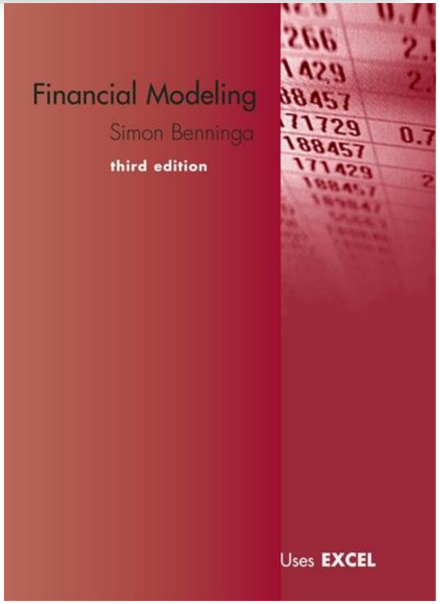
Chapter 10: Computing the variance-covariance matrix

Th. Warin



Basic example: 6 stocks, annual data

	A	B	C	D	E	F	G	H
1	**ANNUAL STOCK PRICE AND RETURN DATA FOR SIX STOCKS** General Electric (GE), Microsoft (MSFT), Johnson & Johnson (JNJ), Kellogg (K), Boeing (BA), IBM							
2	**Price data**							
3	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
4	4-Jan-93	2.36	2.68	6.78	20.37	2.34	11.79	
5	3-Jan-94	4.15	2.64	7.20	18.47	4.21	14.62	
6	3-Jan-95	4.98	3.68	10.91	19.90	4.20	15.53	
7	2-Jan-96	8.80	5.73	19.43	29.03	8.09	20.41	
8	2-Jan-97	13.51	12.64	24.44	27.59	13.93	30.78	
9	2-Jan-98	21.64	18.49	29.15	38.01	20.19	31.60	
10	4-Jan-99	30.57	43.37	38.04	34.14	23.47	30.94	
11	3-Jan-00	40.51	48.51	39.36	20.93	36.27	39.24	
12	2-Jan-01	42.42	30.26	43.80	23.52	48.13	48.78	
13	2-Jan-02	34.82	31.58	55.19	28.70	41.39	51.05	
14	2-Jan-03	22.25	23.52	52.15	32.00	32.81	59.63	
15	2-Jan-04	31.86	28.16	51.49	37.36	48.86	81.95	
16								
17	Shares outstanding	10.56	10.86	2.97	0.41	0.84	0.79	
18	Market value	336.44	305.82	152.93	15.44	41.01	65.13	<-- =G15*G17
19	Percentage of portfolio	36.70%	33.36%	16.68%	1.68%	4.47%	7.10%	<-- =G18/SUM($B$18:$K$18)

We usually use monthly or weekly data, but annual is convenient for visualizing.

From prices to returns

	A	B	C	D	E	F	G	H
21	**Return data**							
22	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
23	3-Jan-94	56.44%	-1.50%	6.01%	-9.79%	58.73%	21.51%	<-- =LN(G5/G4)
24	3-Jan-95	18.23%	33.21%	41.56%	7.46%	-0.24%	6.04%	<-- =LN(G6/G5)
25	2-Jan-96	56.93%	44.28%	57.71%	37.76%	65.55%	27.33%	
26	2-Jan-97	42.87%	79.12%	22.94%	-5.09%	54.34%	41.08%	
27	2-Jan-98	47.11%	38.04%	17.62%	32.04%	37.11%	2.63%	
28	4-Jan-99	34.55%	85.25%	26.62%	-10.74%	15.05%	-2.11%	
29	3-Jan-00	28.15%	11.20%	3.41%	-48.93%	43.53%	23.76%	
30	2-Jan-01	4.61%	-47.19%	10.69%	11.67%	28.29%	21.76%	
31	2-Jan-02	-19.74%	4.27%	23.11%	19.90%	-15.09%	4.55%	
32	2-Jan-03	-44.78%	-29.47%	-5.67%	10.88%	-23.23%	15.54%	
33	2-Jan-04	35.90%	18.01%	-1.27%	15.49%	39.82%	31.80%	
34								
35	Average	23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	<-- =AVERAGE(G23:G33)
36	Standard deviation	32.17%	40.71%	18.97%	23.86%	29.93%	13.56%	<-- =STDEV(G23:G33)
37	Variance	0.1035	0.1657	0.0360	0.0570	0.0896	0.0184	<-- =VAR(G23:G33)

