Internet Appendix for "Maximizing the Sharpe Ratio: A Genetic Programming Approach"

A.1 Flow Charts of GP

Figure A.1 illustrates the flow charts of the training process for genetic programming. Step 1: *Initialization.* Create initial random individuals with the population size of *Pop* and evaluate the fitness of each individual. Step 2: *Transformation.* Apply the genetic operators, such as crossover and mutation, to parent individuals to create offspring, where parent individuals are selected from the current population with the selection probabilities biased in favor of relatively fit individuals. Step 3: *Evaluation and Selection.* Evaluate the fitness of offspring and parent individuals, and those with greater Sharpe ratios will survive in the next generation. Step 4: *Iteration.* Successive generations are iteratively generated in the same way, until the generation number exceeds a pre-defined maximum generation *Gen.*

A.2 Further Empirical Results

We provide the complete results for the robustness checks discussed in the paper. Below, we briefly describe the contents of the appendix tables.

- **Table** A.1: Spread Portfolios: Excluding Micro-Stocks

- **Table** A.2: Spread Portfolios Controlling for Other Models: Alternative Search Depth

- **Table** A.3: Spread Portfolios Controlling for Other Models: Alternative Model Average Number
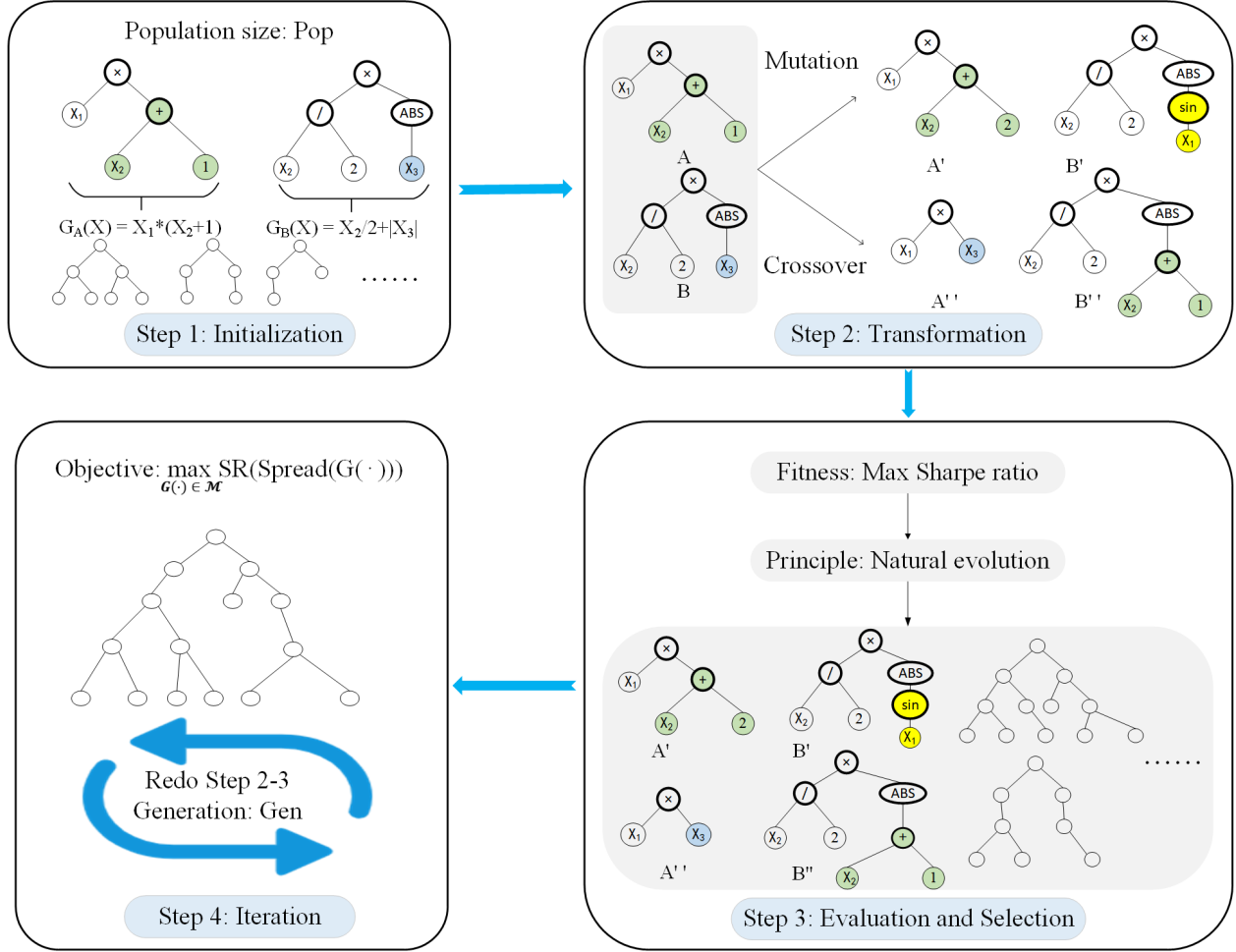
Figure A.1: Flow Charts of GP. This figure illustrates the flow charts of the training process for genetic programming. Step 1: *Initialization*. Create initial random individuals with the population size of *Pop* and evaluate the fitness of each individual. Step 2: *Transformation*. Apply the genetic operators, such as crossover and mutation, to parent individuals to create offspring, where parent individuals are selected from the current population with the selection probabilities biased in favor of relatively fit individuals. Step 3: *Evaluation and Selection*. Evaluate the fitness of offspring and parent individuals, and those with greater Sharpe ratios will survive in the next generation. Step 4: *Iteration*. Successive generations are iteratively generated in the same way, until the generation number exceeds a pre-defined maximum generation *Gen*.

