Factor Mimicking Portfolios

Bernt Arne Ødegaard

24 November 2021

Contents

1 Factor Mimicking Portfolios

In this lecture we are concerned with a useful standard method in empirical finance, replacing some variable with a function of a bunch of other variables.

More specifically, we consider the cases where some variable of interest can be written as a function of a number of *portfolios*.

This is typically because we want to use data about tradeable assets to proxy for some other economic variable that is not observable.

1.1 Economic Tracking Portfolios

A good reference for this type of analysis is Lamont (2001), which explains the idea: Construct, from financial assets traded often, a "matching portfolio" of some economic factor that one wants an estimate of.

Say one want current estimates of GDP or Inflation. Construct the portfolio of financial variables (e.g. industry portfolios, that most closely matches the time series evolution of the macro variable. Use the most recent estimates of stock returns to predict the macro variable.

Note: Lehmann and Modest (1988) has some of the same ideas in the context of the APT.

1.2 Example

We will not go into detail about this, rather illustrate the idea with a simple example.

Consider a value weighted market portfolio for the stocks at the Oslo Stock Exchange.

This is constructed as a sum of returns on individual assets times the market weight of each asset.

What if we don't have the market return, all we have is returns of a bunch of industry portfolios?[1]

We can estimate the market return as a weighted average of the industry portfolios. In this case we know the industry weights, they are shown in Table 1.

But let us ignore that we know this. We can estimate the returns on the market portfolio from a regression of market portfolio returns on the returns on the sector portfolios.

$$r_{tm} = a + \sum_k b_k r_{kt} + \varepsilon_t$$

[1]Of course somewhat unrealistic, usually have both market and industries, but ignore that.

Table 1 The fraction of market values in the different GICS Industry Sectors

Panel A: Subperiod 1980–1989

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Energy and consumption	10.80	9.50	8.46	8.77	8.93	8.17	7.01	10.02	10.42	16.58
Material/labor	8.86	8.95	8.25	10.10	10.81	11.12	11.11	11.75	10.48	12.03
Industrials	57.95	50.83	39.25	36.68	32.59	32.98	34.42	32.95	42.70	40.36
Consumer Discretionary	1.01	1.53	3.19	2.38	3.53	5.39	7.53	6.27	5.30	3.39
Consumer Staples	2.30	4.75	5.50	5.02	6.87	6.47	9.94	11.42	7.79	8.50
Health Care/liability	1.13	1.23	2.34	3.43	3.31	4.45	3.63	5.91	9.34	5.67
Financials	18.29	23.89	27.13	21.40	21.80	20.98	23.65	24.97	14.48	16.56
Information Technology	0.81	3.73	5.96	12.23	12.15	10.53	10.13	5.27	2.74	1.85
Telecommunication Services	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.18	0.10	0.00
Utilities	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: Subperiod 1990–1999

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Energy and consumption	21.72	22.63	19.11	17.80	15.76	16.37	23.57	24.57	15.46	15.37
Material/labor	8.14	6.49	6.04	8.02	8.13	6.66	4.57	2.95	3.69	5.17
Industrials	39.32	40.16	40.23	36.94	40.88	37.95	35.37	27.83	27.40	28.05
Consumer Discretionary	2.63	2.15	4.91	5.82	5.24	5.01	5.70	9.33	15.09	16.45
Consumer Staples	10.45	11.53	15.32	11.81	6.41	6.49	6.84	6.28	6.39	6.23
Health Care/liability	6.58	11.19	12.29	5.69	5.36	6.15	2.59	8.37	14.20	5.61
Financials	16.40	8.54	8.52	16.54	18.16	20.34	17.36	14.70	18.39	17.16
Information Technology	1.81	1.50	1.76	1.98	1.35	3.59	4.62	4.69	5.38	11.27
Telecommunication Services	0.00	0.00	0.00	0.00	0.00	0.00	0.70	1.33	2.04	3.74
Utilities	0.00	0.00	0.00	0.00	0.00	0.00	1.28	0.75	1.04	0.76