Matrix of excess returns

	A	B	C	D	E	F	G	H
35	Average	23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	<-- =AVERAGE(G23:G33)
36	Standard deviation	32.17%	40.71%	18.97%	23.86%	29.93%	13.56%	<-- =STDEV(G23:G33)
37	Variance	0.1035	0.1657	0.0360	0.0570	0.0896	0.0184	<-- =VAR(G23:G33)
38								
39								
40	**Excess returns: Returns minus means**							
41	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
42	3-Jan-94	32.78%	-22.89%	-12.42%	-15.31%	31.11%	3.89%	<-- =G23-G$35
43	3-Jan-95	-5.43%	11.83%	23.13%	1.94%	-27.86%	-11.59%	<-- =G24-G$35
44	2-Jan-96	33.27%	22.90%	39.28%	32.25%	37.93%	9.70%	
45	2-Jan-97	19.21%	57.73%	4.51%	-10.60%	26.72%	23.46%	
46	2-Jan-98	23.45%	16.65%	-0.81%	26.53%	9.49%	-15.00%	
47	4-Jan-99	10.89%	63.87%	8.19%	-16.25%	-12.57%	-19.74%	
48	3-Jan-00	4.49%	-10.18%	-15.02%	-54.44%	15.90%	6.14%	
49	2-Jan-01	-19.05%	-68.58%	-7.74%	6.15%	0.67%	4.14%	
50	2-Jan-02	-43.40%	-17.11%	4.68%	14.39%	-42.71%	-13.08%	
51	2-Jan-03	-68.45%	-50.85%	-24.10%	5.37%	-50.86%	-2.09%	
52	2-Jan-04	12.24%	-3.38%	-19.70%	9.97%	12.20%	14.17%	

The var-cov matrix is based on the matrix of returns minus minus means.

$$\text{var-cov} = \frac{\left[r_{it} - \bar{r}_i\right]^T * \left[r_{it} - \bar{r}_i\right]}{M}$$

Variance-covariance matrix

	A	B	C	D	E	F	G
54	**Uses the array formula {<-- {=MMULT(TRANSPOSE(B42:G52),B42:G52)/10}} to compute the sample var-cov matrix**						
55		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**
56	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123
57	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022
58	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039
59	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046
60	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248
61	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184
62							
63	**Note: To put the array formula into cells B56:G61:**						
64	1. Mark the whole area B56:G61						
65	2. Type <-- {=MMULT(TRANSPOSE(B42:G52),B42:G52)/10} into one of the cells.						
66	3. When finished typing, hit [Ctrl]+[Shift]+[Enter] to put in the formula as an array formula.						

Note the use of the array function to compute the var-cov matrix.

In one step

	A	B	C	D	E	F	G
68	**In one step**						
69	**Uses the array formula <-- {=MMULT(TRANSPOSE(B23:G33-B35:G35),B23:G33-B35:G35)/10} to compute the sample var-cov matrix**						
70		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**
71	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123
72	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022
73	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039
74	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046
75	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248
76	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184

Goes directly from returns & means to variance-covariance.

Note the writing: B23:G33-B35:G35 which creates the excess returns (works only as array function).

