Lecture 5

○ Integer Programming

  ‣ Plant-location example

○ Portfolio Optimization – Part I

  ‣ The Scenario Approach

  ‣ The Mean-Variance Model

○ Summary and Preparation for next class

Integer Programming

Definitions. An *integer program* is a linear program where some or all decision variables are constrained to take on integer values only. A variable is called *integer* if it can take on any value in the range ..., -3, -2,-1, 0, 1, 2, 3,.... A variable is called *binary* if it can take on values 0 and 1 only.

What use?

- ○ Can't build 1.37 aircraft carriers
- ○ Rounding may not give the best, or even a feasible, answer

Selected Applications

- ○ Capital budgeting
  - ‣ invest all or nothing in a project
- ○ Fixed cost/Set-up cost models
- ○ Facility location
  - ‣ build a plant or not (yes/no decision)
- ○ Minimum batch size
  - ‣ if any cars are produced at a plant, then at least 2,000 must be produced
  - ‣ C = 0 or C ≥ 2,000 (either/or decision)
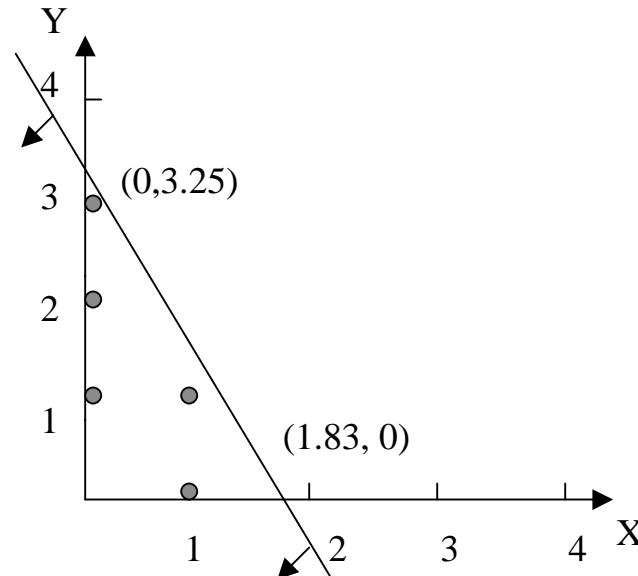
Difficulties in Solving Integer Programs

Example.

max 21 X + 11 Y

subject to:

$$7 X + 4 Y \leq 13$$

$$X, Y \geq 0$$

Y

4

3

2

1

(0,3.25)

(1.83, 0)

1   2   3   4   X

Optimal linear-programming solution: X = 1.83, Y = 0.
Rounded to X = 2, Y = 0 is infeasible.
Rounded to X = 1, Y = 0 is not optimal.
Optimal integer-programming solution: X = 0, Y = 3.

Plant-Location Problem

○ A new company has won contracts to supply a product to customers in Central America, United States, Europe, and South America. The company has determined three potential locations for plants. Relevant cost data are:

Plant Locations	Fixed Costs	Variable Costs	Production Capacity
Brazil	50,000	1,000	30
Philippines	40,000	1,200	25
Mexico	60,000	1,600	35

Fixed costs are in $ per month. Fixed costs are only incurred if the company decides to build and operate the plant. Variable costs are in $ per unit. Production capacities are in units per month. Customer demand (in units per month) is:

	Central America	United States	Europe	South America
Demand	18	15	20	12

In addition to fixed and variable costs, there are shipping costs.

Plant-Location Problem (continued)

Plant

Customer
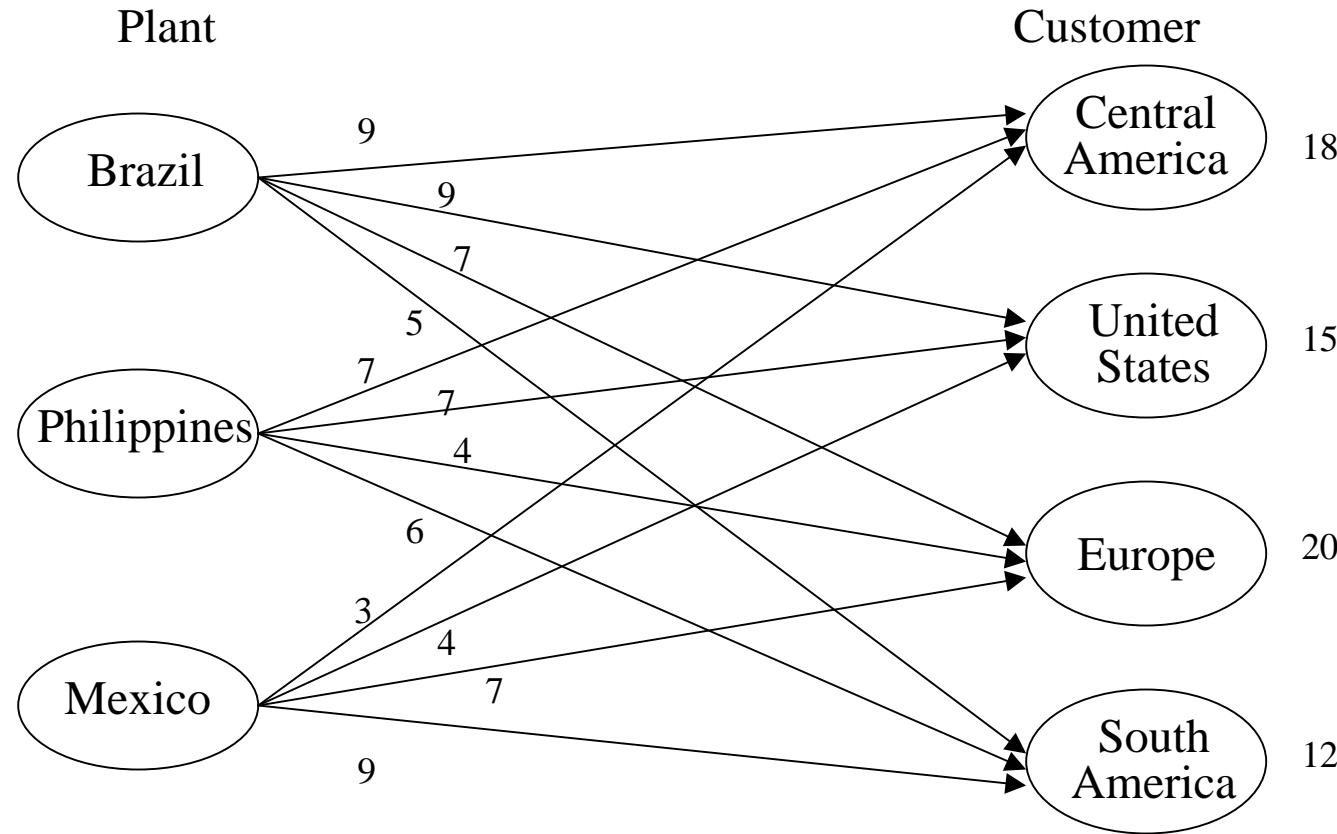
Numbers on arcs represent shipping costs (in $100 per unit).

Which plants and shipping plan minimize monthly production and distribution costs?

Plant-Location Model

○ *Indices:*

Let B represent the Brazil plant, and similarly use P (Philippines), M (Mexico), C (Central America), U (United States), E (Europe), and S (South America).

