Numerical approaches to analysis

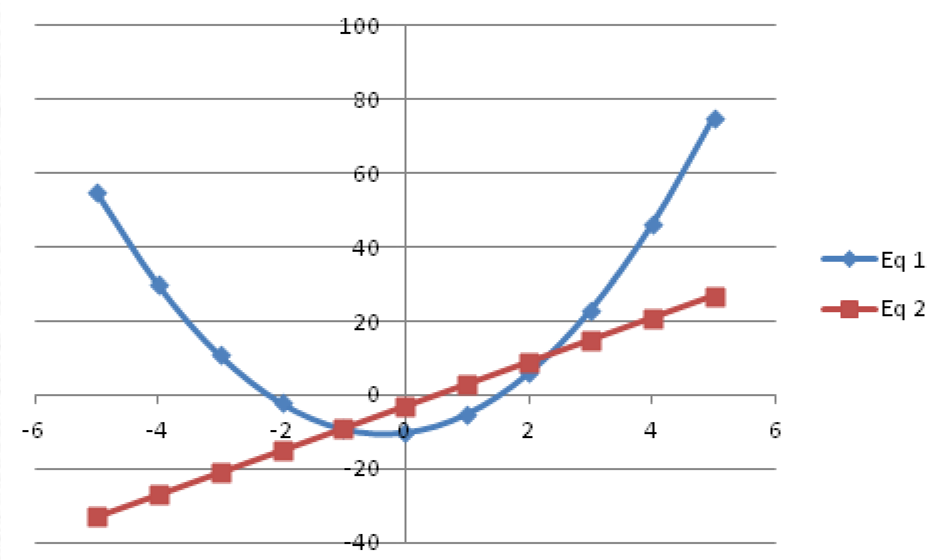
Using numerical solutions to problems
The Solver in Excel

Analytical vs. numerical solutions

- When a mathematical solution can be found by writing equations and solving for the parameter of interest, the solution is an “analytical” solution
- Numerical approach – iterative, sophisticated form of trial and error
- Analytical solutions are exact, numerical solutions are approximate

Solutions to equations

- Two equations, one linear and one quadratic
  - $y = 3x^2 + 2x - 10$
  - $y = 6x - 3$
- Roots of the equations: what value(s) of x yield the same values of y for both equations?
- There is an analytical solution to this – we'll see if a numerical approach agrees with the analytical approach
- First, graph them – how many solutions should we expect?



The analytical solution

$$y = 3x^2 + 2x - 10$$
$$y = 6x - 3$$

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$6x - 3 = 3x^2 + 2x - 10$$

$$0 = 3x^2 - 4x - 7$$

$$x = 2\frac{1}{3}, \quad y = 11$$

$$x = -1, \quad y = -9$$

Quadratic equation in standard form

Numerical approach

- We'll get solutions for both X and Y at the same time
- Re-write the equations as:

$$0 = 3x^2 + 2x - 10 - y$$

$$0 = 6x - 3 - y$$

- Use same values for X and Y in both equations, change them until both equal 0
- When the sum of both equations is 0, do you have your solutions? Problem?

Numerical method – in Excel

- Enter initial "guesses" for x and y
- Calculate F1 and F2 for these guesses
- Square the values of F1 and F2, and sum them
- Change initial guesses for X and Y – if the sum of the squared values gets closer to 0, we are getting closer to a solution
- Stop when the sum of squared differences is equal to 0 (or at least, close enough)

	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	0	
5	y	0	
6			
7	F1	-10	
8	F2	-3	
9			
10	Sum sq	109	
11			

Try a new set of numbers for x and y

- Use x = 1, y = 1
- The sum of squared values for F1 and F2 is smaller, but not 0 yet
- Try again...

	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	1	
5	y	1	
6			
7	F1	-6	
8	F2	2	
9			
10	Sum sq	40	
11			

Not 0 yet...

- Better
- Blind search could take a really long time
- There are good search algorithms that converge on solutions quickly
- Excel's Solver uses these

	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	2	
5	y	10	
6			
7	F1	-4	
8	F2	-1	
9			
10	Sum sq	17	
11			

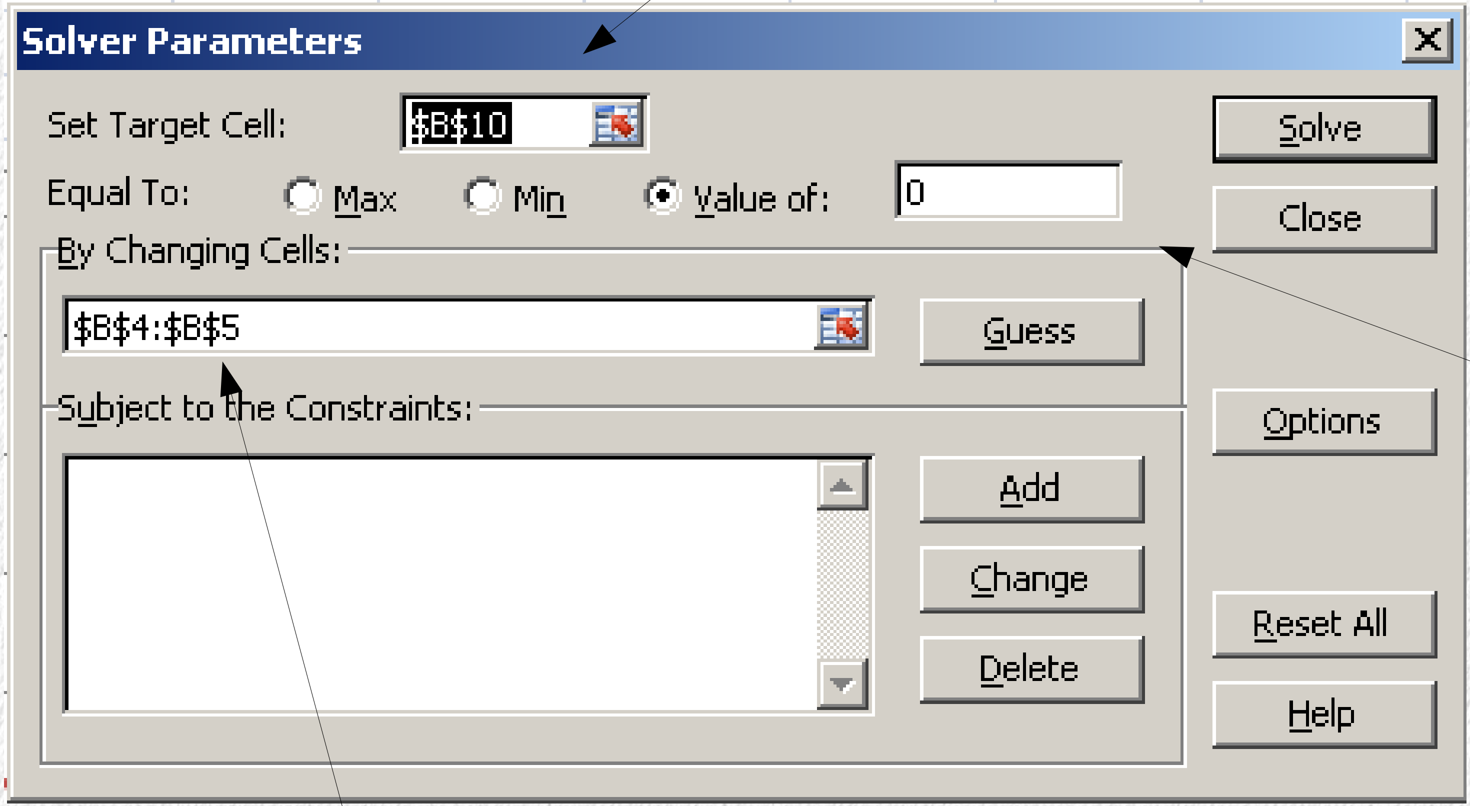
Optimization algorithms used by Excel's Solver