Panel C: Subperiod 2000–2009

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Energy and consumption	9.84	25.94	42.41	42.58	42.96	51.70	51.13	50.39	60.86	52.38
Material/labor	5.25	4.72	3.94	3.59	3.23	2.70	1.72	1.28	0.59	0.57
Industrials	27.17	26.36	9.49	6.85	10.32	11.26	10.49	10.84	11.99	11.51
Consumer Discretionary	10.04	5.78	6.47	8.31	9.07	6.17	4.48	3.46	3.24	3.82
Consumer Staples	7.92	6.73	7.04	5.09	5.62	5.77	6.36	7.01	6.52	6.62
Health Care/liability	7.49	7.94	8.01	8.56	7.59	0.38	0.38	0.70	1.17	0.97
Financials	17.20	14.51	14.91	12.81	13.90	11.21	10.32	9.20	6.48	10.71
Information Technology	10.62	7.17	3.69	4.30	4.40	3.87	6.20	8.58	6.32	4.95
Telecommunication Services	13.26	9.20	8.54	9.94	9.65	7.63	9.91	9.63	7.66	8.88
Utilities	0.60	0.54	0.91	1.03	0.92	1.10	1.33	1.54	1.72	1.14

Panel C: Subperiod 2000–2014

	2010	2011	2012	2013	2014
Energy and consumption	44.04	51.42	46.62	42.28	32.85
Material/labor	1.43	1.14	0.56	0.29	0.44
Industrials	12.72	12.16	13.14	11.22	11.81
Consumer Discretionary	6.52	4.92	5.94	6.40	11.96
Consumer Staples	7.49	5.56	6.66	6.58	8.55
Health Care/liability	0.95	1.16	1.05	1.38	1.16
Financials	14.15	12.92	14.41	17.35	18.45
Information Technology	4.27	2.15	1.95	2.92	3.31
Telecommunication Services	9.56	11.20	11.75	12.26	12.62
Utilities	1.05	1.07	0.84	0.71	0.91

The tables list, for each year, the percentage fraction of the value of the OSE is in each GICS sector. Measurement done at yearend. Data for the period 1980–2014

If we run this regression without a constant term,

$$r_{tm} = \sum_k b_k r_{kt} + \varepsilon_t$$

it looks very much like a portfolio.

Let us do this regression using data 1980-2013, and see what the weights look like.

Download returns for eight norwegian industries (10-45) for 1980-2013.

Similarly download the value weighted portfolio for the same period.

Regress the market on the eight industries.

(This procedure is also termed to *project* the market on the industries)

```
> source ("~/2015_research/empirics_ose/asset_pricing_results/data/read_ose_data.R")
>                                        # only use eight first industries
> IndustryRets <- IndustryRets[,1:8]
> head(IndustryRets)
          Energy10  Material15  Industry20 ConsDisc25 ConsStapl30  Health35
Jan 1980  0.097561  0.01221640  0.02154350  0.0489160  -0.0025490  0.010417
Feb 1980  0.011111  0.07595600  0.04081450  0.1203870   0.0827815  0.025773
Mar 1980 -0.098901 -0.10693300 -0.09349900  0.0128427  -0.0457250 -0.045226
Apr 1980  0.091463  0.02555040 -0.00333583  0.0071930  -0.0174770  0.031579
May 1980  0.131844  0.01895950  0.04345470 -0.0255557   0.0571205  0.102041
Jun 1980 -0.036269  0.00775375 -0.00413267 -0.0219883  -0.0681400 -0.135922
            Finan40      IT45
Jan 1980 -0.0183516  0.400000
Feb 1980  0.0184339  0.178571
Mar 1980 -0.0224918  0.121212
Apr 1980 -0.0055362  0.189189
May 1980  0.0080529  0.636355
Jun 1980  0.0313302 -0.083333
> Ri <- window(IndustryRets,end=as.yearmon("2013-12"))
> Rm <- window(Rmvw,end=as.yearmon("2013-12"))
> tail(Ri)
            Energy10 Material15 Industry20 ConsDisc25 ConsStapl30    Health35
Jul 2013  0.07210540   0.099391  0.0559940  0.0327910 -0.02890040  0.01752880
Aug 2013  0.02612530   0.007380  0.0291508  0.0056002 -0.02740510 -0.01169220
Sep 2013  0.02250170  -0.009158  0.0380361  0.0219844  0.05740760 -0.02395000
Oct 2013  0.00777452   0.044362  0.0018050  0.0190972  0.07184570 -0.01328500
Nov 2013  0.03342590  -0.086726  0.0530893  0.0214134  0.01691990  0.07412920
Dec 2013 -0.00167291  -0.042636  0.0263073  0.0377218  0.00864471 -0.00624583
             Finan40      IT45
Jul 2013  0.04213890 0.0655168
Aug 2013 -0.00436296 0.0986375
Sep 2013  0.00506644 0.0299650
Oct 2013  0.05224100 0.0506128
Nov 2013  0.01714770 0.0538205
Dec 2013  0.01147980 0.0218707
>
> regr <- lm(Rm ~
+            0
+            + Ri$Energy10
+            + Ri$Material15
+            + Ri$Industry20
```

```
+            + Ri$ConsDisc25
+            + Ri$ConsStapl30
+            + Ri$Health35
+            + Ri$Finan40
+            + Ri$IT45 )
> summary(regr)

Call:
lm(formula = Rm ~ 0 + Ri$Energy10 + Ri$Material15 + Ri$Industry20 +
    Ri$ConsDisc25 + Ri$ConsStapl30 + Ri$Health35 + Ri$Finan40 +
    Ri$IT45)

Residuals:
      Min        1Q    Median        3Q       Max
-0.158816 -0.013361  0.002608  0.020336  0.095469

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
Ri$Energy10    0.170445   0.026346   6.469 2.88e-10 ***
Ri$Material15  0.043443   0.016116   2.696  0.00732 **
Ri$Industry20  0.297956   0.044703   6.665 8.78e-11 ***
Ri$ConsDisc25  0.039237   0.030109   1.303  0.19327
Ri$ConsStapl30 0.177604   0.031486   5.641 3.21e-08 ***
Ri$Health35    0.098286   0.020121   4.885 1.50e-06 ***
Ri$Finan40     0.148100   0.046699   3.171  0.00163 **
Ri$IT45        0.004812   0.018245   0.264  0.79212
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0308 on 400 degrees of freedom
Multiple R-squared:  0.7874,Adjusted R-squared:  0.7831
F-statistic: 185.2 on 8 and 400 DF,  p-value: < 2.2e-16
```