○ *Decision Variables:* Let

p_B = # of units to produce in Brazil

and similarly define p_P and p_M. Also let

x_{BC} = # of units to ship from Brazil to Central America,

and define x_{BU} , x_{BE} , ..., x_{MS} similarly.

○ *Objective Function:*

The total cost is the sum of fixed, variable, and shipping costs.

Total variable cost is:

$VAR = 1{,}000\ p_B + 1{,}200\ p_P + 1{,}600\ p_M$.

Total shipping cost is:

$SHIP = 900\ x_{BC} + 900\ x_{BU} + 700\ x_{BE} + 500\ x_{BS}$

$+\ 700\ x_{PC} + 700\ x_{PU} + 400\ x_{PE} + 600\ x_{PS}$

$+300\ x_{MC} + 400\ x_{MU} + 700\ x_{ME} + 900\ x_{MS}$.

We will return to the total fixed cost computation shortly.

Plant-Location Model (continued)

○ *Constraints:*

Plant-production definitions: There are constraints to define total production at each plant. For example, the total production at the Mexico plant is:

$$p_M = x_{MC} + x_{MU} + x_{ME} + x_{MS}$$

This can be thought of as a "flow in = flow out" constraint for the Mexico node.

○ *Demand constraints:*

There are constraints to ensure demand is met for each customer. For example, the constraint for Europe is:

$$x_{BE} + x_{PE} + x_{ME} = 20.$$

This is a "flow in = flow out" constraint for the Europe node.

○ *Plant-Capacity Constraints:*

Production cannot exceed plant capacity, e.g., for Brazil

$$p_B \leq 30$$

Fixed-Cost Computation

○ *Additional Decision Variables:* To compute total fixed cost, define the *binary* plant-open variables:

$$y_B = \begin{cases} 1 \text{ if the Brazil plant is opened (i.e., if } p_B > 0) \\ 0 \text{ if the Brazil plant is not opened (i.e., if } p_B = 0) \end{cases}$$

and define y_P and y_M similarly.
Total fixed cost is:

$$FIX = 50{,}000 \ y_B + 40{,}000 \ y_P + 60{,}000 \ y_M$$

As it currently stands, the optimizer will always set the "plant open" variables to zero (so that no fixed cost will be incurred). We need constraints to enforce the meaning of these variables, e.g.,

$$p_B > 0 \quad \Rightarrow \quad y_B = 1.$$

Why not add constraints to define the plant open variables, e.g., for Brazil,

$$y_B = \text{IF} \ (\ p_B > 0 \ , \ 1, \ 0) \ ?$$

Because =IF statements are *not linear* and they are *discontinuous.* Optimizers cannot solve such problems easily, if at all. What else can be done?

Fixed-Cost Computation (continued)

○ If $y_B = 0$ we want to rule out production at the Brazil plant. If the Brazil plant is not opened (i.e., if $y_B = 0$), its "available" capacity is 0. If $y_B = 1$, the plant is open and its "available" capacity is 30 units per month. The plant capacity constraints can be modified to enforce this meaning of y_B:

$$p_B \leq 30 \, y_B$$

If $y_B = 0$ then the constraint becomes $p_B \leq 0$.

If $y_B = 1$ then the constraint becomes $p_B \leq 30$.

Alternatively, if $p_B > 0$ (and y_B can only take on the values 0 or 1) then $y_B = 1$

This is exactly what is needed!

○ *Modified Plant-Capacity Constraints:*

Production cannot exceed plant capacity, e.g., for Brazil

$$p_B \leq 30 \, y_B$$

Binary variable: $y_B = 0$ or 1.

Similar plant-capacity and binary-variable constraints are needed for the Philippines and Mexico.

Plant Location Integer Programming Model

min $VAR + SHIP + FIX$
- ◯ *Cost definitions:*

  $(VAR$ Def.$)$ $VAR = 1{,}000\ p_B + 1{,}200\ p_P + 1{,}600\ p_M$.

  $(SHIP$ Def.$)$ $SHIP = 900\ x_{BC} + 900\ x_{BU} + 700\ x_{BE} + 500\ x_{BS}$
  $+ 700\ x_{PC} + 700\ x_{PU,} + 400\ x_{PE} + 600\ x_{PS}$
  $+ 300\ x_{MC} + 400\ x_{MU} + 700\ x_{ME} + 900\ x_{MS}$

  $(FIX$ Def.$)$ $FIX = 50{,}000\ y_B + 40{,}000\ y_P + 60{,}000\ y_M$

- ◯ *Plant production definitions:*

  (Brazil) $p_B = x_{BC} + x_{BU} + x_{BE} + x_{BS}$
  (Philippines) $p_P = x_{PC} + x_{PU} + x_{PE} + x_{PS}$
  (Mexico) $p_M = x_{MC} + x_{MU} + x_{ME} + x_{MS}$

- ◯ *Demand constraints:*

  (Central America) $x_{BC} + x_{PC} + x_{MC} = 18$
  (United States) $x_{BU} + x_{PU} + x_{MU} = 15$
  (Europe) $x_{BE} + x_{PE} + x_{ME} = 20$
  (South America) $x_{BS} + x_{PS} + x_{MS} = 12$

- ◯ *Modified plant capacity constraints:*

  (Brazil) $p_B \le 30\ y_B$
  (Philippines) $p_P \le 25\ y_P$
  (Mexico) $p_M \le 35\ y_M$

- ◯ *Binary variables:* y_B , y_P , $y_M = 0$ or 1
- ◯ *Nonnegativity:* All variables ≥ 0

Plant Location Optimized Spreadsheet

`=SUMPRODUCT(B5:B7,E5:E7)`

`=SUMPRODUCT(C5:C7,F17:F19)`

|	A	B	C	D	E	F	G	H	I
1	PLANT.XLS		Plant Location Model						
2									
3		Fixed	Variable	Production	Plant		Fixed cost		1,100
4	Plants	Cost	Cost	Capacity	Open		Variable cost		860
5	Brazil	500	10	30	1		Shipping cost		314
6	Philippines	400	12	25	0		Total cost		2,274
7	Mexico	600	16	35	1				
8							(All costs in $100)		
9	Unit Shipping Costs:		Customers						
10		Central Am.	U.S.	Europe	S.Amer.				
11	Brazil	9	9	7	5				
12	Philippines	7	7	4	6				
13	Mexico	3	4	7	9				
14									
15	Shipping Plan:		Customers				Capacity	Available	
16		Central Am.	U.S.	Europe	S.Amer.	Total	Constraint	Capacity	
17	Brazil	0	0	18	12	30	<=	30	
18	Philippines	0	0	0	0	0	<=	0	
19	Mexico	18	15	2	0	35	<=	35	
20	Total	18	15	20	12				
21	Constraint	=	=	=	=				
22	Demand	18	15	20	12				

`=SUMPRODUCT(B11:E13,B17:E19)`

`=D7*E7`

◯ Decision variables in cells E5:E7 are restricted to 0 or 1, i.e., they are constrained to be *binary*.

◯ Note that many numbers in the spreadsheet were scaled to units of $100. For the optimizer to work properly, it is important (especially with integer programs) to scale the numbers to be about the same magnitude.

◯ Shadow price information is not available with integer programs; the Excel optimizer does not give meaningful sensitivity reports.