- For linear problems, uses the Simplex method
- For non-linear problems it uses a generalized gradient method
- Which to use is judged by Excel based on the formulas in the spreadsheet
- Both require initial guesses of the solutions
- Both can accept constraints (i.e. only positive values considered)
- Both are iterative (i.e. new values chosen until no more improvement possible)

	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	0	
5	y	0	
6			
7	F1	-10	
8	F2	-3	
9			
10	Sum sq	109	
11			

Solver

The sum of the squares of the two equations

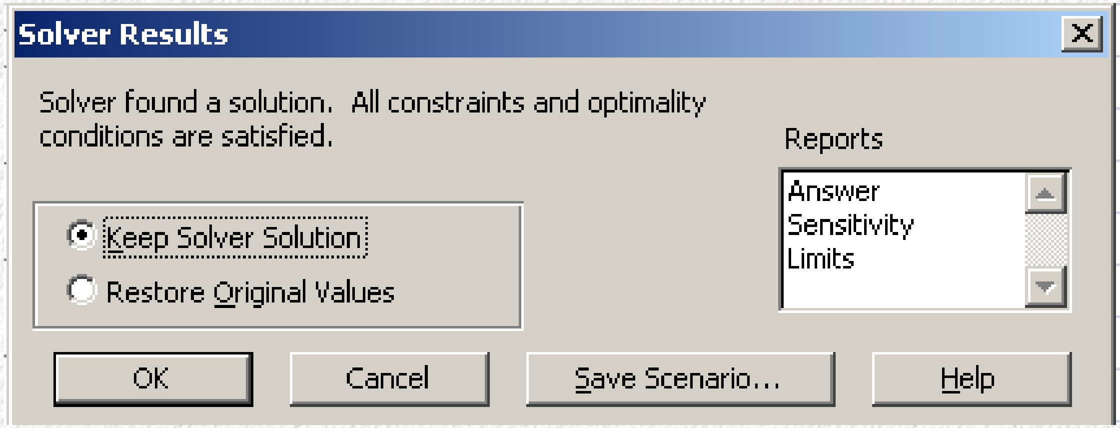


The value we want the sum to be

The values of x and y to change

The first solution

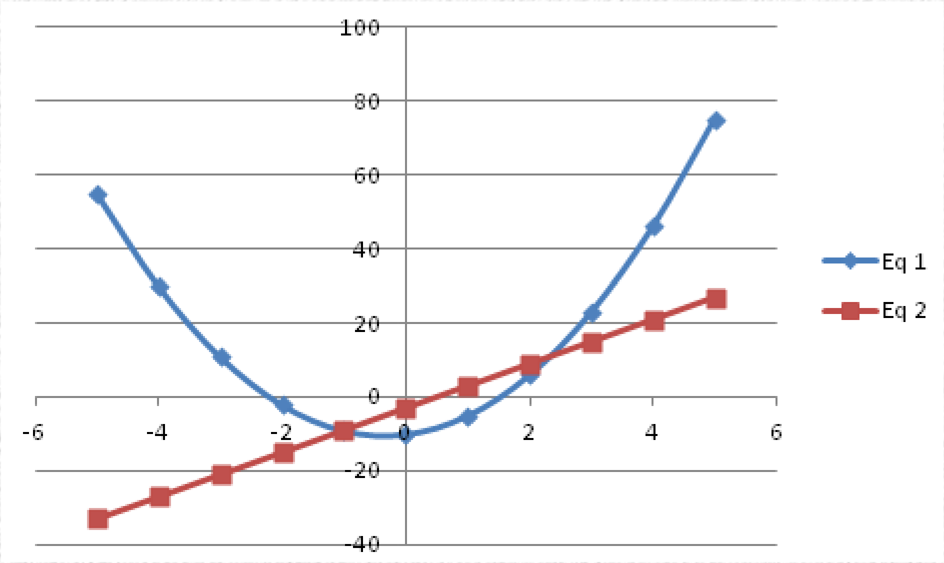
	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	2.333368	
5	y	11.00036	
6			
7	F1	0.000189	
8	F2	-0.00015	
9			
10	Sum sq	5.95E-08	



$$x = 2\frac{1}{3}, \quad y = 11$$

What about the second solution?

- There are two points of intersection of the lines, so two solutions (i.e. two values of x at which the lines have the same y value)
- Our trial and error approach gave us one, now we need the other
- To get the second, try another starting point closer to the other solution and run Solver again



Start closer to the second solution...