Computing var-cov in one step

	A	B	C	D	E	F	G
68	**In one step**						
69	**Uses the array formula** **<-- {=MMULT(TRANSPOSE(B23:G33-B35:G35),B23:G33-B35:G35)/10}** **to compute the sample var-cov matrix**						
70		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**
71	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123
72	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022
73	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039
74	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046
75	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248
76	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184

Writing B23:G33-B35:G35 as an array function subtracts B35 from each of B23:B33, C35 from each of C23:C33, … , thus creating the matrix of excess returns.

	A	B	C	D	E	F	G	H
1	**SOME CONFUSION ABOUT M VERSUS M-1 IN EXCEL?**							
2	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
3	3-Jan-94	56.44%	-1.50%	6.01%	-9.79%	58.73%	21.51%	
4	3-Jan-95	18.23%	33.21%	41.56%	7.46%	-0.24%	6.04%	
5	2-Jan-96	56.93%	44.28%	57.71%	37.76%	65.55%	27.33%	
6	2-Jan-97	42.87%	79.12%	22.94%	-5.09%	54.34%	41.08%	
7	2-Jan-98	47.11%	38.04%	17.62%	32.04%	37.11%	2.63%	
8	4-Jan-99	34.55%	85.25%	26.62%	-10.74%	15.05%	-2.11%	
9	3-Jan-00	28.15%	11.20%	3.41%	-48.93%	43.53%	23.76%	
10	2-Jan-01	4.61%	-47.19%	10.69%	11.67%	28.29%	21.76%	
11	2-Jan-02	-19.74%	4.27%	23.11%	19.90%	-15.09%	4.55%	
12	2-Jan-03	-44.78%	-29.47%	-5.67%	10.88%	-23.23%	15.54%	
13	2-Jan-04	35.90%	18.01%	-1.27%	15.49%	39.82%	31.80%	
14								
15	**Mean**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
16		23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	<-- =AVERAGE(G3:G13)
17		23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	<-- =SUM(G3:G13)/COUNT(G3:G13)
18								
19	**Variance**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
20		0.0941	0.1507	0.0327	0.0518	0.0814	0.0167	<-- =COVAR(G3:G13,G3:G13)
21		0.0941	0.1507	0.0327	0.0518	0.0814	0.0167	<-- =VARP(G3:G13)
22		0.1035	0.1657	0.0360	0.0570	0.0896	0.0184	<-- =VAR(G3:G13)
23								
24	**Standard deviation**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
25		0.3067	0.3882	0.1808	0.2275	0.2854	0.1293	<-- =SQRT(G20)
26		0.3067	0.3882	0.1808	0.2275	0.2854	0.1293	<-- =STDEVP(G3:G13)
27		0.3217	0.4071	0.1897	0.2386	0.2993	0.1356	<-- =STDEV(G3:G13)
28								
29	**Covariance(GE,MSFT)**							
30		0.0690	<-- =COVAR(B3:B13,C3:C13)					
31		0.0690	<-- {=MMULT(TRANSPOSE(B3:B13-B16),C3:C13-C16)/11}					
32		0.0758	<-- {=MMULT(TRANSPOSE(B3:B13-B16),C3:C13-C16)/10}					
33		0.0758	<-- =COVAR(B3:B13,C3:C13)*11/10					

In *Financial Modeling*, *N* = number of assets, *M* = number of observations

The Excel function **Covar** is based on dividing by M, not M-1

M versus M-1?? Who cares?

"There is a long story about why the denominator is *M-1* instead of *M*. If you have never heard that story, you may consult any good statistics text. Here we will be content to note that the *M-1 should* be changed to *M* if you are ever in the situation of measuring the variance of a distribution whose mean is known *a priori* rather than being estimated from the data. (We might also comment that if the difference between *M* and *M-1* ever matters to you, then you are <u>probably up to no good anyway</u>—e.g., trying to substantiate a questionable hypothesis with marginal data.)"

Source: *Numerical Recipes: The Art of Scientific Computing*, by William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, Cambridge University Press, 1986.