Table A.1
Spread Portfolios: Excluding Micro-Stocks

The table reports the summary statistics for the decile spread portfolios generated by the GP and other models. We exclude the micro-stocks whose market capitalization is below the 20th percentile of the market capitalization of the NYSE stocks. For each model, we report the average monthly return in percentage points, the Newey and West (1987) robust t-statistic, standard deviation in percentage points, the annualized Sharpe ratio (Sharpe), the skewness (Skew), and the maximum drawdown (MDD) in percentage points. The sample period is from 1991:01 to 2021:12.

	GP	Ridge	LASSO	ENET	PCR	PLS	NN1	NN2	NN3	NN4	NN5
Low	0.36	0.78	0.74	0.76	0.76	0.76	0.87	0.83	0.77	0.58	0.49
2	0.57	0.70	0.80	0.79	0.81	0.72	0.83	0.84	0.95	0.76	0.78
3	0.66	0.92	0.85	0.87	0.88	0.94	0.87	0.88	0.97	0.84	0.85
4	0.68	1.02	1.11	1.07	0.96	1.00	1.00	0.96	1.03	0.97	0.95
5	0.90	1.08	1.04	1.04	1.16	1.13	1.14	1.09	1.19	0.89	1.01
6	1.06	1.20	1.25	1.21	1.20	1.17	1.18	1.21	1.09	0.99	0.90
7	1.19	1.24	1.29	1.31	1.36	1.22	1.13	1.27	1.17	1.13	1.05
8	1.46	1.46	1.43	1.44	1.44	1.47	1.12	1.32	1.25	1.21	1.14
9	1.47	1.51	1.40	1.42	1.29	1.54	1.14	1.41	1.44	1.34	1.23
High	1.77	1.61	1.68	1.66	1.68	1.60	1.54	1.70	1.76	1.52	1.48
H-L	1.41***	0.82***	0.94***	0.90***	0.91***	0.83***	0.66***	0.86***	0.98***	0.94***	0.98***
t-stat	(5.30)	(2.96)	(3.26)	(3.16)	(3.19)	(2.97)	(2.91)	(2.98)	(3.77)	(4.50)	(4.03)
Std. dev.	4.23	4.90	4.77	4.77	4.66	4.91	4.37	4.65	5.09	4.28	4.03
Sharpe	1.16	0.59	0.68	0.65	0.68	0.59	0.53	0.64	0.67	0.76	0.84
Skew	0.87	0.35	0.22	0.26	0.26	0.33	-0.49	1.01	0.99	-0.24	-0.06
MDD	21.73	41.96	39.40	40.87	35.29	43.94	37.32	61.04	47.91	26.22	28.85