	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	-1	
5	y	-10	
6			
7	F1	1	
8	F2	1	
9			
10	Sum sq	2	
11			

Second solution

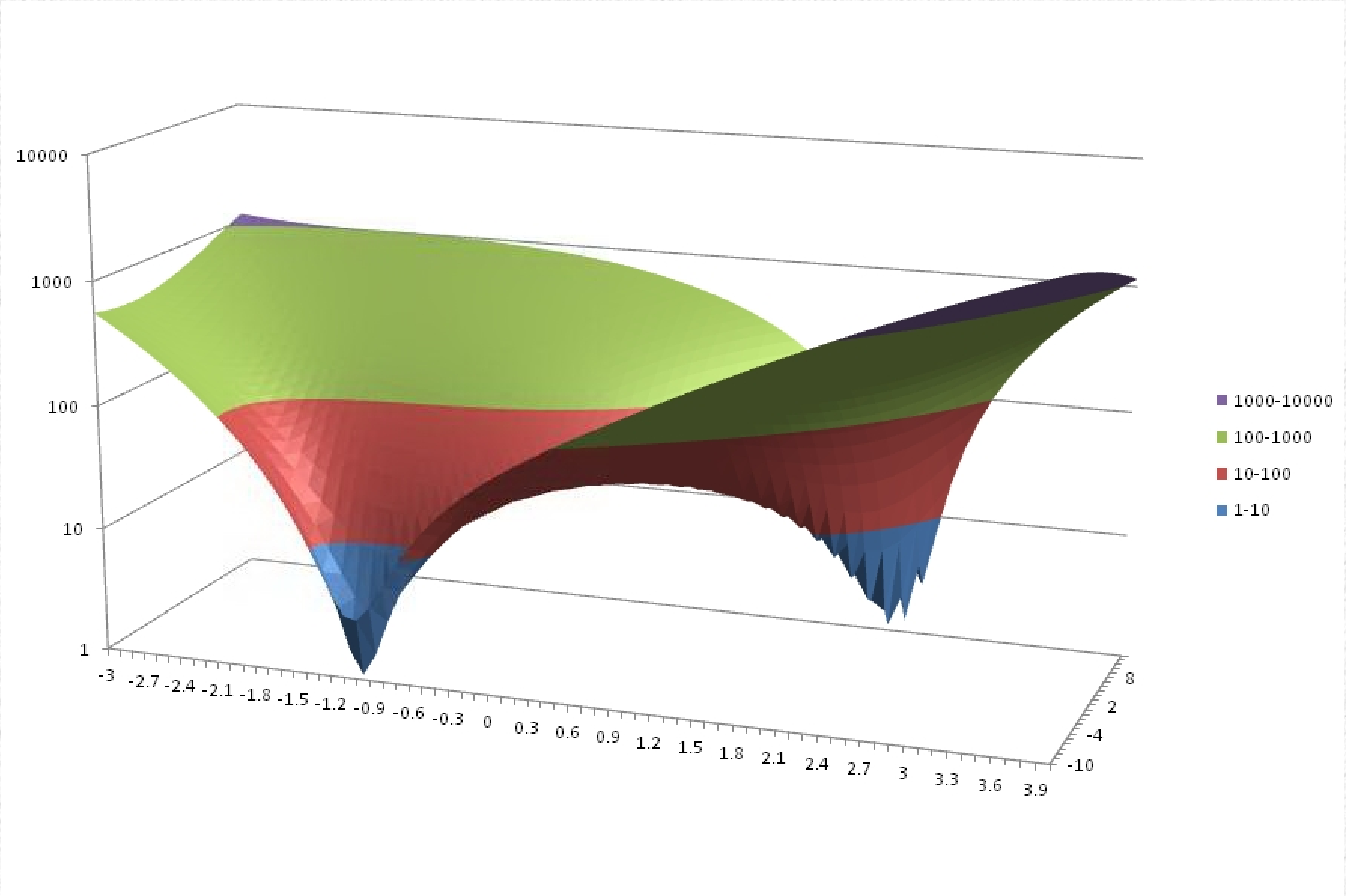
	A	B	C
1	F1	3x^2+2x -10 -y	
2	F2	6x - 3 - y	
3			
4	x	-1	
5	y	-8.99996	
6			
7	F1	-5.6E-05	
8	F2	-1.4E-05	
9			
10	Sum sq	3.32E-09	
11			

Note that this works best when you know how many solutions to expect – graphing is a very important first step

If you don't know, choose several different starting positions to increase the chance you'll find the solutions

$$x = -1, \quad y = -9$$

Sum of squared functional values at different x,y values



A biological example

- Data on age-specific birth and death rates for populations = “life table”
- The basic data are:
  - The number of individuals alive each year, starting from birth until all are dead = n_x
  - The number of female offspring per females of age x = b_x
- From this we can calculate “intrinsic population growth rate” (birth rate-death rate) for the population as a whole

r and λ

- Both are measures of population increase
- λ is the finite rate of increase = N_{t+1}/N_t
- r is the instantaneous rate of change = birth rate – death rate
- They are related:

$$\lambda = e^r$$

Life table for a squirrel population

Age	n(x)	b(x)
0	1000	0
1	458	1.28
2	352	2.28
3	229	2.28
4	154	2.28
5	99	2.28
6	87	2.28

Convert number alive to proportion alive (l_x)

	A	B	C	D
1	Age	n(x)	b(x)	lx
2	0	1000	0	1
3	1	458	1.28	0.458
4	2	352	2.28	0.352
5	3	229	2.28	0.229
6	4	154	2.28	0.154
7	5	99	2.28	0.099
8	6	87	2.28	0.087
9				

Multiply proportion alive by birth rate ($l_x b_x$)

	A	B	C	D	E
1	**Age**	**n(x)**	**b(x)**	**lx**	**lxbx**
2	0	1000	0	1	0
3	1	458	1.28	0.458	0.58624
4	2	352	2.28	0.352	0.80256
5	3	229	2.28	0.229	0.52212
6	4	154	2.28	0.154	0.35112
7	5	99	2.28	0.099	0.22572
8	6	87	2.28	0.087	0.19836
9					

Euler's equation

- Growth rate is the balance between birth and death rate,
  r = birth rate – death rate
- If r is positive, the population is growing
- The best estimate of r from a life table is the value that satisfies Euler's equation:

$$1 = \sum l_x b_x e^{-rx}$$

- x is age
- x, l_x, b_x are all known, e is a constant
- This is calculated for each age, summed
- Equation can't be solved analytically, but we can find r numerically with the Solver

In Excel

F8 =E8*EXP(-F$10*A8)