VBA function varcov (I)

```
'This is Amir Kirsh's variance-covariance function
'Benjamin Czaczkes made an improvement in 2007

Function VarCovar(rng As Range) As Variant
    Dim i As Integer
    Dim j As Integer
    Dim numCols As Integer
    numCols = rng.Columns.Count
    numRows = rng.Rows.Count
    Dim matrix() As Double
    ReDim matrix(numCols - 1, numCols - 1)

    For i = 1 To numCols
        For j = 1 To numCols
            matrix(i - 1, j - 1) = _
         Application.WorksheetFunction.Covar(rng.Columns(i), _
         rng.Columns(j)) _
            * numRows / (numRows - 1)
        Next j
    Next i
    VarCovar = matrix
End Function
```

VBA function varcov (2)

	A	B	C	D	E	F	G	H
1	**USING A VBA FUNCTION TO COMPUTE THE COVARIANCE MATRIX** General Electric (GE), Microsoft (MSFT), Johnson & Johnson (JNJ), Kellogg (K), Boeing (BA), IBM							
21	**Return data**							
22	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
23	3-Jan-94	56.44%	-1.50%	6.01%	-9.79%	58.73%	21.51%	<-- =LN(G5/G4)
24	3-Jan-95	18.23%	33.21%	41.56%	7.46%	-0.24%	6.04%	<-- =LN(G6/G5)
25	2-Jan-96	56.93%	44.28%	57.71%	37.76%	65.55%	27.33%	
26	2-Jan-97	42.87%	79.12%	22.94%	-5.09%	54.34%	41.08%	
27	2-Jan-98	47.11%	38.04%	17.62%	32.04%	37.11%	2.63%	
28	4-Jan-99	34.55%	85.25%	26.62%	-10.74%	15.05%	-2.11%	
29	3-Jan-00	28.15%	11.20%	3.41%	-48.93%	43.53%	23.76%	
30	2-Jan-01	4.61%	-47.19%	10.69%	11.67%	28.29%	21.76%	
31	2-Jan-02	-19.74%	4.27%	23.11%	19.90%	-15.09%	4.55%	
32	2-Jan-03	-44.78%	-29.47%	-5.67%	10.88%	-23.23%	15.54%	
33	2-Jan-04	35.90%	18.01%	-1.27%	15.49%	39.82%	31.80%	
34								
38								
69	**Uses the homemade array formula {=varcovar(B23:G33)} to compute the sample var-cov matrix**							
70		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
71	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123	
72	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022	
73	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039	
74	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046	
75	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248	
76	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184	

Global minimum variance portfolio

	A	B	C	D	E	F	G	H	I
1	**COMPUTING THE GLOBAL MINIMUM VARIANCE PORTFOLIO USING THE SAMPLE VARIANCE-COVARIANCE MATRIX**								
2	**Sample variance-covariance matrix**								
3		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**		**One**
4	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123		1
5	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022		1
6	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039		1
7	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046		1
8	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248		1
9	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184		1
10									
11	**Mean**	23.66%	21.38%	18.43%	5.51%	27.63%	17.63%		
12									
13	**The global minimum variance portfolio (GMVP) is computed below with the formula {=MMULT(TRANSPOSE(I4:I9),MINVERSE(B4:G9))/MMULT(MMULT(TRANSPOSE(I4:I9),MINVERSE(B4:G9)),I4:I9)}**								
14		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**		
15		0.6105	-0.1034	0.2074	0.0539	-0.7704	1.0019		
16									
17	**Sum**	1.0000	<-- =SUM(B15:G15)						
18									
19	**GMVP mean**	0.1273	<-- =SUMPRODUCT(B11:G11,B15:G15)						
20	**GMVP variance**	0.0060	<-- {=MMULT(MMULT(B15:G15,B4:G9),TRANSPOSE(B15:G15))}						
21	**GMVP standard deviation**	0.0773	<-- =SQRT(B20)						

$$GMVP = \frac{\{1,\ldots,1\}\, S^{-1}}{\{1,\ldots,1\}\, S^{-1} \begin{Bmatrix} 1 \\ \vdots \\ 1 \end{Bmatrix}}$$