	Dependent variable:
	Rm
Energy10	0.170^{***}
	(0.026)
Material15	0.043^{***}
	(0.016)
Industry20	0.298^{***}
	(0.045)
ConsDisc25	0.039
	(0.030)
ConsStapl30	0.178^{***}
	(0.031)
Health35	0.098^{***}
	(0.020)
Finan40	0.148^{***}
	(0.047)
IT45	0.005
	(0.018)
Observations	408
R^2	0.787
Adjusted R^2	0.783
Residual Std. Error	0.031 (df = 400)
F Statistic	185.171^{***} (df = 8; 400)
Note:	*p<0.1; **p<0.05; ***p<0.01

If this was a portfolio, the weight should sum to one. Let us look at how close we get:

```
 > sum(coefficients(regr))
[1] 0.9798838
```

Now, to the typical usage of this kind of procedure:
Prediction into the future.
Download the industry returns for 2014.
Use the estimated relationship to predict the return to the value weighted market portfolio.
Compare your prediction with the actual market returns.

```
>                                           #now look at prediction for 2014
> Ri <- window(IndustryRets,start=as.yearmon("2014-01"))
> rm <- window(Rmvw,start=as.yearmon("2014-01"))
> Rmpred <- predict.lm(regr,Ri)
> data <- merge(rm,Rmpred)
> print(data)
               rm       Rmpred
```

```
Jan 2014 -0.017441  0.024426508
Feb 2014  0.031665  0.011566819
Mar 2014  0.015775  0.012166395
Apr 2014  0.029255  0.010763492
May 2014  0.049615  0.033303480
Jun 2014  0.025168  0.023292721
Jul 2014 -0.001637  0.016756078
Aug 2014 -0.001261 -0.016250066
Sep 2014  0.003305  0.003561559
Oct 2014 -0.032450 -0.015193497
Nov 2014 -0.026977 -0.017774517
Dec 2014  0.028808  0.028986437
```

	R_m (actual)	R_m (predicted)
2014 Jan	-0.0174	0.0244
2014 Feb	0.0317	0.0116
2014 Mar	0.0158	0.0122
2014 Apr	0.0293	0.0108
2014 May	0.0496	0.0333
2014 Jun	0.0252	0.0233
2014 Jul	-0.0016	0.0168
2014 Aug	-0.0013	-0.0163
2014 Sep	0.0033	0.0036
2014 Oct	-0.0324	-0.0152
2014 Nov	-0.0270	-0.0178
2014 Dec	0.0288	0.0290

This kind of procedure is often called construction of "factor mimicking" portfolios.

In the example the "factor" we are constructing is the value weighted market portfolio.

This type of procedure obviously extends to non-traded "factors," and that is the usage one typically runs into it.

2 Readings

Lamont (2001) looks at this in a setting closer related to forecasting.

References

Owen A Lamont. Economic tracking portfolios. *Journal of Econometrics*, 105:161–184, 2001.

B N Lehmann and David M Modest. The empirical foundations of the Arbitrage Pricing Theory. *Journal of Financial Economics*, 21:213–254, 1988.