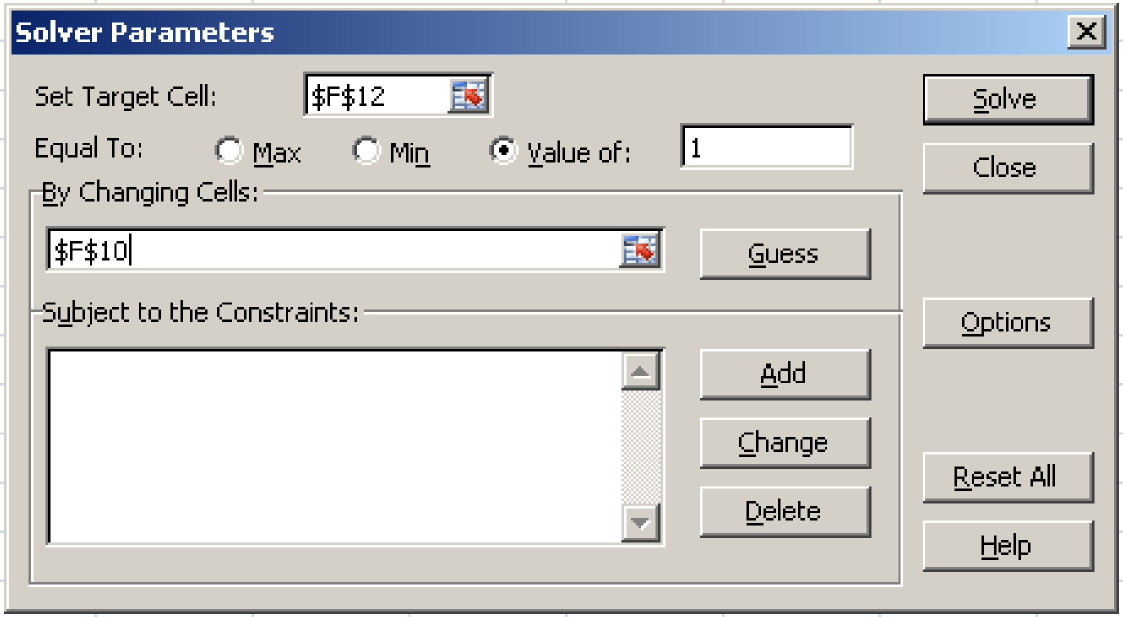
	A	B	C	D	E	F	G
1	Age	n(x)	b(x)	l(x)	l(x)b(x)	Euler	
2	0	1000	0	1	0	0	
3	1	458	1.28	0.458	0.58624	0.58624	
4	2	352	2.28	0.352	0.80256	0.80256	
5	3	229	2.28	0.229	0.52212	0.52212	
6	4	154	2.28	0.154	0.35112	0.35112	
7	5	99	2.28	0.099	0.22572	0.22572	
8	6	87	2.28	0.087	0.19836	0.19836	
9							
10					r	0	
11							
12					Sum Euler	2.68612	
13							

1. Solver will change this...

2. ...until this is equal to 1

Solver setup

	A	B	C	D	E	F
1	Age	n(x)	b(x)	l(x)	l(x)b(x)	Euler
2	0	1000	0	1	0	0
3	1	458	1.28	0.458	0.58624	0.58624
4	2	352	2.28	0.352	0.80256	0.80256
5	3	229	2.28	0.229	0.52212	0.52212
6	4	154	2.28	0.154	0.35112	0.35112
7	5	99	2.28	0.099	0.22572	0.22572
8	6	87	2.28	0.087	0.19836	0.19836
9						
10					r	0
11						
12					Sum Euler	2.68612
13						



Solution

Age	n(x)	b(x)	l(x)	l(x)b(x)	Euler			
0	1000	0	1	0	0			
1	458	1.28	0.458	0.58624	0.3873143			
2	352	2.28	0.352	0.80256	0.350311			
3	229	2.28	0.229	0.52212	0.1505687			
4	154	2.28	0.154	0.35112	0.0668972			
5	99	2.28	0.099	0.22572	0.0284126			
6	87	2.28	0.087	0.19836	0.0164962			
				r	0.4144927			1
				Sum Euler	0.9999999			

Solver Results

Solver found a solution. All constraints and optimality conditions are satisfied.

Reports: Answer, Sensitivity, Limits

- Keep Solver Solution
- Restore Original Values

OK | Cancel | Save Scenario... | Help

r = 0.414 results in next year's population being 1.51 times as big as last year's ($e^{0.414} = 1.51$)

What if... analysis

- We could ask, how low does adult birth rate have to go for the population to stop growing?
- As population size increases females can't get enough food to reproduce successfully
- Assuming survival doesn't change, we can estimate what the reproductive rate would be when the population growth is zero

Set up in Excel

1. Set these to all point to cell c10

C4 | f_x | =C$10

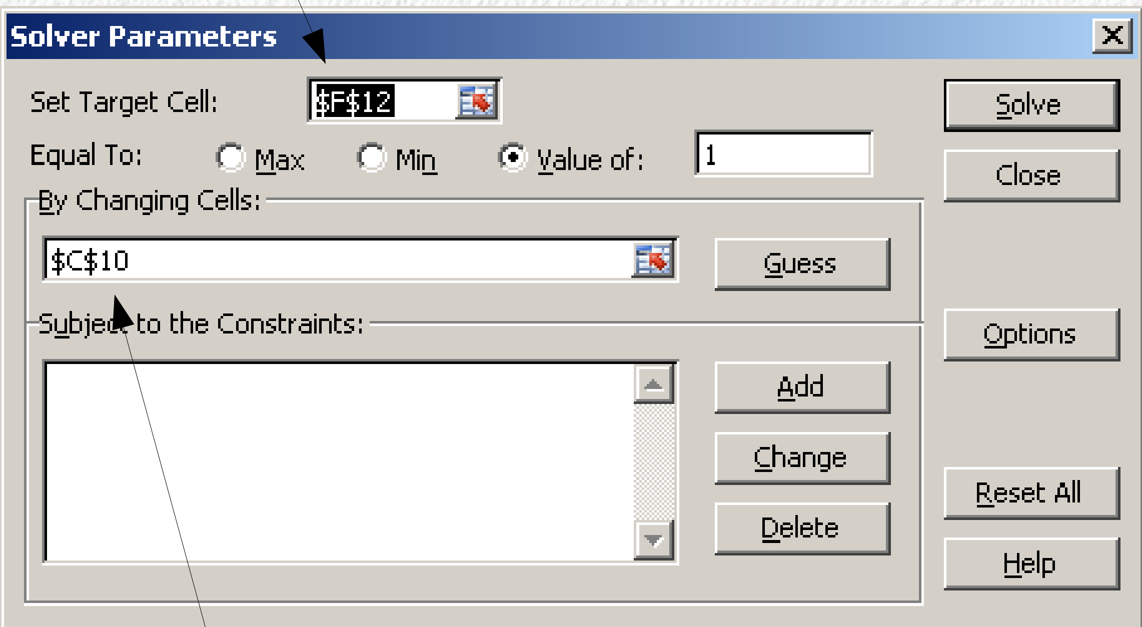
	A	B	C	D	E	F
1	Age	n(x)	b(x)	lx	lxbx	Euler
2	0	1000	0	1	0	0
3	1	458	1.28	0.458	0.58624	0.58624
4	2	352	2.28	0.352	0.80256	0.80256
5	3	229	2.28	0.229	0.52212	0.52212
6	4	154	2.28	0.154	0.35112	0.35112
7	5	99	2.28	0.099	0.22572	0.22572
8	6	87	2.28	0.087	0.19836	0.19836
9						
10	Adult birth rate		2.28		r	0
11						
12					Sum Euler	2.68612
13						
14						