4

Table A.2
Spread Portfolios Controlling for Other Models: Alternative Search Depth

This table reports the average monthly return for the decile spread portfolios of each model controlling for one of the other models. Panel A provides the results for the GP controlling for one of the other models. We first regress the expected stock returns generated by GP on those generated by another model, and then construct the long-short spread portfolio sorted by the associated residuals. Panel B provides the results for other models controlling for the GP. We report the results under various GP search depth parameters of $<Pop, Gen>$. We take the average of the top M (M=5) GP models with the highest training sample Sharpe ratio as the final GP model. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

	Ridge	LASSO	ENET	PCR	PLS	NN1	NN2	NN3	NN4	NN5
Panel A: GP with various parameters of $<Pop, Gen>$, controlling for other models										
<100, 10>	0.75***	0.77***	0.77***	0.78***	0.75***	0.82***	0.67***	0.72***	0.74***	0.75***
	(4.29)	(4.26)	(4.27)	(4.22)	(4.25)	(5.27)	(4.16)	(4.53)	(4.46)	(4.58)
<100, 20>	0.30**	0.30**	0.30**	0.31**	0.30**	0.44***	0.36***	0.28**	0.45***	0.56***
	(2.39)	(2.34)	(2.33)	(2.41)	(2.39)	(3.70)	(2.83)	(2.57)	(3.41)	(3.96)
<100, 40>	0.69***	0.71***	0.72***	0.77***	0.69***	0.76***	0.66***	0.80***	0.77***	0.72***
	(3.72)	(3.75)	(3.79)	(4.03)	(3.69)	(4.42)	(3.88)	(4.66)	(5.21)	(5.07)
<200, 10>	0.40***	0.44***	0.42***	0.44***	0.40***	0.51***	0.44***	0.42***	0.58***	0.62***
	(2.71)	(2.84)	(2.94)	(2.77)	(2.75)	(3.20)	(2.96)	(2.73)	(3.41)	(3.42)
<200, 20>	0.42***	0.45***	0.44***	0.40**	0.42***	0.51***	0.39***	0.42***	0.57***	0.65***
	(2.73)	(2.81)	(2.74)	(2.43)	(2.71)	(3.89)	(3.08)	(2.84)	(3.54)	(3.71)
<200, 40>	0.63***	0.60***	0.62***	0.70***	0.64***	0.70***	0.58***	0.60***	0.71***	0.76***
	(4.08)	(3.97)	(3.81)	(4.27)	(4.14)	(4.10)	(3.38)	(4.15)	(4.68)	(4.37)
<400, 10>	0.80***	0.74***	0.75***	0.74***	0.79***	0.76***	0.66***	0.72***	0.79***	0.77***
	(4.27)	(4.21)	(4.25)	(4.06)	(4.21)	(4.65)	(4.28)	(4.47)	(4.58)	(4.30)