Solver Parameters
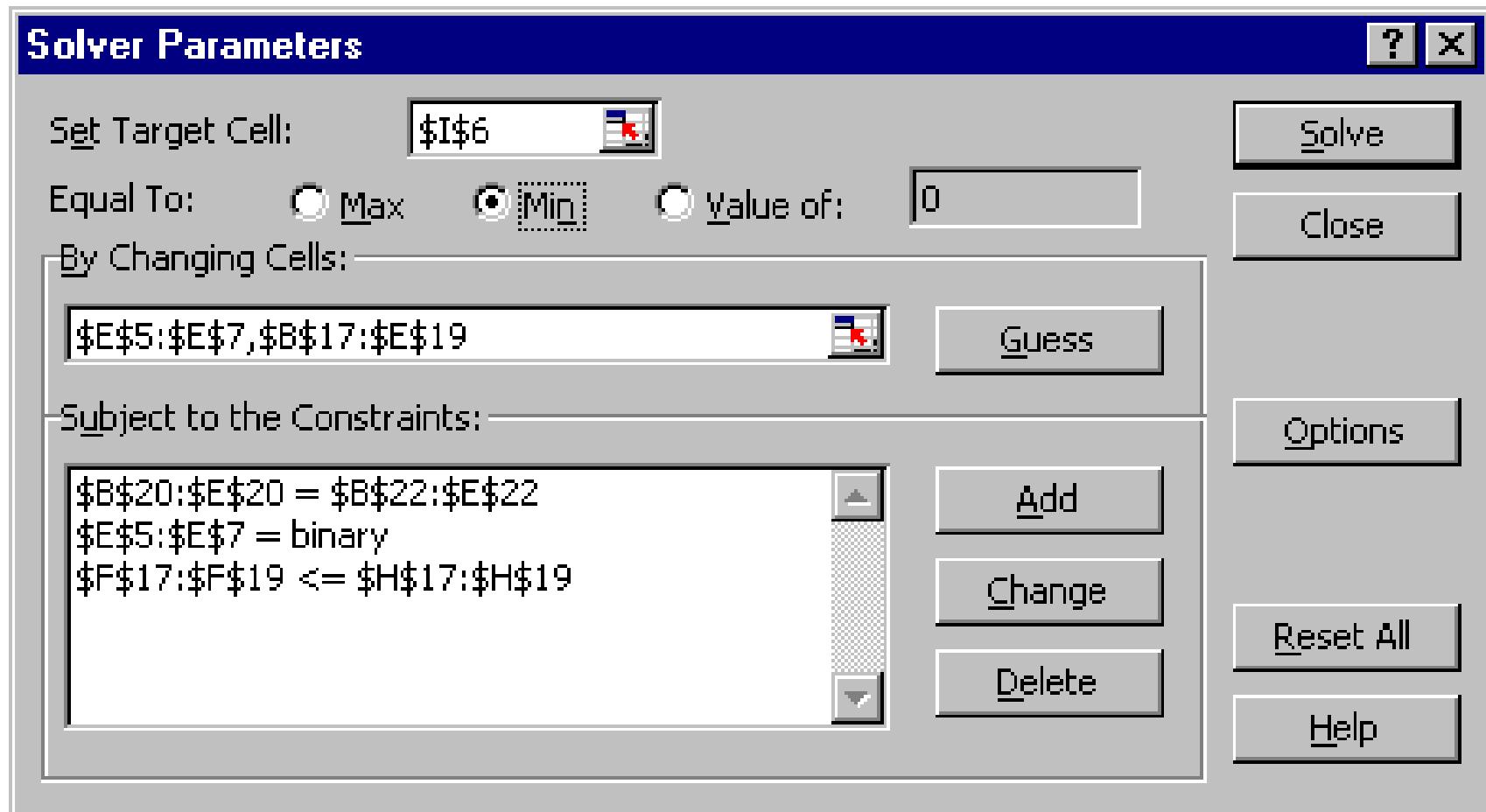
The Solver Parameters Dialog Box

Solver Options Dialog Box
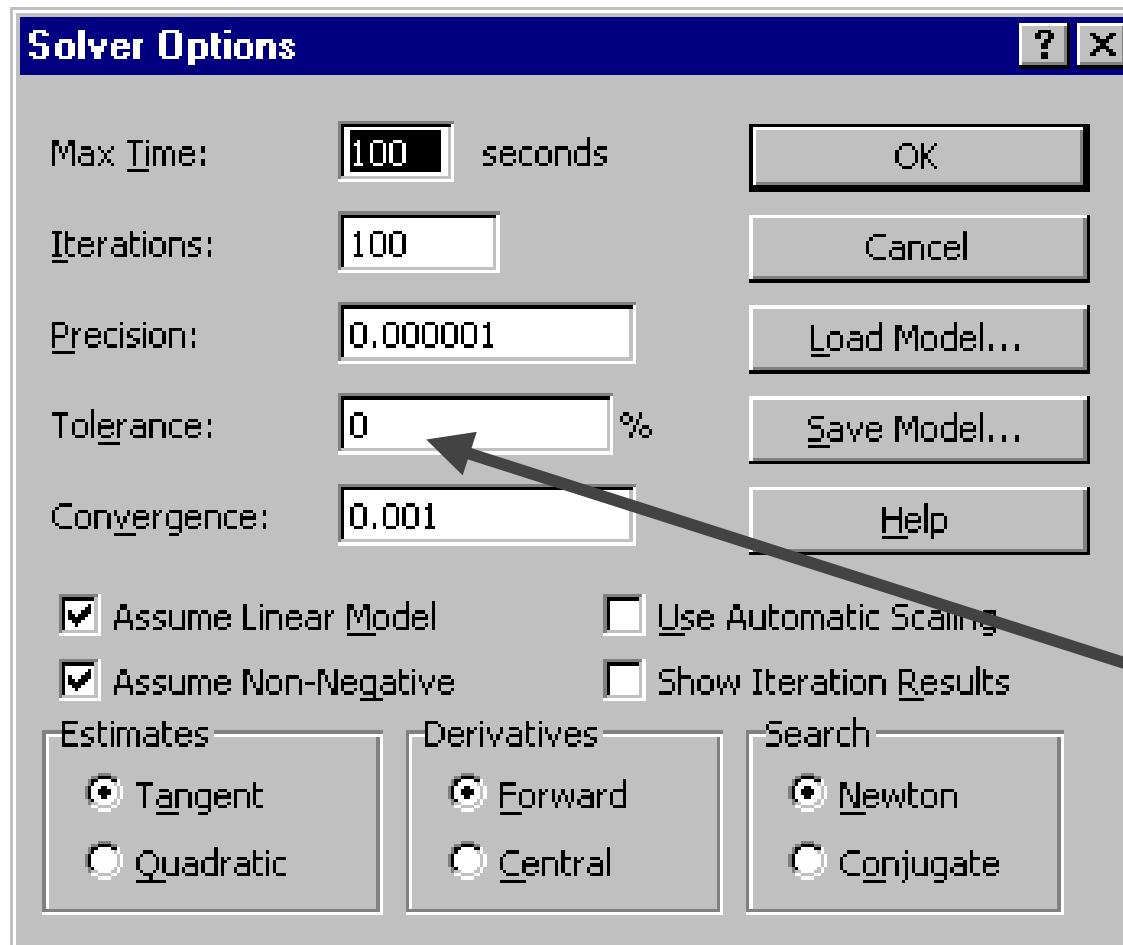
Default is 5%,

change to 0

❍ In the Solver Options Dialog Box, make sure to change the "Tolerance" setting to 0%. This is necessary will ensure that Solver finds the best possible solution.