2. Varying this will now change all the adult birth rates

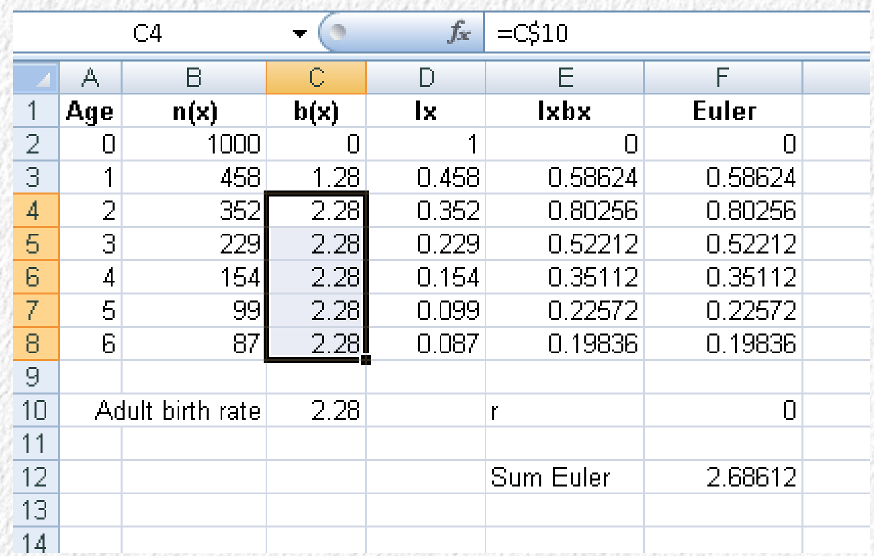
3. Set r to 0, don't vary it

Like before, set the sum of
Euler's equation to 1

Solver setup



But, now change adult birth
rate instead of growth rate



C4 =C$10

	A	B	C	D	E	F
1	Age	n(x)	b(x)	lx	lxbx	Euler
2	0	1000	0	1	0	0
3	1	458	1.28	0.458	0.58624	0.58624
4	2	352	2.28	0.352	0.80256	0.80256
5	3	229	2.28	0.229	0.52212	0.52212
6	4	154	2.28	0.154	0.35112	0.35112
7	5	99	2.28	0.099	0.22572	0.22572
8	6	87	2.28	0.087	0.19836	0.19836
9						
10		Adult birth rate	2.28		r	0
11						
12					Sum Euler	2.68612
13						
14						

Solver's solution

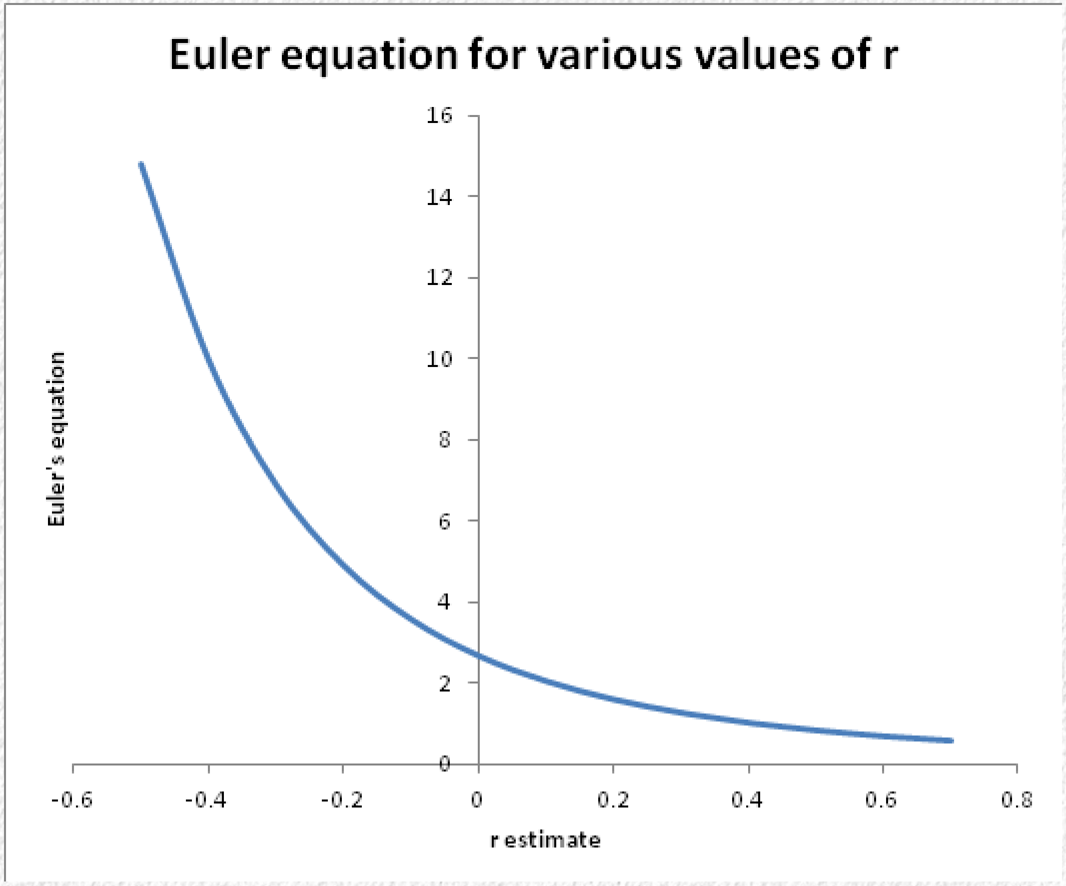
F12 | f_x | =SUM(F2:F8)

	A	B	C	D	E	F
1	**Age**	**n(x)**	**b(x)**	**lx**	**lxbx**	**Euler**
2	0	1000	0	1	0	0
3	1	458	1.28	0.458	0.58624	0.58624
4	2	352	0.4492	0.352	0.1581359	0.1581359
5	3	229	0.4492	0.229	0.10287819	0.10287819
6	4	154	0.4492	0.154	0.06918446	0.06918446
7	5	99	0.4492	0.099	0.04447572	0.04447572
8	6	87	0.4492	0.087	0.03908473	0.03908473
9						
10		Adult birth rate	0.4492		r	0
11						
12					Sum Euler	0.999999
13						
14						

