Replication of Cong et al. [2020]

Haksoo Lee

December 17, 2021

Abstract

I replicate the main findings of Cong et al. [2020]. Namely, I replicate Table 3; Table 4; and Figure 4 of Cong et al. [2020].

Table 3: Out-of-Sample Performance of Long-Only AP

This table presents the OOS performance of a long-only AlphaPortfolio taking positions only in the highest decile of winner scores, with Return, Std.Dev., and Sharpe ratio all annualized. Portfolio returns are further adjusted by the CAPM, Fama-French-Carhart 4-factor model (FFC), Fama-French-Carhart 4-factor and Pastor-Stambaugh liquidity factor model (FFC+PS), Fama-French 5-factor model (FF5), Fama-French 6-factor model (FF6), Stambaugh-Yuan 4-factor model (SY), and Hou-Xue-Zhang 4-factor model (Q4). The first columns present the alphas for the overall sample. The remaining four columns present alphas for subsamples excluding microcap firms in the smallest decile and quintile, respectively. q_n symbolizes the n^{th} NYSE size percentile. "*," "**," and "***" denote significance at the 10%, 5%, and 1% level, respectively.

	AP Performance				AP Excess Alpha					
	(1)	(2)	(3)		(4)	(5)	(6)	(7)	(8)	(9)
Firms	All	$> q_{10}$	$> q_{20}$	Factor Models	All		$> q_{10}$		$> q_{20}$	
					$\alpha(\%)$	R^2	$\alpha(\%)$	R^2	$\alpha(\%)$	R^2
Return(%)	22.41	37.60	44.27	CAPM	12.88***	0.477	25.80***	0.508	31.79***	0.485
Std.Dev.(%)	19.28	25.08	27.63	FFC	12.54***	0.791	25.79***	0.798	30.97***	0.733
Sharpe	1.16	1.50	1.60	FFC+PS	11.02***	0.795	22.20***	0.806	28.20***	0.738
Skewness	0.67	1.03	1.55	FF5	9.91***	0.731	21.62***	0.755	27.84***	0.696
Kurtosis	5.64	4.72	7.55	FF6	12.10***	0.791	24.12***	0.802	30.31***	0.734
Turnover	0.38	0.44	0.48	SY	15.38***	0.720	26.29***	0.739	31.68***	0.681
MDD	0.24	0.27	0.25	Q4	14.51***	0.698	26.83***	0.714	33.01***	0.656

Figure 1: Cong et al. [2020] Table 3

1 Findings of Cong et al. [2020]

In this section, I present the findings of Cong et al. [2020] that I replicate. Figures 1-6 show the authors' results.

Table 4: Out-of-Sample Performance in Recent Years

This table reports alphas for portfolios of long/short stocks in the highest/lowest decile of winner scores from 2001 to 2016. Return, Std.Dev., and Sharpe ratio are all annualized. Portfolio returns are further adjusted by the CAPM, Fama-French-Carhart 4-factor model (FFC), Fama-French-Carhart 4-factor and Pastor-Stambaugh liquidity factor model (FFCPS), Fama-French 5-factor model (FF5), Fama-French 6-factor model (FF6), Stambaugh-Yuan 4-factor model (SY), and Hou-Xue-Zhang 4-factor model (Q4). The first columns present the alphas for the overall sample. The remaining four columns present alphas for subsamples excluding microcap firms in the smallest decile and quintile, respectively. q_n symbolizes the n^{th} NYSE size percentile. "*," "**," and "***" denote significance at the 10%, 5% and 1% level, respectively.