<400, 20>	0.58***	0.57***	0.57***	0.60***	0.58***	0.68***	0.50***	0.66***	0.60***	0.61***
	(4.65)	(4.43)	(4.35)	(4.51)	(4.58)	(4.56)	(3.55)	(4.32)	(4.28)	(4.07)
<400, 40>	0.51***	0.63***	0.58***	0.52***	0.52***	0.61***	0.48***	0.39***	0.46***	0.47***
	(4.35)	(5.42)	(5.05)	(4.24)	(4.42)	(4.13)	(3.51)	(3.35)	(3.43)	(3.40)
Average	0.56***	0.58***	0.57***	0.58***	0.57***	0.64***	0.53***	0.56***	0.63***	0.66***
	(3.69)	(3.78)	(3.73)	(3.66)	(3.68)	(4.21)	(3.51)	(3.74)	(4.11)	(4.10)
Panel B: Other Models, controlling for GP with various parameters of $<Pop, Gen>$										
<100, 10>	-0.18	-0.18	-0.24	-0.21	-0.18	-0.22	-0.09	0.23	-0.06	-0.19
	(-0.73)	(-0.65)	(-0.94)	(-0.82)	(-0.74)	(-1.11)	(-0.37)	(1.11)	(-0.15)	(-0.54)
<100, 20>	-0.12	-0.28	-0.27	-0.31	-0.12	-0.09	-0.14	0.05	0.01	-0.06
	(-0.52)	(-1.22)	(-1.13)	(-1.34)	(-0.51)	(-0.80)	(-0.85)	(0.33)	(0.05)	(-0.23)
<100, 40>	-0.01	-0.15	-0.01	-0.03	-0.06	-0.11	0.04	0.42	0.20	0.76
	(-0.07)	(-0.57)	(-0.05)	(-0.13)	(-0.26)	(-0.53)	(0.19)	(2.01)	(0.56)	(1.77)
<200, 10>	-0.05	0.09	0.08	-0.19	-0.06	-0.06	0.02	0.34	0.19	-0.22
	(-0.18)	(0.30)	(0.27)	(-0.70)	(-0.22)	(-0.32)	(0.12)	(1.48)	(0.51)	(-0.49)
<200, 20>	-0.17	-0.18	-0.18	-0.20	-0.15	-0.16	-0.00	0.05	-0.48	-0.66
	(-0.74)	(-0.74)	(-0.74)	(-0.82)	(-0.66)	(-1.22)	(-0.05)	(0.28)	(-1.61)	(-1.80)
<200, 40>	-0.14	-0.18	-0.25	-0.14	-0.13	-0.13	-0.00	0.26	-0.06	-0.24
	(-0.58)	(-0.69)	(-0.95)	(-0.54)	(-0.57)	(-0.69)	(-0.02)	(1.24)	(-0.16)	(-0.77)
<400, 10>	-0.20	-0.19	-0.21	-0.38	-0.22	-0.16	-0.01	0.09	-0.05	-0.13
	(-0.72)	(-0.67)	(-0.80)	(-1.47)	(-0.83)	(-0.77)	(-0.07)	(0.43)	(-0.15)	(-0.35)
<400, 20>	-0.16	-0.11	-0.24	-0.36	-0.15	-0.16	-0.09	0.24	0.20	0.12
	(-0.64)	(-0.39)	(-0.92)	(-1.35)	(-0.61)	(-0.93)	(-0.39)	(1.10)	(0.55)	(0.42)
<400, 40>	0.10	0.04	0.07	0.02	0.04	0.00	0.17	0.39	0.35	0.46
	(0.38)	(0.15)	(0.26)	(0.09)	(0.18)	(0.04)	(0.75)	(1.97)	(0.97)	(1.60)
Average	-0.10	-0.13	-0.14	-0.20	-0.11	-0.12	-0.01	0.23	0.03	-0.02
	(-0.42)	(-0.50)	(-0.56)	(-0.79)	(-0.47)	(-0.70)	(-0.08)	(1.11)	(0.06)	(-0.04)

