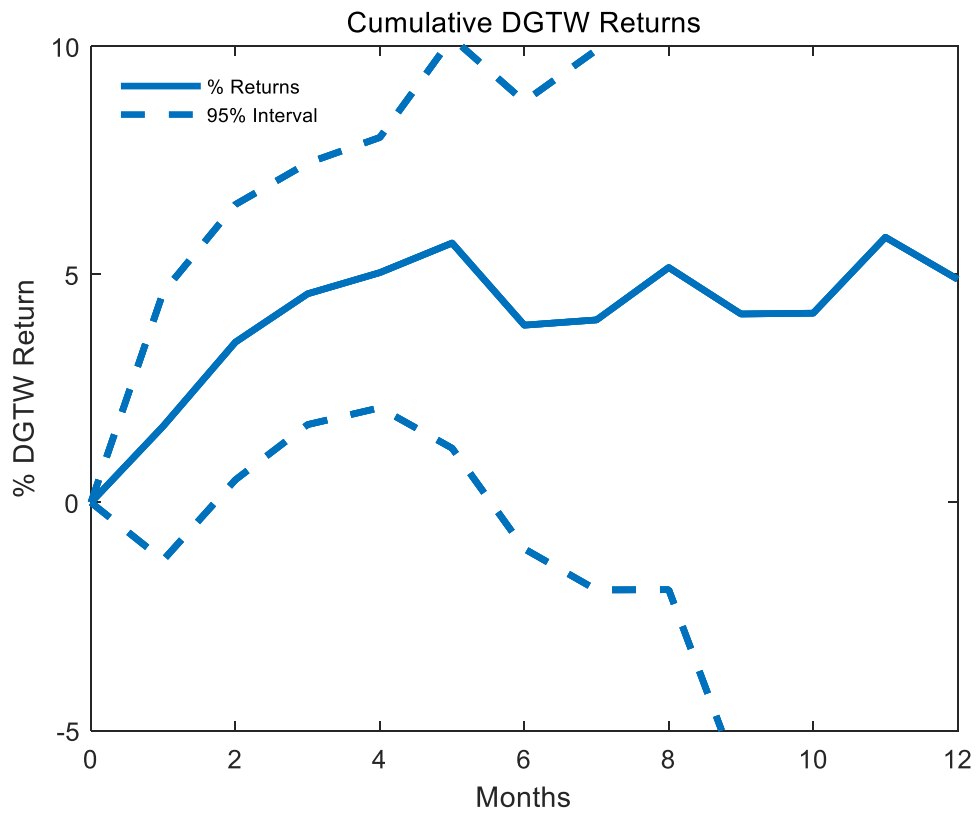
- 1) The same cipher would visit similar forms in multiple days and months. For example, the same IP address, with *aca* as the cipher, would visit AIG filings in January 2009, February 2009, and so on.
- 2) We find that ciphers with the lowest frequency tend to be the same two: *ghf* and *ghg* throughout the data. We ended up matching *ghf* and *ghg* to 0 and 255 respectively. IP addresses ending with 0 and 255 tend to be reserved for network administrative purposes, and are generally less frequent than other addresses.

We obtain another dataset of server logs from a private but well-trafficked website in order to decrypt the EDGAR IP address. Specifically, we identify all instances where the first 9 digits of an IP address in the Edgar files match an IP address from this private server.

If 1.1.1.aaa visited the SEC server in its history and 1.1.1.111 visited this website in 2016, then the cipher *aaa* will have one match to 111. Since there are multiple matches between a cipher and a set of octets, we compose a scoring system that counts the number of times each cipher is matched to each octet. The most frequent match is our candidate for the cipher link. The link between a cipher and its most frequently matched octet is distinct for the vast majority of the ciphers (236 out of 256 ciphers). For example, *aba*'s most frequent octet match is 009, no other cipher matches to 009 as their most frequent match. In this set of 236 matches, the identified octet is on average 96% more frequent than the next most frequent match. I.e. *aej* is matched to its more frequent octet, 71, 298 times. The next most frequent octet, 135, has about 139 matches to *aej*.

The remainder of the matching - where multiple ciphers are matched to a single octet - is done through a process of elimination. For example, if both *aaa* and *aab* score 001 as their most frequent match, we distinguish the two by examining which of the two ciphers are more frequently linked to 001. If 001 has many more matches with *aaa* than *aab*, then *aaa* is matched to 001. In this case 001 is removed as a potential choice for *aab*'s matching. *aab* is then linked to its next most frequent octet. We only had to iterate this process once in order to match the remaining 20 pairs. When we examine the frequency of octet to cipher matches in this subset, the most frequent cipher for each of these octet accounts for significantly more matches than the next most frequent cipher (116% more frequent).

At the end of our decryption, we obtain 256 pairs of unique matchings between the hidden octet and the actual octet for the EDGAR database. We present our results in table 1.



Appendix Figure A1. This figure shows the long-run returns (along with the 95% confidence interval) to tracked-insider trades by fund managers. Specifically, it is the difference between the portfolio described in row 5 of Table 6 Panel A, and that in row 5 of Table 6 Panel B over the next 12 months.