	AP Performance				AP Excess Alpha					
	(1)	(2)	(3)		(4)	(5)	(6)	(7)	(8)	(9)
Firms	All	$> q_{10}$	$> q_{20}$	Factor Models	All $\alpha(\%)$	R^2	$> q_{10}$ $\alpha(\%)$	R^2	$> q_{20}$ $\alpha(\%)$	R^2
Return(%)	18.10	16.10	16.60	CAPM	16.5^{***}	0.007	13.6^{***}	0.136	13.8^{***}	0.176
Std.Dev.(%)	9.20	7.90	8.90	FFC	16.3^{***}	0.078	12.9^{***}	0.497	13.3^{***}	0.594
Sharpe	1.97	2.04	1.87	FFC+PS	15.7^{***}	0.080	11.7^{***}	0.506	11.7^{***}	0.606
Skewness	1.67	1.53	1.61	FF5	18.0^{***}	0.151	13.8^{***}	0.426	14.6^{***}	0.432
Kurtosis	5.95	4.23	3.59	FF6	17.8^{***}	0.174	13.3^{***}	0.560	14.0^{***}	0.620
Turnover	0.25	0.23	0.25	SY	18.9^{***}	0.065	15.3^{***}	0.428	16.5^{***}	0.502
MDD	0.05	0.03	0.04	Q4	16.9^{***}	0.121	13.7^{***}	0.532	14.6^{***}	0.551

Figure 2: Cong et al. [2020] Table 4

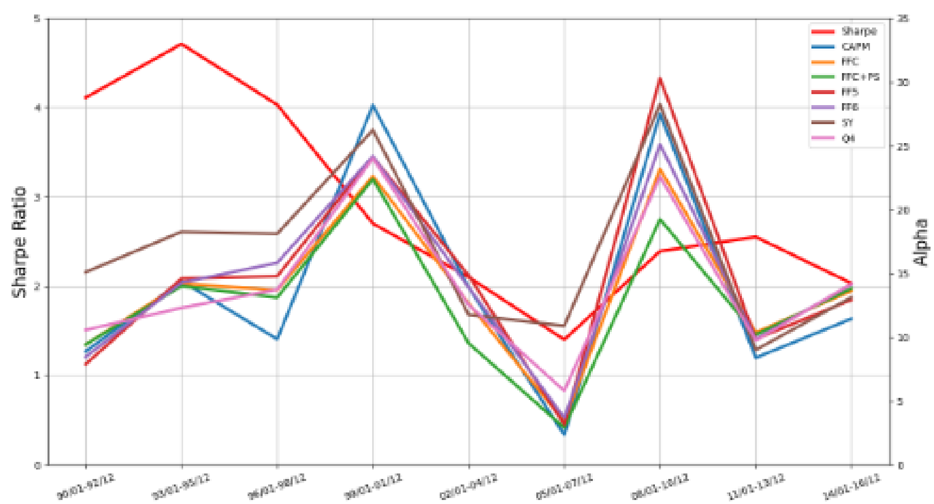

Figure 4: Trends of AlphaPortfolio Performance.

Figure 3: Cong et al. [2020] Figure 4

2 Replication

2.1 Learning problem definition, algorithm, and data

Due to computational complexity of Cong et al. [2020] (and the complexity of the reinforcement learning models that the authors' employ), I consider a smaller RL model with a subset of features and an alternate learning algorithm. I define the learning problem for the RL agent to be maximizing the Sharpe ratio of her long-only portfolio measured at quarterly frequency for the duration of one year. The RL agent learns to choose the Sharpe ratio maximizing portfolio weight $a_t \in [0, 1]$ on an asset (such that a_t of her wealth is invested in the asset and $1 - \alpha_t$ is invested in the risk-free asset). Each learning episode consists of 252 trading days, in which the RL agent receives feedback/rewards in terms of the Sharpe ratio of her portfolio performance in the previous 63 days. As such, the RL agent's action has delayed effects on future rewards. Note, as in Cong et al. [2020], I consider long-only portfolios as the RL agent's action space is limited to $[0, 1]$.