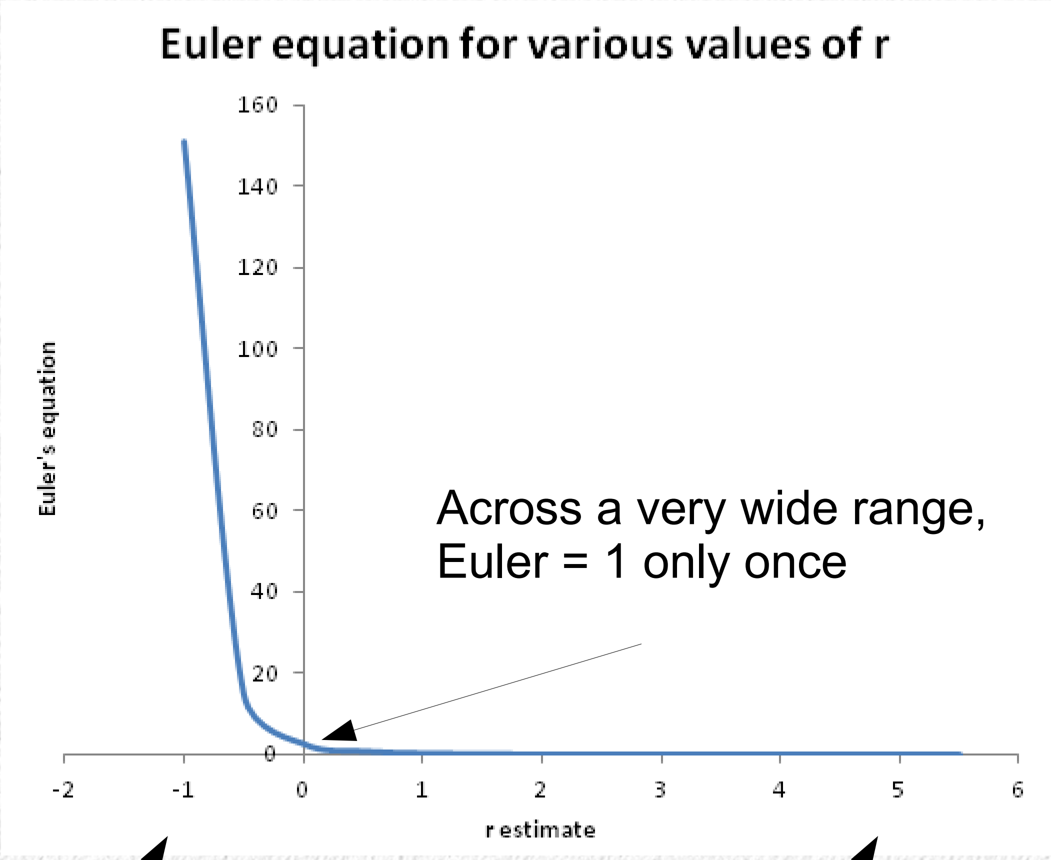
Birth rate would need to be 0.4492 for the population to stop growing

How do we know there is only one solution for r?

- For a numerical result, can't know for sure
- Some ways to check
  - Graph the result across all plausible values of r
  - Try different starting values to see if the solution is always the same



Wider range of possible r's



Across a very wide range,
Euler = 1 only once

40% as many next year as this year

245 times as many next year as this year

Very different starting values always converge on the same solution

	A	B	C	D	E	F	G
1	Age		n(x)	b(x)	l(x)	l(x)b(x)	Euler
2	0		1000	0	1	0	0
3	1		458	1.28	0.458	0.58624	0.0291872
4	2		352	2.28	0.352	0.80256	0.0019893
5	3		229	2.28	0.229	0.52212	6.443E-05
6	4		154	2.28	0.154	0.35112	2.157E-06
7	5		99	2.28	0.099	0.22572	6.905E-08
8	6		87	2.28	0.087	0.19836	3.021E-09
9							
10		Birth rate		2.28		r	3
11							
12						Sum Euler	0.0312432

	A	B	C	D	E	F	G
1	Age		n(x)	b(x)	l(x)	l(x)b(x)	Euler
2	0		1000	0	1	0	0
3	1		458	1.28	0.458	0.58624	11.774945
4	2		352	2.28	0.352	0.80256	323.77581
5	3		229	2.28	0.229	0.52212	4230.7822
6	4		154	2.28	0.154	0.35112	57146.462
7	5		99	2.28	0.099	0.22572	737882.6
8	6		87	2.28	0.087	0.19836	13024311
9							
10		Birth rate		2.28		r	-3
11							
12						Sum Euler	13823907
13							

r	0.4144928
Sum Euler	0.9999996

Optimal foraging

- How do animals choose what to eat and what not to eat?
- A bird foraging in a patch of clams along the shore has a range of sizes to choose from
  - Little ones may be more abundant, easier to open, but give a smaller caloric reward
  - Large ones give a bigger caloric reward, but are less common and harder to break into
- If a bird is trying maximize the profitability of its foraging, what sizes should it eat?