Correlation matrix

- A simple variation of the **varcovar** macro will give the correlation matrix:

```
'Adaptation of variance-covariance function in FM3
Function CorrMatrix(rng As Range) As Variant
    Dim i As Integer
    Dim j As Integer
    Dim numCols As Integer
    numCols = rng.Columns.Count
    numRows = rng.Rows.Count
    Dim matrix() As Double
    ReDim matrix(numCols - 1, numCols - 1)

    For i = 1 To numCols
        For j = 1 To numCols
        matrix(i - 1, j - 1) = _
         Application.WorksheetFunction.Correl(rng.Columns(i), _
              rng.Columns(j))
        Next j
    Next i
    CorrMatrix = matrix
End Function
```

Envelope portfolio

By Proposition 1 of Chapter 9, an envelope portfolio *x* is the solution of the equation:

$$x = \frac{\Sigma^{-1} * \left\{ E\left(\tilde{r}\right) - c \right\}}{\sum \Sigma^{-1} * \left\{ E\left(\tilde{r}\right) - c \right\}}$$

c is an arbitrary constant.

Note: I often blur the distinction between *envelope*_ and *efficient* portfolios. This is incorrect but convenient. ("Envelope" is, of course, correct.)

Efficient portfolio

	A	B	C	D	E	F	G	H	I
1	**COMPUTING AN EFFICIENT PORTFOLIO**								
2		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**		**Means**
3	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.0123		23.66%
4	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	-0.0022		21.38%
5	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	-0.0039		18.43%
6	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	-0.0046		5.51%
7	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0248		27.63%
8	**IBM**	0.0123	-0.0022	-0.0039	-0.0046	0.0248	0.0184		17.63%
9									
10	Risk-free rate	2%							
11									
12	**The efficient portfolio is computed by the array formula {=TRANSPOSE(MMULT(MINVERSE(B3:G8),I3:I8-B10)/SUM(MMULT(MINVERSE(B3:G8),I3:I8-B10)))}. The cells below use TRANSPOSE( ) to make this a row vector.**								
13		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**		
14	Efficient portfolio	26.37%	-6.05%	36.98%	-4.81%	-33.87%	81.39%		
15									
16	Market value ($billion)	336.44	305.82	152.93	15.44	41.01	16.98		
17	Market proportions	38.73%	35.21%	17.61%	1.78%	4.72%	1.96%	<-- =G16/SUM($B$16:$K$16)	

Note that the efficient portfolio is very different from the market proportions!

In our example: All envelope portfolios contain negative stock positions

	A	B	C	D	E	F	G	H	I
20	**Data table: Optimal portfolio for various c's: the table header (hidden) in row 22 refers to the computation of the efficient portfolio in row 14**								
21		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	**Largest long position**	**Smallest position**
22	c								
23	-10%	44.68%	-8.31%	28.41%	0.57%	-56.66%	91.32%	91.32%	-56.66%
24	-9%	43.92%	-8.22%	28.76%	0.35%	-55.72%	90.91%	90.91%	-55.72%
25	-8%	43.10%	-8.12%	29.15%	0.11%	-54.69%	90.46%	90.46%	-54.69%
26	-7%	42.19%	-8.01%	29.57%	-0.16%	-53.56%	89.97%	89.97%	-53.56%
27	-6%	41.18%	-7.88%	30.04%	-0.46%	-52.31%	89.42%	89.42%	-52.31%
28	-5%	40.06%	-7.74%	30.57%	-0.79%	-50.91%	88.81%	88.81%	-50.91%
29	-4%	38.80%	-7.59%	31.16%	-1.15%	-49.35%	88.13%	88.13%	-49.35%
30	-3%	37.39%	-7.41%	31.82%	-1.57%	-47.59%	87.37%	87.37%	-47.59%
31	-2%	35.78%	-7.21%	32.57%	-2.04%	-45.59%	86.50%	86.50%	-45.59%
32	-1%	33.94%	-6.98%	33.43%	-2.58%	-43.30%	85.50%	85.50%	-43.30%
33	0%	31.81%	-6.72%	34.43%	-3.21%	-40.65%	84.34%	84.34%	-40.65%
34	1%	29.32%	-6.41%	35.59%	-3.94%	-37.55%	82.99%	82.99%	-37.55%
35	2%	26.37%	-6.05%	36.98%	-4.81%	-33.87%	81.39%	81.39%	-33.87%
36	3%	22.80%	-5.61%	38.65%	-5.86%	-29.44%	79.46%	79.46%	-29.44%
37	4%	18.42%	-5.06%	40.70%	-7.15%	-23.99%	77.08%	77.08%	-23.99%
38	5%	12.91%	-4.38%	43.28%	-8.77%	-17.13%	74.10%	74.10%	-17.13%
39	7%	-3.89%	-2.30%	51.14%	-13.71%	3.77%	64.99%	64.99%	-13.71%
40	9%	-38.68%	2.00%	67.42%	-23.94%	47.07%	46.13%	67.42%	-38.68%
41	11%	-153.79%	16.25%	121.30%	-57.80%	190.31%	-16.27%	190.31%	-153.79%