Plant Location "Optimized" Spreadsheet
Using = IF statements

`=IF(F17>0,1,0)`

	A	B	C	D	E	F	G	H	I
1	PLANT_IF.XLS		Plant Location Model						
2									
3		Fixed	Variable	Production	Plant		Fixed cost		1500
4	Plants	Cost	Cost	Capacity	Open		Variable cost		760
5	Brazil	500	10	30	1		Shipping cost		367
6	Philippines	400	12	25	1		Total cost		2,627
7	Mexico	600	16	35	1				
8							(All costs in $100)		
9	Unit Shipping Costs:		Customers						
10		Central Am.	U.S.	Europe	S.Amer.				
11	Brazil	9	9	7	5				
12	Philippines	7	7	4	6				
13	Mexico	3	4	7	9				
14									
15	Shipping Plan:		Customers				Capacity	Available	
16		Central Am.	U.S.	Europe	S.Amer.	Total	Constraint	Capacity	
17	Brazil	8	10	0	12	30	<=	30	
18	Philippines	0	5	20	0	25	<=	25	
19	Mexico	10	0	0	0	10	<=	35	
20	Total	18	15	20	12				
21	Constraint	=	=	=	=				
22	Demand	18	15	20	12				

❍ In this spreadsheet, the plant-open cells, E5:E7, are computed with =IF statements.

❍ The optimizer returns an *incorrect* optimal solution because of the =IF statements.

❍ *This is not an Excel bug.* It is simply a difficult problem for any optimizer to solve because =IF statements represent discontinuous functions.

Portfolio Optimization

Problem: What portfolio to invest in today given an uncertain future?

This investment problem is often called an *asset-allocation* or *portfolio-selection* decision. The assets or securities could include Treasury bonds, options, mortgage-backed securities, foreign stocks, real estate, etc.

Example. Suppose an investor is considering investing in 3 asset classes:
(1) stocks, (2) bonds, and (3) T-bills.

Suppose the investor has a budget of $2,000,000 and the investor's portfolio consists of $1,200,000 in stocks, $600,000 in bonds, and $200,000 in T-bills.

○ Index the asset classes by $j = 1, ..., n$. Define the *decision variables*

x_j = fraction of budget invested in asset class j.

For this example, the investor's portfolio is $(x_1, x_2, x_3) = (0.6, 0.3, 0.1)$.

Definition: A *portfolio* is an allocation x_j, $j = 1, ..., n$, satisfying $\sum_{j=1}^{n} x_j = 1$
and $x_j \geq 0$ for $j = 1, ..., n$.

Note: $x_j \geq 0$ prohibits *short sales*.

A Model of the Uncertain Future

Consider a 1-period model with a finite number of future scenarios.

Price

Scenario 1

Scenario 2

Scenario 3

Scenario 4

today

1 month
later

Time

Make portfolio
decision

Random future
returns
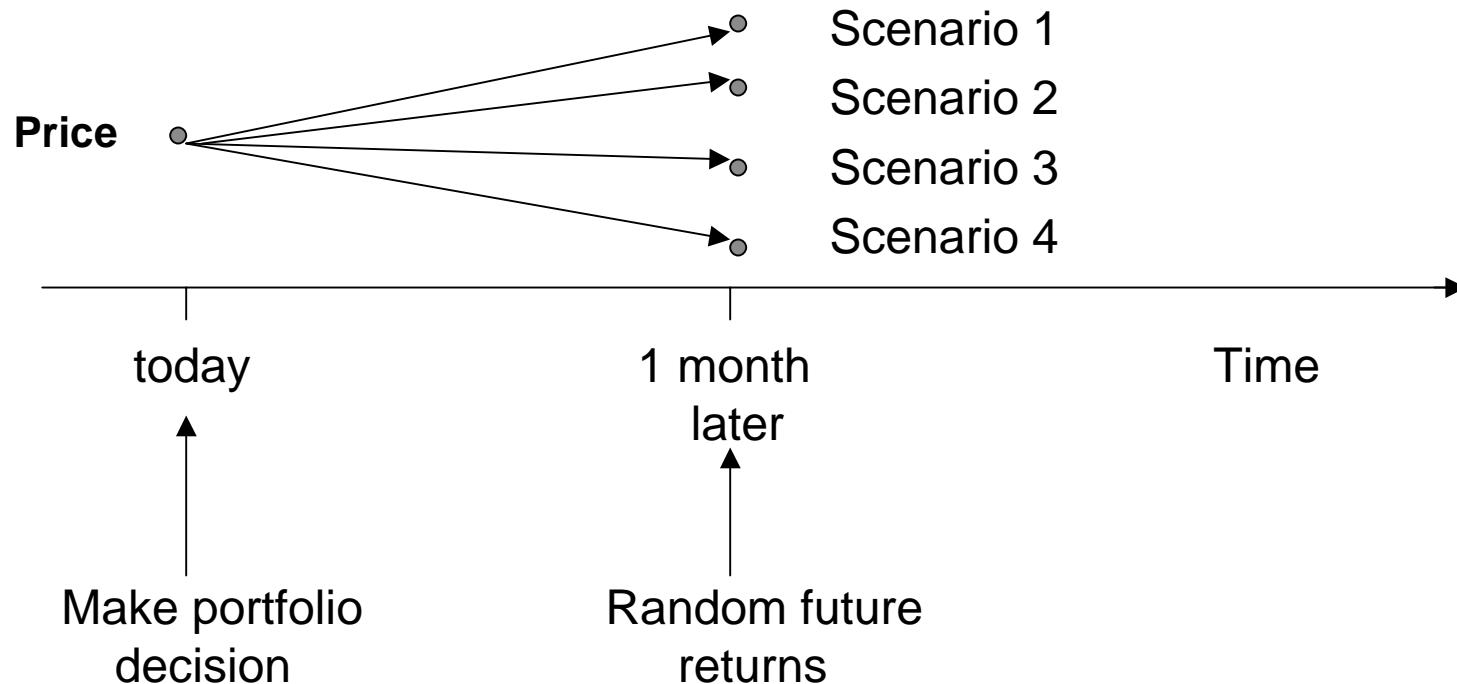
p_i = probability scenario *i* occurs

Definition: A *scenario* is a list of returns for the *n* securities.

Scenario Returns and Probabilities

Table. (Monthly returns)

	Prob.	Security 1	Security 2	Security 3
Scenario 1	0.25	5.51%	1.95%	2.56%
2	0.25	−1.24%	2.26%	0.16%
3	0.25	5.46%	−4.07%	−0.64%
4	0.25	−1.90%	3.59%	0.30%

Let r_{ij} denote the return of security j if scenario i occurs. E.g., $r_{32}=-4.07\%$.

Where do the scenarios come from?

‣ Historical returns

‣ Security analysts' forecasts

‣ Economic/Financial models

‣ A combination of the above

Portfolio Returns

If scenario i occurs, what is the return of the portfolio $(x_1, ..., x_n)$?