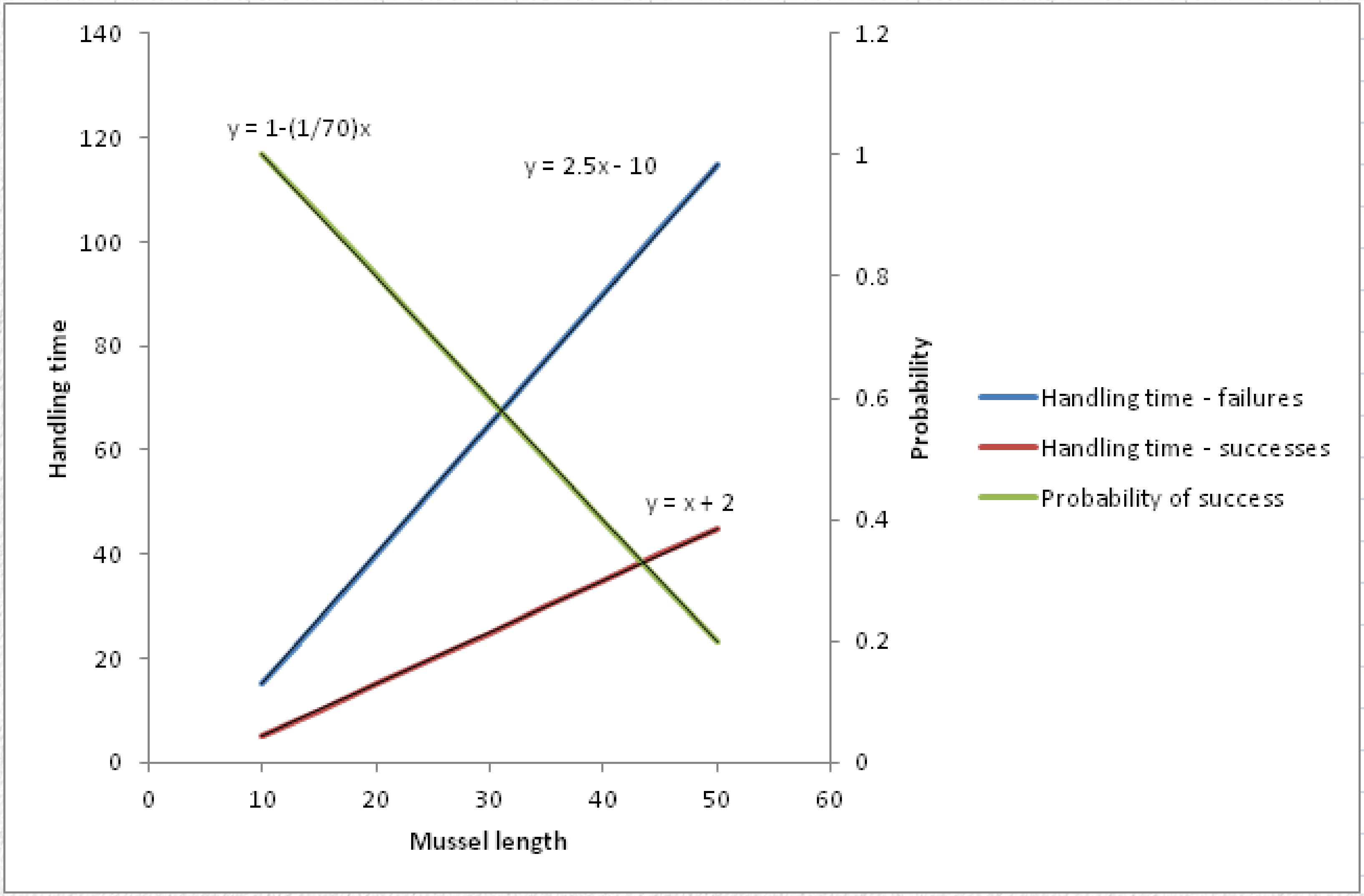
Profitability of foraging

- Profit = revenue – expenses
- Revenue in optimal foraging is usually expressed as energy acquired (in cal or kcal)
- Expenses come from time spent in foraging doing something other than eating
  - Searching for food items
  - "Handling time" = time devoted to opening/consuming the food
- We can measure profit as energy consumed per unit time of foraging – organisms should maximize this quantity

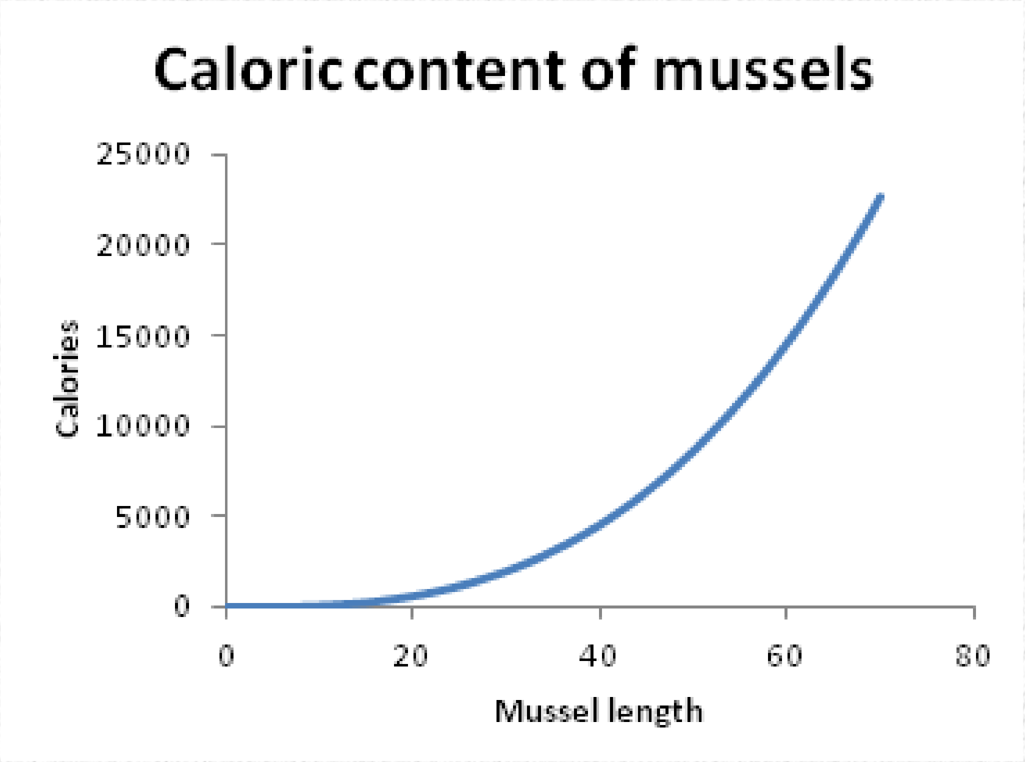
Oystercatchers foraging on oysters

- Oystercatchers eat mussels and other attached shellfish exposed during low tide
- They are easily able to open small mussels
- They are unable to open mussels above 70 mm in length
- The handling time depends on the size of the mussel, and if they are able to successfully open it
  - If they open it, the handling time is H
  - If they try and fail, the handling time is W
- Big mussels have more energy content
- Question is: what size of mussel is most profitable?

Handling time equations – the costs



Caloric content – the benefits



Caloric content (E) increases with length as kcal = $0.12\,(\text{length})^{2.86}$