With respect to the learning algorithm, I consider a Monte Carlo policy gradient algorithm, REINFORCE. The Monte Carlo policy gradient algorithm appears more effective than other RL algorithms–such as the Q-learning algorithm–for the current learning problem for several reasons. First, Monte Carlo policy gradient algorithms have strong theoretical convergence properties and exhibit more robust convergence in the learning problems of the paper experimentally. (Phansalkar and Thathachar [1995] prove local and global theorems for a branch of the Monte Carlo policy gradient algorithm.) Second, the policy gradient algorithm can effectively adapt to continuous action spaces where the number of possible actions is infinite. In contrast, the Q-learning algorithm (e.g., the DQN agent of Mnih et al. [2013]), which is based on the action-value function (Q-function), requires discretization of the action space which may hinder learning and which is unnatural for the learning problem. Third, with $\pi(a|x, \theta) \in (0, 1)$, exploration is naturally ensured and implemented. The form of exploration induced by the policy gradient algorithm may be particularly expedient for locating an optimal policy in the present learning problems. Lastly, the policy gradient algorithm can learn optimal policies that are stochastic in nature. In poker, for instance, a mixed strategy that bluffs with a positive probability may be optimal. Yet, finding such policy may be difficult for canonical action-value function based algorithms without awkward adaptations. The Monte Carlo policy gradient algorithm is disadvantageous in that, as a Monte Carlo method, it suffers from high variance of its estimates and from slow learning, which hinders real-time online learning.

With respect to features, I consider the set of variables that Cong et al. [2020] find to be dominant–that is, useful (in a Kalman gain sense) for the learning problem–from the economic distillation procedure as outline in Section 5. Namely, I consider: INVT (inventory changes), NPM (net profit margin), Q (Tobin's Q), C (cash and short-term investments), IVOL (idiosyncratic volatility); along with price-based signals, namely costs of equity and debt, and the stock's past returns. The RL agent has memory of 63 trading days such that she observes the 63-day history of the features in each period to make an appropriate decision.

All variables are downloaded from WRDS, making use of CRSP, WRDS Beta Suite, and Compustat. The fundamental variables are in quarterly frequency and the priced-based features are at daily frequency. The training data consists of Apple, Microsoft, and Tesla from 4 April 2014 to 31 December 2020.

2.2 Replication results

Table 1 show the replication of Table 4 of Cong et al. [2020]. The return profile is comparable to column (1) of Cong et al. [2020]. The replicating RL model, however, generates larger Sharpe ratios. Figures 4-6 show replication of Figure 4 of Cong et al. [2020].