The portfolio return if scenario i occurs, denoted r_i , is

$$r_i = \sum_{j=1}^{n} r_{ij} x_j.$$

(1)

Portfolio Returns (continued)

○ Example. Suppose the investor's portfolio is $(x_1, x_2, x_3) = (0.5, 0.5, 0)$.
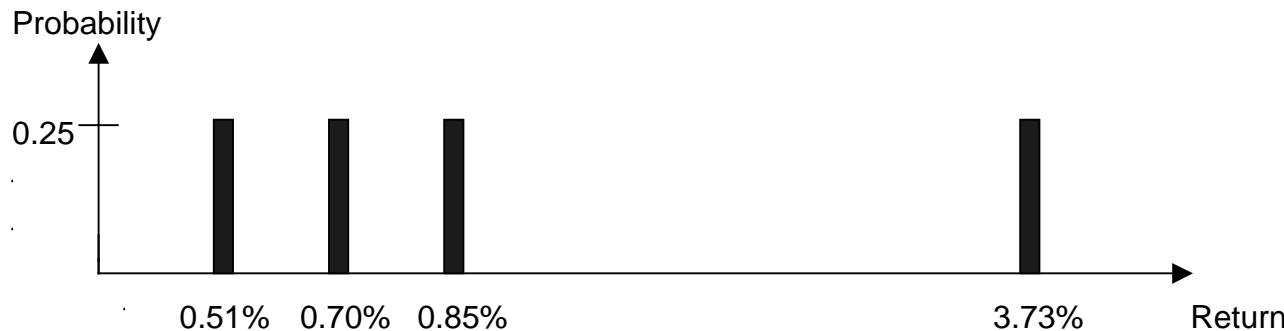Then, from equation (1), the portfolio returns in the four scenarios are:

Scenario 1: $r_1 = \ \ 5.51(0.5) \ + 1.95(0.5) + 2.56(0) = \ \ 3.73$

Scenario 2: $r_2 = -1.24(0.5) \ + 2.26(0.5) + 0.16(0) = \ \ 0.51$

Scenario 3: $r_3 = \ \ 5.46(0.5) \ - 4.07(0.5) - 0.64(0) = \ \ 0.70$

Scenario 4: $r_4 = -1.90(0.5) \ + 3.59(0.5) + 0.30(0) = \ \ 0.85$

This distribution of returns can be plotted as follows:

○ Different portfolios will have different distributions of returns. How can
an investor express a preference for one distribution over another?

Preferences for Return Distributions
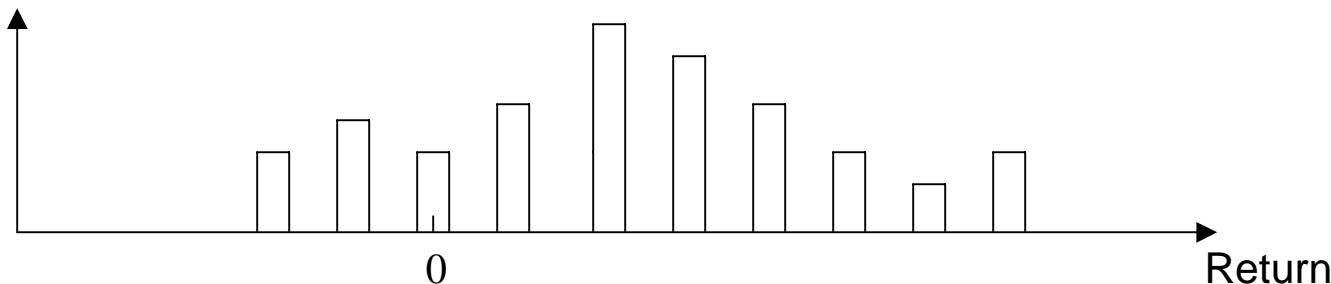
❍   Consider two return distributions:

Probability   Distribution 1:

0   Return

Distribution 2:

Probability

0   Return

❍   The returns in Distribution 2 are higher than the returns in Distribution 1. Hence, most rational investors would prefer 2 to 1. Generally, though, one distribution will not dominate another in this way. So how can we express a preference over complicated distributions?

One way is to summarize a distribution is by its *average return*.

Average Portfolio Return

Definition: A portfolio's *average return* of a portfolio, denoted r_P, is

$$r_P = \sum_{i=1}^{m} p_i r_i. \qquad (2)$$

○ The average return is the return of the portfolio in each scenario (r_i) weighted by the probability that the scenario occurs (p_i). In the example,

  $r_P = 0.25(3.73) + 0.25(0.51) + 0.25(0.70) + 0.25(0.85) = 1.45\%$.

  Or, in the case of equal probability scenarios, we can use the shortcut:

  $r_P = \text{AVERAGE} (r_1 , r_2 , r_3 , r_4) = 1.45\%$.

○ The average summarizes the *location* of a distribution with a single number:

Probability

0

Return
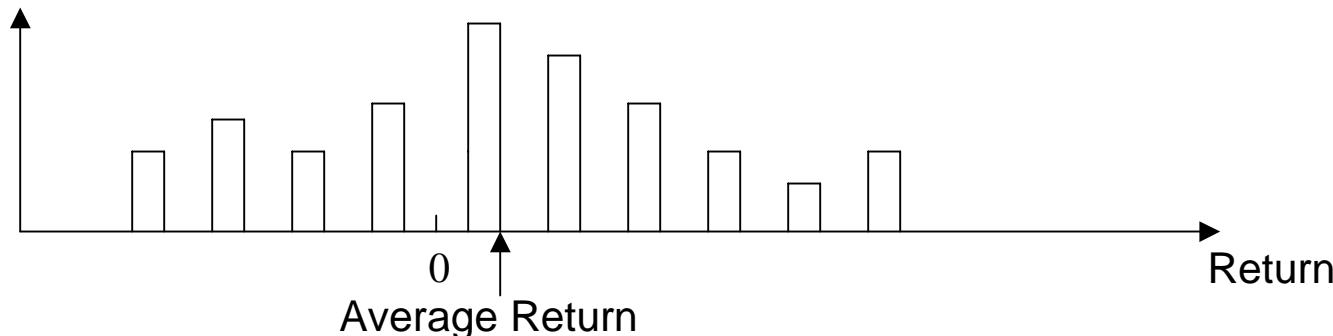
Average Return

○ Most investors would prefer r_P to be as large as possible, everything else equal.

○ What else matters?

Standard Deviation of Return

○ Suppose r_P = 1%. This is the average, and the actual return could differ substantially from that value. *Risk* can be measured by the *uncertainty*. One measure of risk is the *standard deviation (SD)* of returns.

○ With equal probability scenarios, we can get the standard deviation by using the Excel function =STDEVP. For example:

▸ Portfolio Standard Deviation (SD) = STDEVP(r_1, r_2, r_3, r_4).

▸ For portfolio (x_1, x_2, x_3) = (0.5, 0.5, 0): r_P = 1.45% and SD = 1.53%.

○ What about the portfolio (x_1, x_2, x_3) = (0, 1, 0), i.e., all in security 2?

▸ Well, the average portfolio return is r_P = 0.93% and SD = 3.41%.