Equation to maximize

Energy in a mussel of size i

Probability of opening a mussel of size i

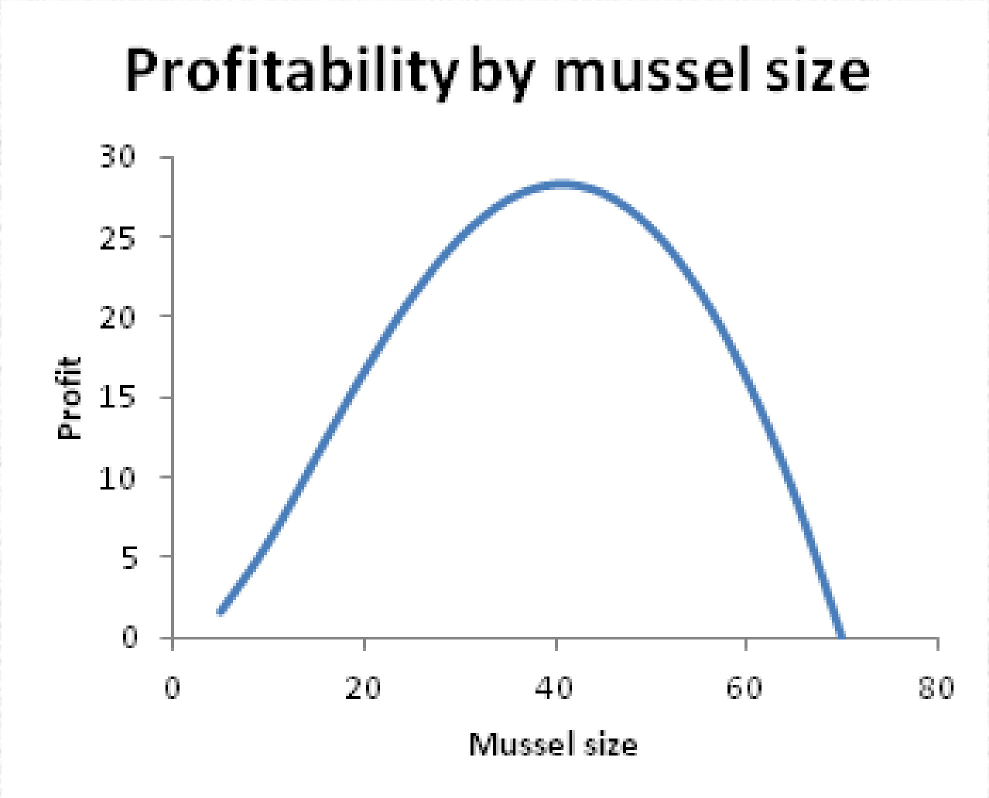
$$Profit = \frac{E_i \times P_i}{H_i \times P_i + W_i \times (1 - P_i)}$$

Time spent opening a mussel of size i

Time wasted failing to open a mussel of size i

Probability of failing to open a mussel of size i

Graph of equation



Setup in Excel

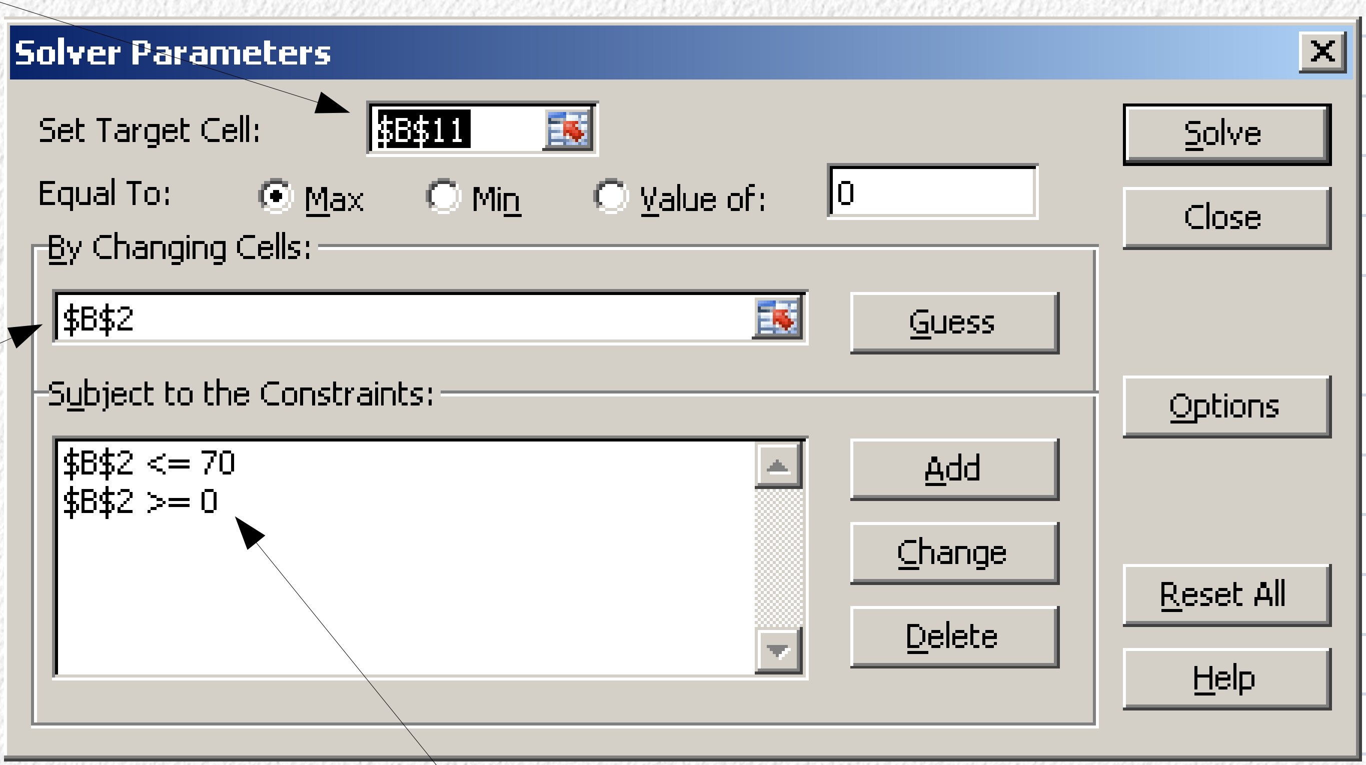
	A	B	C	D	E
1		Solution			
2	Size	35			
3					
4	Terms in the profit equation	Values given size	Cell formulas used		
5	E	3127.633476	=0.12*B2^2.86		
6	P	0.5	=1-(1/70)*B2		
7	H	37	=B2+2		
8	W	77.5	=2.5*B2-10		
9					
10		Value given size			
11	Profit	27.31557621	=(B5*B6)/(B7*B6+B8*(1-B6))		
12					

2. ...by changing this

1. Have solver maximize this...

Solver setup

Maximize
the profit



Try
different
mussel
sizes

Don't bother with sizes that are too big to eat
or negative

Solver's solution

	A	B	C	D	E
1		Solution			
2	Size	40.70668067			
3					
4	Terms in the profit equation	Values given size	Cell formulas used		
5	E	4817.537684	=0.12*B2^2.86		
6	P	0.41847599	=1-(1/70)*B2		
7	H	42.70668067	=B2+2		
8	W	91.76670169	=2.5*B2-10		
9					
10		Value given size			
11	Profit	28.30052884	=(B5*B6)/(B7*B6+B8*(1-B6))		
12					

So, neither the biggest or smallest are most profitable
The optimal size is 40.7 mm
At this size, will only successfully open 42% of mussels
But, bigger mussels are so nutritious that it's worth the high failure rate

Beware of local optima

- What if profit looked like this...
- If oystercatchers do really well on really big mussels, but we stop searching at 70 mm, we will find the smaller peak only
- This can happen!