The computation in rows 22:41 is a data table in Excel. The table header, in row 22, is hidden.

4 ALTERNATIVE VAR-COVAR MODELS

Four alternative var-cov models

1. Sample variance-covariance (illustrated thus far)
2. Single-index model (SIM)
3. Constant correlation model
4. Shrinkage models

Single-index model (1)

	A	B	C	D	E	F	G	H	I
1	ESTIMATING THE VARIANCE-COVARIANCE MATRIX USING THE SINGLE-INDEX MODEL								
2	**Return data**								
3	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	**SP500**	
4	3-Jan-94	56.44%	-1.50%	6.01%	-9.79%	58.73%	21.51%	-2.35%	
5	3-Jan-95	18.23%	33.21%	41.56%	7.46%	-0.24%	6.04%	30.16%	
6	2-Jan-96	56.93%	44.28%	57.71%	37.76%	65.55%	27.33%	21.19%	
7	2-Jan-97	42.87%	79.12%	22.94%	-5.09%	54.34%	41.08%	22.07%	
8	2-Jan-98	47.11%	38.04%	17.62%	32.04%	37.11%	2.63%	26.65%	
9	4-Jan-99	34.55%	85.25%	26.62%	-10.74%	15.05%	-2.11%	8.59%	
10	3-Jan-00	28.15%	11.20%	3.41%	-48.93%	43.53%	23.76%	-2.06%	
11	2-Jan-01	4.61%	-47.19%	10.69%	11.67%	28.29%	21.76%	-18.95%	
12	2-Jan-02	-19.74%	4.27%	23.11%	19.90%	-15.09%	4.55%	-27.82%	
13	2-Jan-03	-44.78%	-29.47%	-5.67%	10.88%	-23.23%	15.54%	27.91%	
14	2-Jan-04	35.90%	18.01%	-1.27%	15.49%	39.82%	31.80%	4.34%	
15									
16	**Average**	23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	8.16%	<-- =AVERAGE(H4:H14)
17	**Standard deviation**	32.17%	40.71%	18.97%	23.86%	29.93%	13.56%	19.59%	<-- =STDEV(H4:H14)
18	**Variance**	0.1035	0.1657	0.0360	0.0570	0.0896	0.0184	0.0384	<-- =VAR(H4:H14)
19									
20	**Beta**	0.3411	0.9185	0.2598	0.2344	0.1046	0.0186		<-- =SLOPE(G4:G14,$H$4:$H$14)

Single-index model (2)