▸ This portfolio has a smaller average return and larger risk (as measured by *SD*) compared to the portfolio (0.5, 0.5, 0). <u>Portfolio (0.5, 0.5, 0) *dominates* portfolio (0, 1, 0).</u>

Summary

Table. (Monthly returns expressed in percent)

	Prob.	Security 1	Security 2	Security 3
Scenario1	0.25	5.51	1.95	2.56
2	0.25	−1.24	2.26	0.16
3	0.25	5.46	−4.07	−0.64
4	0.25	−1.90	3.59	0.30

A portfolio of these three securities, denoted (x_1, x_2, x_3), must satisfy:

▸ *(Budget)* $\quad\quad\quad\quad x_1 + x_2 + x_3 = 1$

▸ *(No short sales)* $\quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0,$

and the portfolio returns in each scenario are given by:

Scenario 1: $r_1 = \ \ 5.51\ x_1 + 1.95\ x_2 + 2.56\ x_3$

Scenario 2: $r_2 = -1.24\ x_1 + 2.26\ x_2 + 0.16\ x_3$

Scenario 3: $r_3 = \ \ 5.46\ x_1 - 4.07\ x_2 - 0.64\ x_3$

Scenario 4: $r_4 = -1.90\ x_1 + 3.59\ x_2 + 0.30\ x_3$

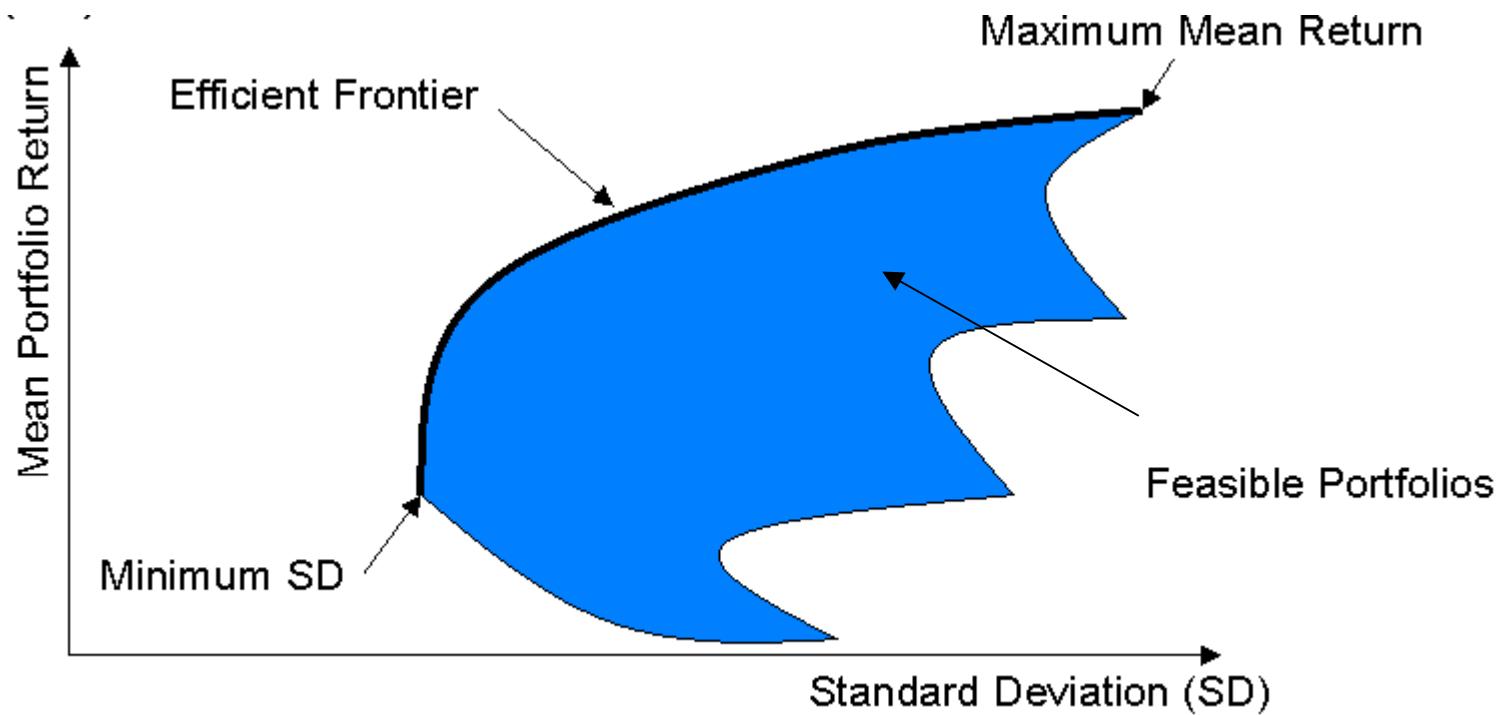The average portfolio return is given by:

$$r_P = \text{AVERAGE}(r_1, r_2, r_3, r_4)$$

The standard deviation of the portfolio return (i.e., the risk) is:

$$SD = \text{STDEVP}(r_1, r_2, r_3, r_4)$$

Efficient Frontier

○ For any portfolio $(x_1, ..., x_n)$ with $\sum_{j=1}^{n} x_j = 1$ and $x_j \geq 0$, we can compute the corresponding average portfolio return r_P and standard deviation (*SD*). The set of all feasible portfolios is as follows:

○ Average return and risk are *two conflicting objectives*. Since we can't have two objective functions in an optimization model, choose one to be the objective and the other to be a constraint.