Table A.3
Spread Portfolios Controlling for Other Models: Alternative Model Average Number

This table reports the average monthly return for the decile spread portfolios of each model controlling for one of the other models. Panel A provides the results for the GP controlling for one of the other models. We first regress the expected stock returns generated by GP on those generated by another model, and then construct the long-short spread portfolio sorted by the associated residuals. Panel B provides the results for other models controlling for the GP. We take the average of the top M GP models with the highest training sample Sharpe ratio as the final GP model. We report the results for $M = 3$, 5, and 10. The GP search depth parameter pair $<Pop, Gen>$ is determined by the validation sample and is $<200, 40>$. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

	Ridge	LASSO	ENET	PCR	PLS	NN1	NN2	NN3	NN4	NN5
Panel A: GP with various model average number M, controlling for other models										
M=3	0.64***	0.61***	0.61***	0.66***	0.63***	0.67***	0.58***	0.65***	0.73***	0.76***
	(4.13)	(4.08)	(3.81)	(4.09)	(3.92)	(4.00)	(3.60)	(4.23)	(4.48)	(4.50)
M=5	0.63***	0.60***	0.62***	0.70***	0.64***	0.70***	0.58***	0.60***	0.71***	0.76***
	(4.08)	(3.97)	(3.81)	(4.27)	(4.14)	(4.10)	(3.38)	(4.15)	(4.68)	(4.37)
M=10	0.65***	0.73***	0.67***	0.76***	0.64***	0.79***	0.59***	0.66***	0.75***	0.80***
	(3.47)	(3.88)	(3.64)	(3.99)	(3.48)	(5.13)	(4.02)	(4.48)	(4.34)	(4.48)
Average	0.64***	0.65***	0.63***	0.71***	0.64***	0.72***	0.58***	0.64***	0.73***	0.77***
	(3.89)	(3.98)	(3.75)	(4.12)	(3.85)	(4.41)	(3.67)	(4.29)	(4.50)	(4.45)
Panel B: Other models, controlling for GP with various model average number M										
M=3	-0.12	-0.19	-0.17	-0.15	-0.16	-0.13	-0.01	0.33	0.00	-0.08
	(-0.50)	(-0.74)	(-0.65)	(-0.59)	(-0.69)	(-0.70)	(-0.06)	(1.51)	(0.01)	(-0.23)
M=5	-0.14	-0.18	-0.25	-0.14	-0.13	-0.13	-0.00	0.26	-0.06	-0.24
	(-0.58)	(-0.69)	(-0.95)	(-0.54)	(-0.57)	(-0.69)	(-0.02)	(1.24)	(-0.16)	(-0.77)
M=10	-0.20	-0.14	-0.19	-0.14	-0.21	-0.21	-0.09	0.11	-0.35	-0.52
	(-0.87)	(-0.51)	(-0.76)	(-0.54)	(-0.92)	(-1.14)	(-0.42)	(0.52)	(-0.95)	(-1.61)
Average	-0.15	-0.17	-0.20	-0.14	-0.17	-0.16	-0.03	0.23	-0.14	-0.28
	(-0.65)	(-0.65)	(-0.79)	(-0.56)	(-0.73)	(-0.84)	(-0.17)	(1.09)	(-0.37)	(-0.87)

Table A.4
Risk-adjusted Returns under GP: Alternative Search Depth

The table reports the risk-adjusted returns of the spread portfolios generated by other methods under the CAPM augmented by the GPF factor. Newey and West (1987) robust t-statistics are reported in parentheses. We report the results under various GP search depth parameters of $<Pop, Gen>$. We take the average of the top M (M=5) GP models with the highest training sample Sharpe ratio as the final GP model. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