	A	B	C	D	E	F	G	H
23	**The SIM var-cov matrix uses the array formula {=IF(B24:G24=A25:A30,B18:G18,MMULT(TRANSPOSE(B20:G20),B20:G20)*H18)} to compute the sample var-cov matrix**							
24		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
25	**GE**	0.1035	0.0120	0.0034	0.0031	0.0014	0.0002	
26	**MSFT**	0.0120	0.1657	0.0092	0.0083	0.0037	0.0007	
27	**JNJ**	0.0034	0.0092	0.0360	0.0023	0.0010	0.0002	
28	**K**	0.0031	0.0083	0.0023	0.0570	0.0009	0.0002	
29	**BA**	0.0014	0.0037	0.0010	0.0009	0.0896	0.0001	
30	**IBM**	0.0002	0.0007	0.0002	0.0002	0.0001	0.0184	
31								
32								
33	**Check--doing this another way: In the cells below we use the formula =IF($A36=B$35,B$18,B$20*$H36*$H$18) Row 34 and column H contain the firm betas**							
34		0.3411	0.9185	0.2598	0.2344	0.1046	0.0186	
35		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
36	**GE**	0.1035	0.0120	0.0034	0.0031	0.0014	0.0002	0.3411
37	**MSFT**	0.0120	0.1657	0.0092	0.0083	0.0037	0.0007	0.9185
38	**JNJ**	0.0034	0.0092	0.0360	0.0023	0.0010	0.0002	0.2598
39	**K**	0.0031	0.0083	0.0023	0.0570	0.0009	0.0002	0.2344
40	**BA**	0.0014	0.0037	0.0010	0.0009	0.0896	0.0001	0.1046
41	**IBM**	0.0002	0.0007	0.0002	0.0002	0.0001	0.0184	0.0186

Constant correlation model

- Assumes that all covariances have the form

$$\sigma_{ij} = \rho * \sigma_i \sigma_j$$

- Where ρ is constant (typically ρ between 0.2 and 0.3)

	A	B	C	D	E	F	G	H
1	**ESTIMATING THE VARIANCE-COVARIANCE MATRIX USING THE CONSTANT-CORRELATION APPROACH**							
2	**Return data**							
3	**Date**	**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
4	3-Jan-94	56.44%	-1.50%	6.01%	-9.79%	58.73%	21.51%	
5	3-Jan-95	18.23%	33.21%	41.56%	7.46%	-0.24%	6.04%	
6	2-Jan-96	56.93%	44.28%	57.71%	37.76%	65.55%	27.33%	
7	2-Jan-97	42.87%	79.12%	22.94%	-5.09%	54.34%	41.08%	
8	2-Jan-98	47.11%	38.04%	17.62%	32.04%	37.11%	2.63%	
9	4-Jan-99	34.55%	85.25%	26.62%	-10.74%	15.05%	-2.11%	
10	3-Jan-00	28.15%	11.20%	3.41%	-48.93%	43.53%	23.76%	
11	2-Jan-01	4.61%	-47.19%	10.69%	11.67%	28.29%	21.76%	
12	2-Jan-02	-19.74%	4.27%	23.11%	19.90%	-15.09%	4.55%	
13	2-Jan-03	-44.78%	-29.47%	-5.67%	10.88%	-23.23%	15.54%	
14	2-Jan-04	35.90%	18.01%	-1.27%	15.49%	39.82%	31.80%	
15								
16	**Average**	23.66%	21.38%	18.43%	5.51%	27.63%	17.63%	<-- =AVERAGE(G4:G14)
17	**Standard deviation**	32.17%	40.71%	18.97%	23.86%	29.93%	13.56%	<-- =STDEV(G4:G14)
18	**Variance**	0.1035	0.1657	0.0360	0.0570	0.0896	0.0184	<-- =VAR(G4:G14)
19	**Average correlation**	0.1999	<-- =AVERAGE(B23:G28)-1/6					
20								
21	**Uses the array formula {=MMULT(TRANSPOSE(B4:G14-B16:G16),B4:G14-B16:G16)/10/MMULT(TRANSPOSE(B17:G17),B17:G17)} to compute the correlations.**							
22		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
23	**GE**	1.0000	0.5791	0.3632	-0.0560	0.8905	0.2819	
24	**MSFT**	0.5791	1.0000	0.5340	-0.0532	0.3113	-0.0406	
25	**JNJ**	0.3632	0.5340	1.0000	0.4002	0.1780	-0.1529	
26	**K**	-0.0560	-0.0532	0.4002	1.0000	-0.1067	-0.1427	
27	**BA**	0.8905	0.3113	0.1780	-0.1067	1.0000	0.6116	
28	**IBM**	0.2819	-0.0406	-0.1529	-0.1427	0.6116	1.0000	