Portfolio-Optimization Model

○ One formulation of the portfolio-optimization model is: over all feasible portfolios, minimize "risk" (e.g., *SD*) subject to "reward" (e.g., r_P) at least some user-specified level. That is,
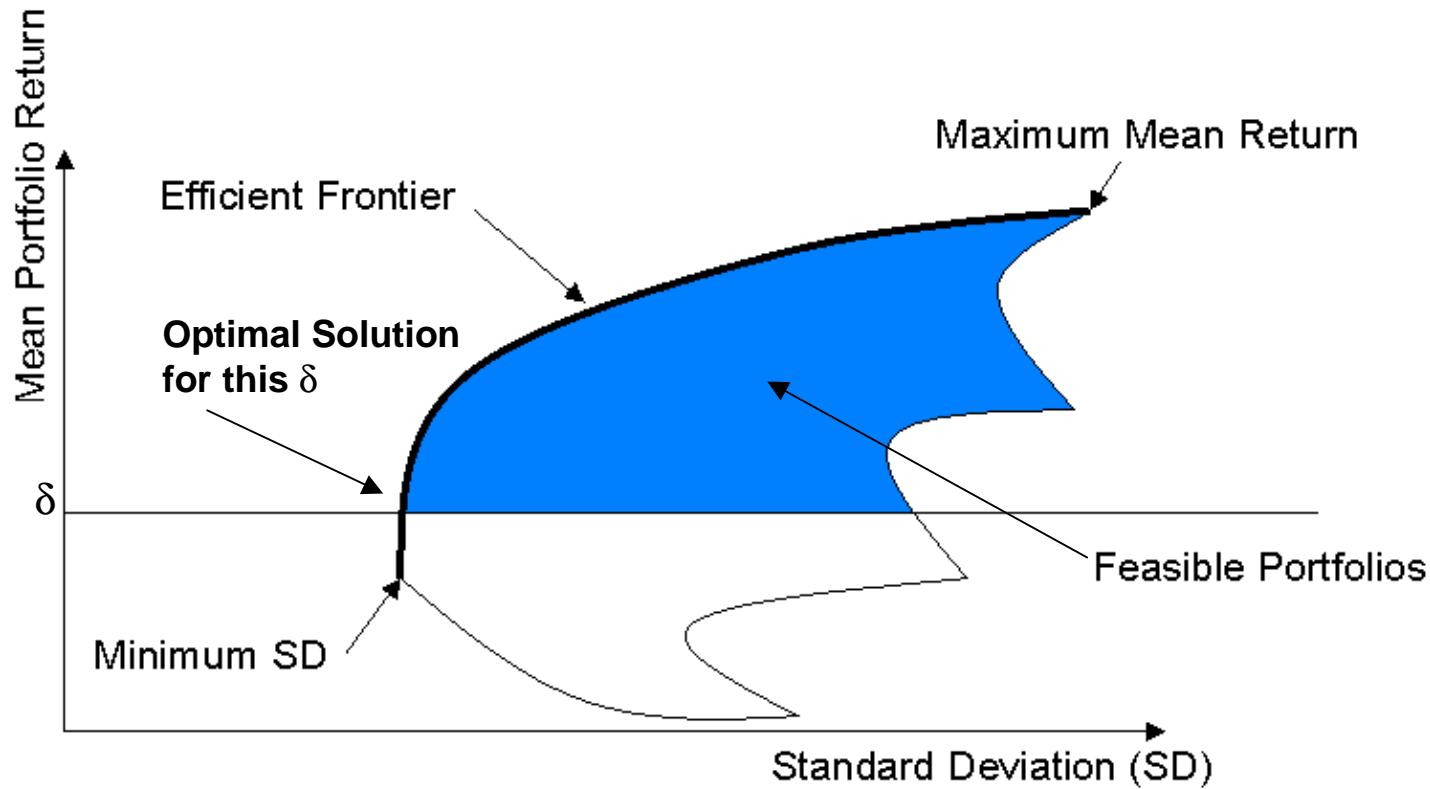
$$\min \quad SD$$

subject to:

(Average return) $\quad r_P \geq \delta$

(Budget) $\quad x_1 + x_2 + x_3 + \ldots + x_n = 1$

(No short sales) $\quad x_j \geq 0$ for all j

○ δ is a user-supplied constant, indicating the minimum level of average return that the investor is willing to accept.

○ This is a *non-linear* model.

Portfolio-Optimization Model (continued)

❍ Next we specify the details of the optimization model.

Mean-Variance Portfolio-Optimization Model

The complete *non-linear* optimization model can be written as:

min SD

subject to:

$(r_1$ def.$)$ $r_1 = 5.51\ x_1 + 1.95\ x_2 + 2.56\ x_3$

$(r_2$ def.$)$ $r_2 = -1.24\ x_1 + 2.26\ x_2 + 0.16\ x_3$

$(r_3$ def.$)$ $r_3 = 5.46\ x_1 - 4.07\ x_2 - 0.64\ x_3$

$(r_4$ def.$)$ $r_4 = -1.90\ x_1 + 3.59\ x_2 + 0.30\ x_3$

$(r_P$ def.$)$ $r_P = 0.25\ r_1 + 0.25\ r_2 + 0.25\ r_3 + 0.25\ r_4$

(Min. r_P) $r_P \geq \delta$

  (Risk) $SD = \mathrm{STDEVP}(r_1, r_2, r_3, r_4)$

 (Budget) $x_1 + x_2 + x_3 = 1$

(nonneg.) $x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0.$

This formulation can easily be set up in a spreadsheet, but it is a *non-linear* model since the standard deviation involves squares and square-roots.

Spreadsheet Solution

=AVERAGE(D11:D14) =STDEVP(D11:D14) Objective Function

=SUM(F5:H5)

	A	B	C	D	E	F	G	H	I	J
1	INVEST.XLS			**Investment Non-Linear Program: Equal Probability Scenarios**						
2										
3			Avg. Portfolio	Portfolio		Portfolio Weights x(j)				Sum of Portfolio
4			Return	Stnd. Dev.		1	2	3		Weights
5			1.00	**1.24**		**23.2%**	**26.4%**	**50.4%**		100%
6			>=							=
7		Min Return	1.00							100%
8				Portfolio						
9				Return		Security returns by scenario				
10			Scenario	by Scenario		1	2	3		
11			1	3.08		5.51	1.95	2.56		
12			2	0.39		-1.24	2.26	0.16		
13			3	-0.13		5.46	-4.07	-0.64		
14			4	0.66		-1.90	3.59	0.30		
15										
16					Average	1.96	0.93	0.60		
17					StdDev	3.54	2.95	1.19		
18										
19						Correlations				
20					Sec1 vs. Sec2	-0.68				
21					Sec1 vs. Sec3	0.31				
22					Sec2 vs. Sec3	0.48				

Decision Variables

=SUMPRODUCT(F14:H14,$F$5:$H$5)

=AVERAGE(H11:H14)

=STDEVP(H11:H14)

=CORREL(F11:F14,G11:G14)