$<Pop, Gen>$	Ridge	LASSO	ENET	PCR	PLS	NN1	NN2	NN3	NN4	NN5
$< 100, 10 >$	-0.45	-0.38	-0.40	-0.35	-0.47	-0.30	-0.15	-0.34	-0.07	-0.15
	(-1.59)	(-1.28)	(-1.35)	(-1.12)	(-1.64)	(-1.05)	(-0.54)	(-1.35)	(-0.31)	(-0.68)
$< 100, 20 >$	-0.08	-0.04	-0.06	-0.09	-0.10	-0.17	0.19	0.10	0.10	-0.00
	(-0.30)	(-0.14)	(-0.24)	(-0.31)	(-0.36)	(-0.63)	(0.67)	(0.41)	(0.45)	(-0.04)
$< 100, 40 >$	-0.47	-0.41	-0.41	-0.39	-0.48	-0.25	-0.10	-0.39	-0.02	-0.16
	(-1.77)	(-1.46)	(-1.53)	(-1.34)	(-1.79)	(-0.94)	(-0.39)	(-1.57)	(-0.08)	(-0.78)
$< 200, 10 >$	-0.20	-0.14	-0.16	-0.19	-0.21	-0.00	0.21	-0.05	0.20	0.04
	(-0.80)	(-0.52)	(-0.63)	(-0.65)	(-0.82)	(-0.01)	(0.79)	(-0.21)	(0.84)	(0.20)
$< 200, 20 >$	-0.01	0.03	0.01	-0.01	-0.02	0.02	0.30	0.11	0.25	0.05
	(-0.05)	(0.11)	(0.05)	(-0.05)	(-0.10)	(0.07)	(1.07)	(0.46)	(1.05)	(0.23)
$< 200, 40 >$	-0.45	-0.35	-0.35	-0.38	-0.46	-0.05	-0.16	-0.30	-0.10	-0.18
	(-1.64)	(-1.20)	(-1.24)	(-1.29)	(-1.65)	(-0.18)	(-0.57)	(-1.25)	(-0.43)	(-0.89)
$< 400, 10 >$	-0.36	-0.32	-0.33	-0.31	-0.38	-0.31	0.07	-0.15	0.03	-0.14
	(-1.31)	(-1.08)	(-1.18)	(-1.01)	(-1.34)	(-1.15)	(0.25)	(-0.60)	(0.15)	(-0.64)
$< 400, 20 >$	-0.34	-0.29	-0.31	-0.31	-0.35	-0.26	0.07	-0.11	0.05	-0.12
	(-1.24)	(-1.03)	(-1.11)	(-1.05)	(-1.26)	(-0.99)	(0.26)	(-0.45)	(0.24)	(-0.62)
$< 400, 40 >$	-0.39	-0.36	-0.36	-0.41	-0.39	-0.33	-0.02	-0.15	-0.03	-0.17
	(-1.51)	(-1.34)	(-1.39)	(-1.44)	(-1.51)	(-1.30)	(-0.07)	(-0.63)	(-0.17)	(-0.90)
Average	-0.31	-0.25	-0.26	-0.27	-0.32	-0.18	0.05	-0.14	0.05	-0.09
	(-1.13)	(-0.88)	(-0.96)	(-0.92)	(-1.16)	(-0.69)	(0.16)	(-0.58)	(0.19)	(-0.46)

Table A.5
Risk-adjusted Returns under GP: Alternative Model Average Number

The table reports the risk-adjusted returns of the spread portfolios generated by other methods under the CAPM augmented by the GPF factor. Newey and West (1987) robust t-statistics are reported in parentheses. We take the average of the top M (M=5) GP models with the highest training sample Sharpe ratio as the final GP model. We report the results for $M = 3$, 5, and 10. The GP search depth parameter pair $<Pop, Gen>$ is determined by the validation sample and is $<200, 40>$. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

	Ridge	LASSO	ENET	PCR	PLS	NN1	NN2	NN3	NN4	NN5
M=3	-0.44	-0.35	-0.36	-0.37	-0.46	-0.04	-0.17	-0.30	-0.11	-0.19
	(-1.60)	(-1.19)	(-1.24)	(-1.26)	(-1.62)	(-0.15)	(-0.60)	(-1.22)	(-0.45)	(-0.94)
M=5	-0.45	-0.35	-0.35	-0.38	-0.46	-0.05	-0.16	-0.30	-0.10	-0.18
	(-1.64)	(-1.20)	(-1.24)	(-1.29)	(-1.65)	(-0.18)	(-0.57)	(-1.25)	(-0.43)	(-0.89)
M=10	-0.32	-0.25	-0.26	-0.25	-0.33	-0.13	-0.00	-0.17	-0.06	-0.15
	(-1.13)	(-0.83)	(-0.90)	(-0.80)	(-1.16)	(-0.44)	(-0.03)	(-0.74)	(-0.28)	(-0.73)
Average	-0.40	-0.32	-0.32	-0.33	-0.42	-0.07	-0.11	-0.26	-0.09	-0.17
	(-1.46)	(-1.07)	(-1.13)	(-1.12)	(-1.48)	(-0.26)	(-0.40)	(-1.07)	(-0.39)	(-0.85)