Shrinkage model

- Assumes that variance covariance matrix is convex combination of sample varcov and diagonal (i.e. variance-only) varcov
- Or some plausible alternative

	A	B	C	D	E	F	G	H
20	**Shrinkage factor λ**	0.3	<-- This is the weight put on the sample var-cov					
21								
22	**Shrinkage matrix** **Uses the array formula {=B20*B34:G39+(1-B20)*B44:G49} to compute the shrinkage covariance matrix**							
23		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
24	**GE**	0.1035	0.0228	0.0066	-0.0013	0.0257	0.0424	
25	**MSFT**	0.0228	0.1657	0.0124	-0.0016	0.0114	0.0420	
26	**JNJ**	0.0066	0.0124	0.0360	0.0054	0.0030	0.0137	
27	**K**	-0.0013	-0.0016	0.0054	0.0570	-0.0023	0.0037	
28	**BA**	0.0257	0.0114	0.0030	-0.0023	0.0896	0.0257	
29	**IBM**	0.0424	0.0420	0.0137	0.0037	0.0257	0.2993	
30								
31								
32	**Uses the array formula {=MMULT(TRANSPOSE(B4:G14-B16:G16),B4:G14-B16:G16)/10} to compute the constant sample covariance matrix. In the shrinkage var-cov, this matrix is given weight lambda.**							
33		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
34	**GE**	0.1035	0.0758	0.0222	-0.0043	0.0857	0.1414	
35	**MSFT**	0.0758	0.1657	0.0412	-0.0052	0.0379	0.1400	
36	**JNJ**	0.0222	0.0412	0.0360	0.0181	0.0101	0.0455	
37	**K**	-0.0043	-0.0052	0.0181	0.0570	-0.0076	0.0122	
38	**BA**	0.0857	0.0379	0.0101	-0.0076	0.0896	0.0856	
39	**IBM**	0.1414	0.1400	0.0455	0.0122	0.0856	0.2993	
40								
41								
42	**Uses the array formula {=MMULT(TRANSPOSE(B4:G14-B16:G16),B4:G14-B16:G16)/10*IF(A44:A49=B43:G43,1,0)} to compute a matrix with only variances on diagonal and zeros elsewhere. In the shrinkage var-cov this matrix is given weight 1-lambda.**							
43		**GE**	**MSFT**	**JNJ**	**K**	**BA**	**IBM**	
44	**GE**	0.1035	0.0000	0.0000	0.0000	0.0000	0.0000	
45	**MSFT**	0.0000	0.1657	0.0000	0.0000	0.0000	0.0000	
46	**JNJ**	0.0000	0.0000	0.0360	0.0000	0.0000	0.0000	
47	**K**	0.0000	0.0000	0.0000	0.0570	0.0000	0.0000	
48	**BA**	0.0000	0.0000	0.0000	0.0000	0.0896	0.0000	
49	**IBM**	0.0000	0.0000	0.0000	0.0000	0.0000	0.2993	

Are alternatives plausible?

- Appears that all shrinkage methods improve portfolio optimization.
- Which is best? Apparently the simplest!

Reference: David Disatnik and Simon Benninga, "Shrinking the Covariance Matrix—Simpler is Better," *Journal of Portfolio Management*, Summer 2007, Vol 33, No. 4, pp. 56-63.