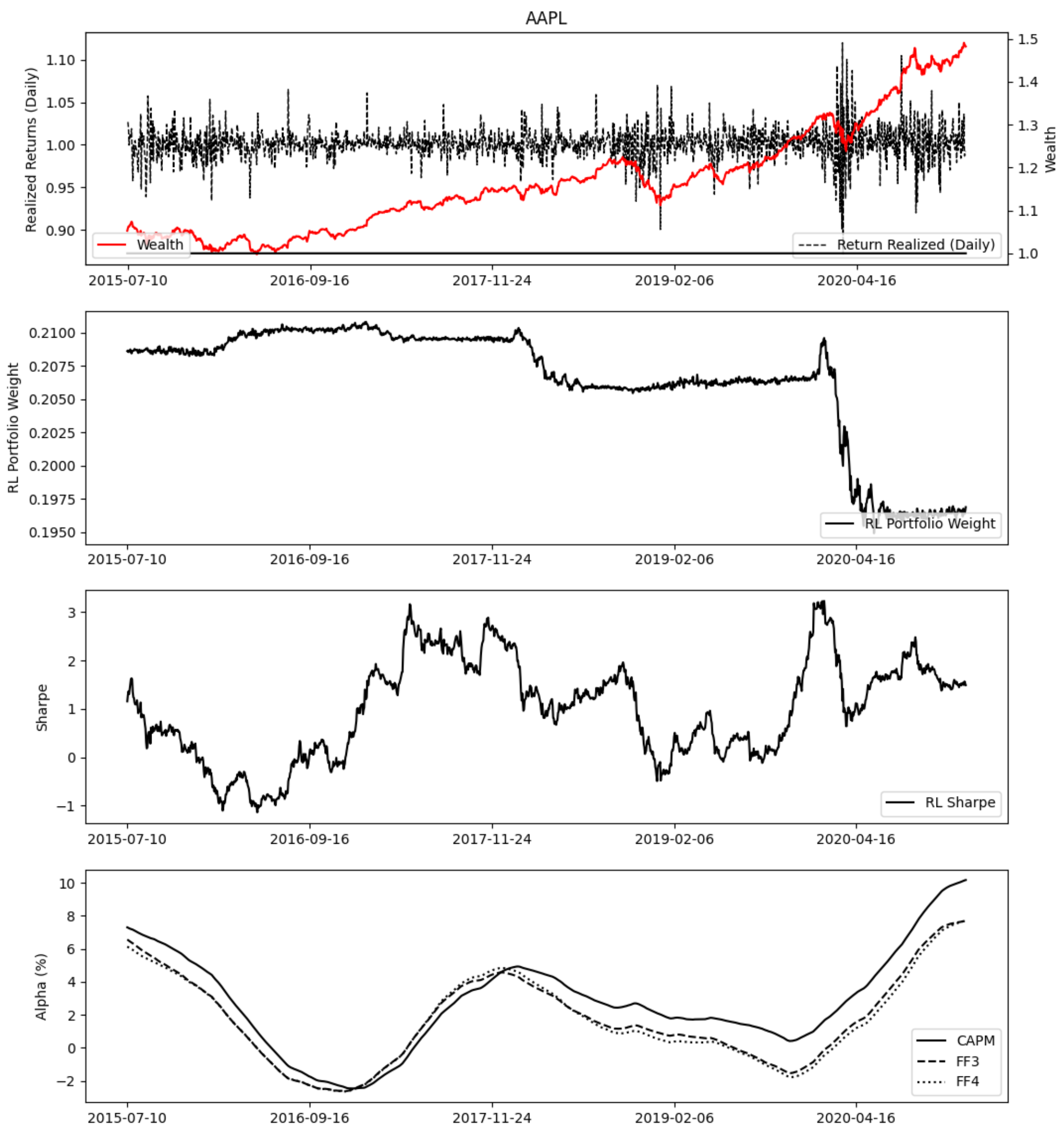

Figure 4: Replicating Cong et al. [2020] Figure 4

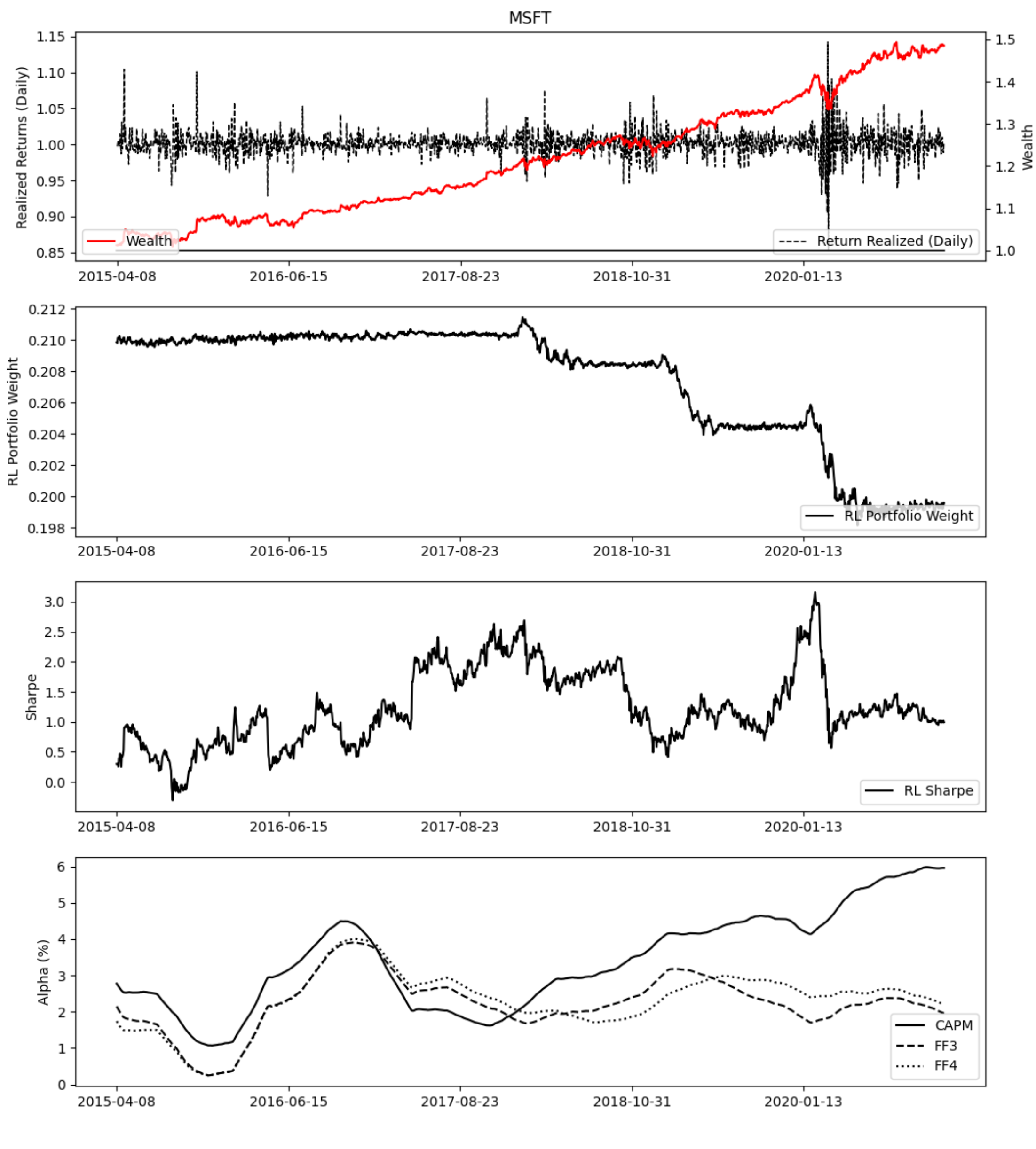

Figure 5: Replicating Cong et al. [2020] Figure 4

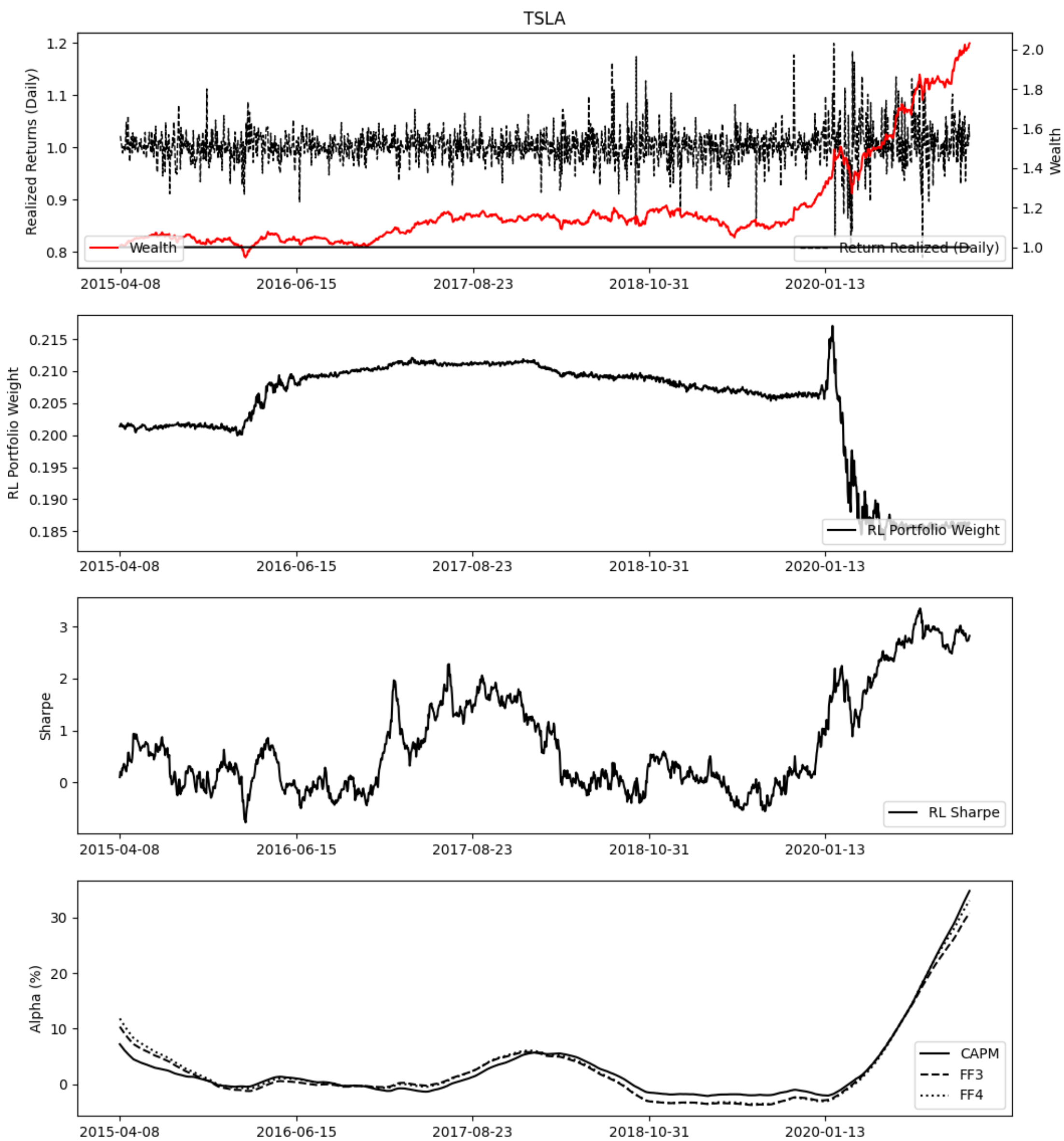

Figure 6: Replicating Cong et al. [2020] Figure 4

Table 1: Relication of Table 4 of Cong et al. [2020]

	AAPL	MSFT	TSLA		AAPL	MSFT	TSLA
Return	0.14	0.24	0.22	Factor Model	Alpha (%)	Alpha (%)	Alpha (%)
Std	0.11	0.12	0.24	CAPM	2.7	3.39	2.66
Sharpe	10.46	10.12	5.17	FF3	1.69	2.23	2.34
Skewness	0.81	0.3	1.81	FF4	1.58	2.3	2.57
Kurtosis	-0.04	-1	2.25				
Turnover	0.04	0.04	0.09				
MDD	0.1	0.08	0.16				

References

Lin William Cong, Ke Tang, Jingyuan Wang, and Yang Zhang. AlphaPortfolio: Direct Construction through Reinforcement Learning and Interpretable AI. *Ssrn*, 2020. ISSN 1556-5068. doi: 10.2139/ssrn.3554486.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with Deep Reinforcement Learning. 12 2013.

V. V. Phansalkar and M. A. L. Thathachar. Local and Global Optimization Algorithms for Generalized Learning Automata. *Neural Computation*, 7(5):950–973, 9 1995. ISSN 0899-7667. doi: 10.1162/neco.1995.7.5.950.