○ The spreadsheet shows the optimal solution corresponding to $\delta = 1.0$ (where δ is set in cell C7).
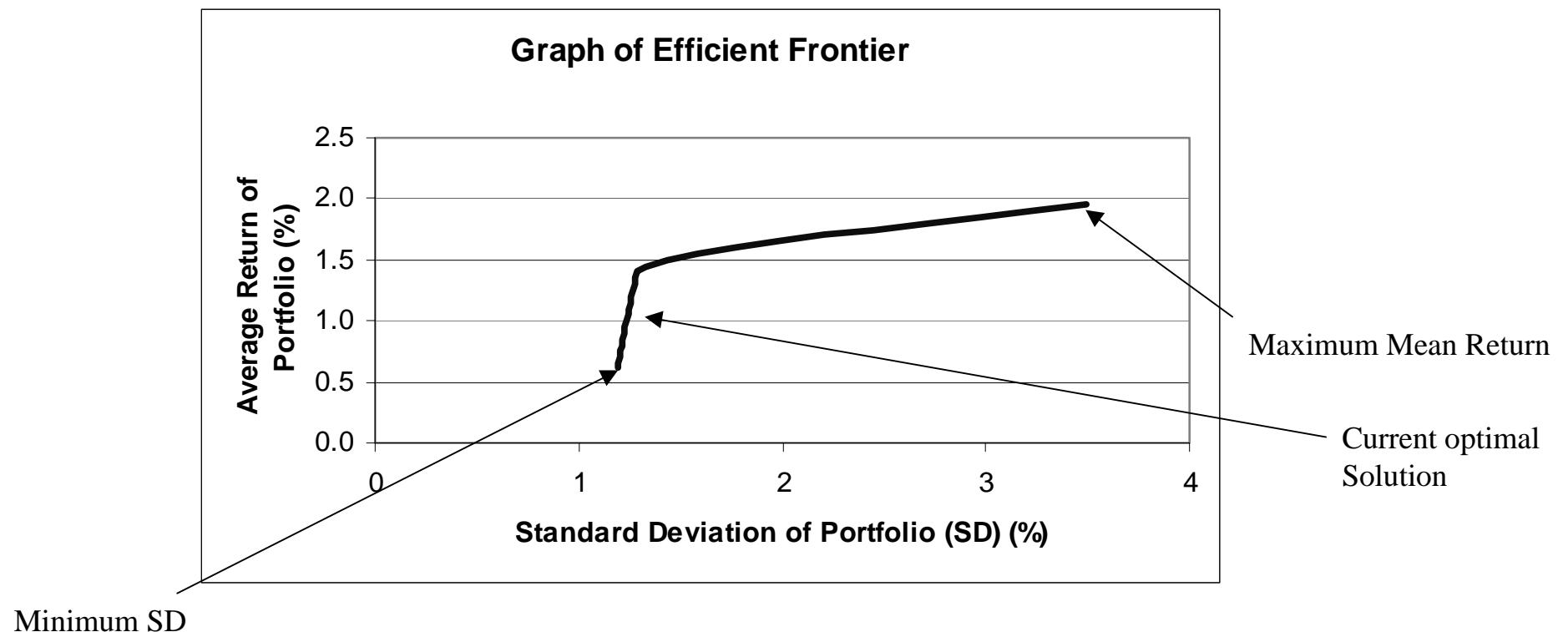
Solver Parameters
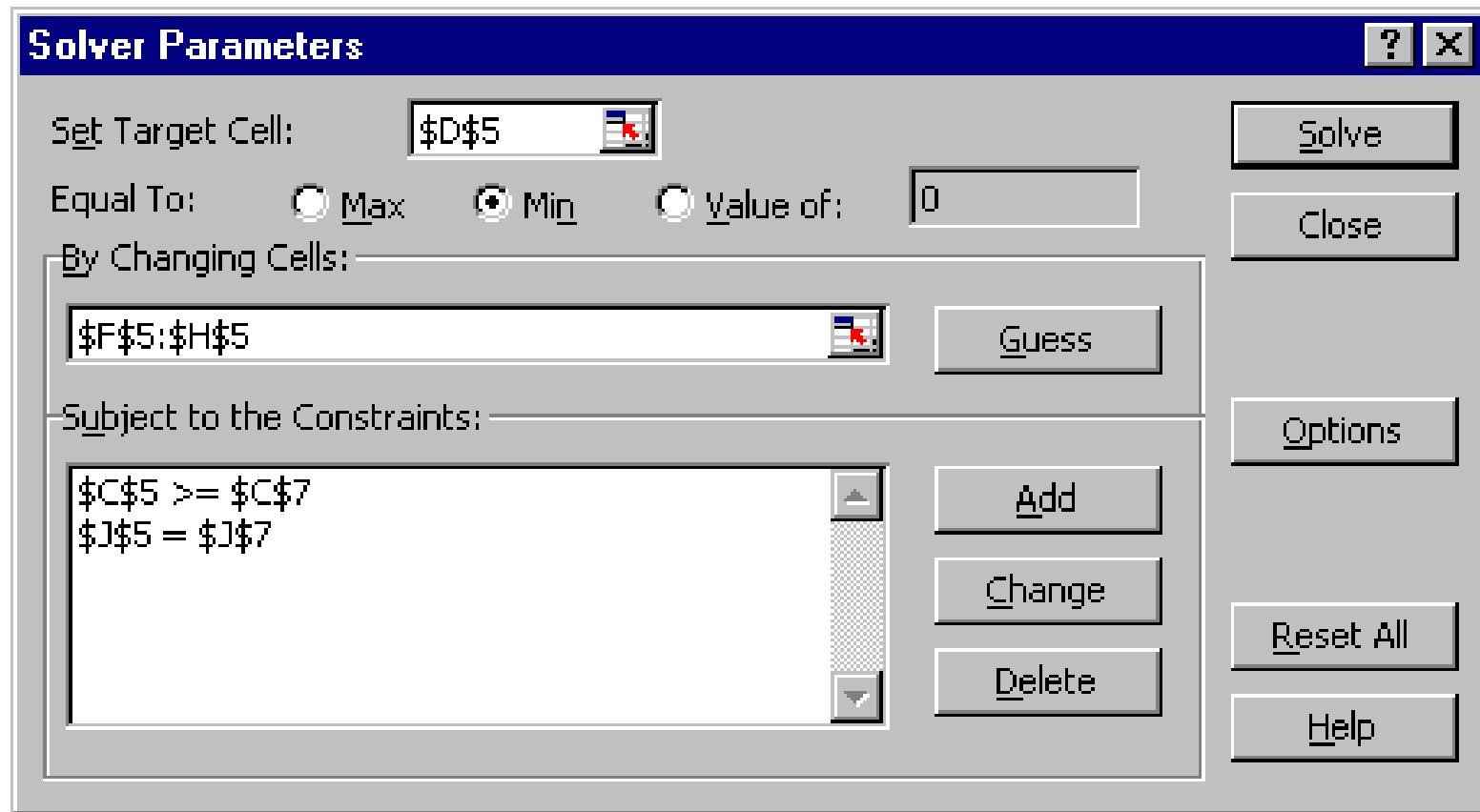
The solver parameters dialog box.

❍ Remember: do not click on "Assume Linear Model" (in "Options") since it is a non-linear model.

Optimization-Model Results

○ For $\delta = 1.0$, the optimal solution is:

$$x_1 = 23.2\%, \quad x_2 = 26.4\%, \quad x_3 = 50.4\%$$

$$r_1 = 3.08\%, \quad r_2 = 0.39\%, \quad r_3 = -0.13\%, \quad r_4 = 0.66\%$$

with $SD = 1.24\%$ and $r_P = 1.00\%$.

○ Using SolverTable, we can vary δ and graph the optimal solutions to the problem. These trace out the *efficient frontier*.

Graph of Efficient Frontier

Average Return of Portfolio (%)

Standard Deviation of Portfolio (SD) (%)

Maximum Mean Return

Current optimal Solution

Minimum SD

Comments on the Mean-Variance Model

❍ Alternate formulation: maximize return subject to a user-specified maximum risk (SD).

❍ The mean-variance approach leads to a nonlinear model

- ▶ This non-linear model is more difficult to solve than a linear one, but Excel can solve it.

- ▶ Variance penalizes upside and downside returns

- ▶ Less sensitivity-analysis information available with nonlinear programs

- ▶ Right-hand side ranges are not given for nonlinear models

❍ Alternative models: Use a measure of risk, e.g. Average Downside Risk (ADR), which can be formulated as a linear model. See the readings book for details ("Portfolio Optimization Using Linear Programming").

Comparison of Scenario Approach
with Variance-Covariance Approach

❍ An alternate method of formulating a mean-variance portfolio optimization model involves computing the variance-covariance matrix of the security returns.

❍ Because the scenario approach uses the security returns directly, it does not require computation of the variance-covariance matrix.

❍ *The two approaches give the same answer! (They are really identical.)*

❍ Since the scenario-based model uses the data directly, the historical correlations between security returns are used implicitly, and need not be computed.

Summary

❍ Integer Programming - A Plant Location Example

❍ Portfolio Optimization

  ▶ Modeling uncertainty with scenarios

  ▶ Definitions of reward and risk

  ▶ Tradeoff between two conflicting objectives

  ▶ The Efficient Frontier

  ▶ Setting up the Model

  ▶ Solving the Model

  ▶ Looking at the Results

For next class

❍ *Optional readings*: "Exploring the New Efficient Frontier" and "Asset Allocation in a Downside-Risk Framework" in the readings book.