Table A.6
Risk-adjusted Returns of GP: Alternative Search Depth

The table reports the risk-adjusted returns of the GP spread portfolios under existing factor models, including the CAPM, the Fama and French (1993) 3-factor model (FF-3), Fama and French (2015) 5-factor model (FF-5), Hou, Xue, and Zhang (2015) 4-factor model (HXZ-4), Stambaugh and Yuan (2016) mispricing-factor model (SY-4), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model model (DHS-3). Newey and West (1987) robust t-statistics are reported in parentheses. We report the results under various GP search depth parameters of $<Pop, Gen>$. We take the average of the top M (M=5) GP models with the highest training sample Sharpe ratio as the final GP model. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

	CAPM	FF-3	FF-5	HXZ-4	DHS-3	SY-4
<100, 10>	1.49***	1.54***	1.64***	1.39***	1.46***	1.35***
	(5.40)	(5.51)	(4.81)	(4.60)	(6.11)	(4.55)
<100, 20>	1.38***	1.39***	1.38***	1.38***	1.30***	1.49***
	(5.25)	(5.18)	(5.04)	(4.79)	(5.05)	(4.92)
<100, 40>	1.34***	1.34***	1.27***	1.03***	1.09***	0.91***
	(4.49)	(4.59)	(3.78)	(3.25)	(3.81)	(2.69)
<200, 10>	1.06***	1.11***	1.20***	1.04***	1.17***	1.00***
	(3.81)	(3.90)	(3.68)	(3.16)	(3.98)	(2.70)
<200, 20>	1.17***	1.23***	1.29***	1.08***	1.28***	1.07***
	(4.62)	(4.73)	(4.59)	(4.13)	(4.63)	(3.49)
<200, 40>	1.32***	1.33***	1.49***	1.39***	1.43***	1.43***
	(5.95)	(5.74)	(5.27)	(5.42)	(6.63)	(5.83)
<400, 10>	1.48***	1.52***	1.53***	1.34***	1.42***	1.25***
	(5.20)	(5.43)	(5.01)	(4.42)	(5.56)	(3.73)
<400, 20>	1.50***	1.55***	1.66***	1.40***	1.48***	1.33***
	(5.57)	(5.80)	(5.07)	(4.54)	(5.89)	(4.05)
<400, 40>	1.08***	1.11***	1.26***	1.06***	1.25***	1.15***
	(4.27)	(4.28)	(4.35)	(3.50)	(5.02)	(3.28)
Average	1.31***	1.35***	1.41***	1.23***	1.32***	1.22***
	(4.95)	(5.02)	(4.62)	(4.20)	(5.19)	(3.92)

Table A.7
Risk-adjusted Returns of GP: Alternative Model Average Number

The table reports the risk-adjusted returns of the GP spread portfolios under existing factor models, including the CAPM, the Fama and French (1993) 3-factor model (FF-3), Fama and French (2015) 5-factor model (FF-5), Hou, Xue, and Zhang (2015) 4-factor model (HXZ-4), Stambaugh and Yuan (2016) mispricing-factor model (SY-4), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model model (DHS-3). Newey and West (1987) robust t-statistics are reported in parentheses. We take the average of the top M (M=5) GP models with the highest training sample Sharpe ratio as the final GP model. We report the results for $M = 3$, 5, and 10. The GP search depth parameter pair $<Pop, Gen>$ is determined by the validation sample and is $<200, 40>$. The last row "Average" reports the average statistics across the parameters. The sample period is from 1991:01 to 2021:12.

	CAPM	FF-3	FF-5	HXZ-4	DHS-3	SY-4
M=3	1.33***	1.35***	1.50***	1.44***	1.46***	1.48***
	(5.68)	(5.50)	(5.05)	(5.28)	(6.45)	(5.62)
M=5	1.32***	1.33***	1.49***	1.39***	1.43***	1.43***
	(5.95)	(5.74)	(5.27)	(5.42)	(6.63)	(5.83)
M=10	1.27***	1.30***	1.37***	1.33***	1.34***	1.31***
	(5.11)	(5.00)	(4.78)	(4.62)	(5.66)	(4.64)
Average	1.31***	1.33***	1.45***	1.39***	1.41***	1.41***
	(5.58)	(5.41)	(5.03)	(5.11)	(6.25)	(5.36)