SWINBURNE UNIVERSITY OF TECHNOLOGY

Non-linear regression and seemingly unrelated regression

A \ Prof. Steve Quinn

Department of Statistics, Data Science and Epidemiology,
Swinburne University of Technology, Melbourne, Australia

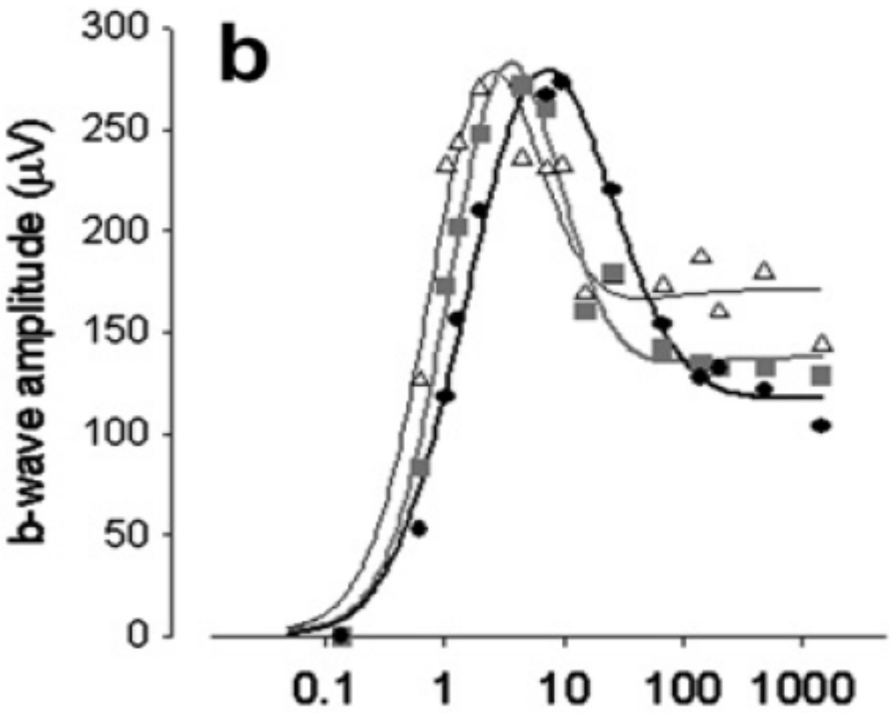
sjquinn@swin.edu.au

Two part analysis

- The problem
  - Trying to identify autism
  - believe that ocular response to a flash at different frequencies is different in autistic vs normal children
  - Want to identify the best flash frequency

Mixed model and Non-linear regression

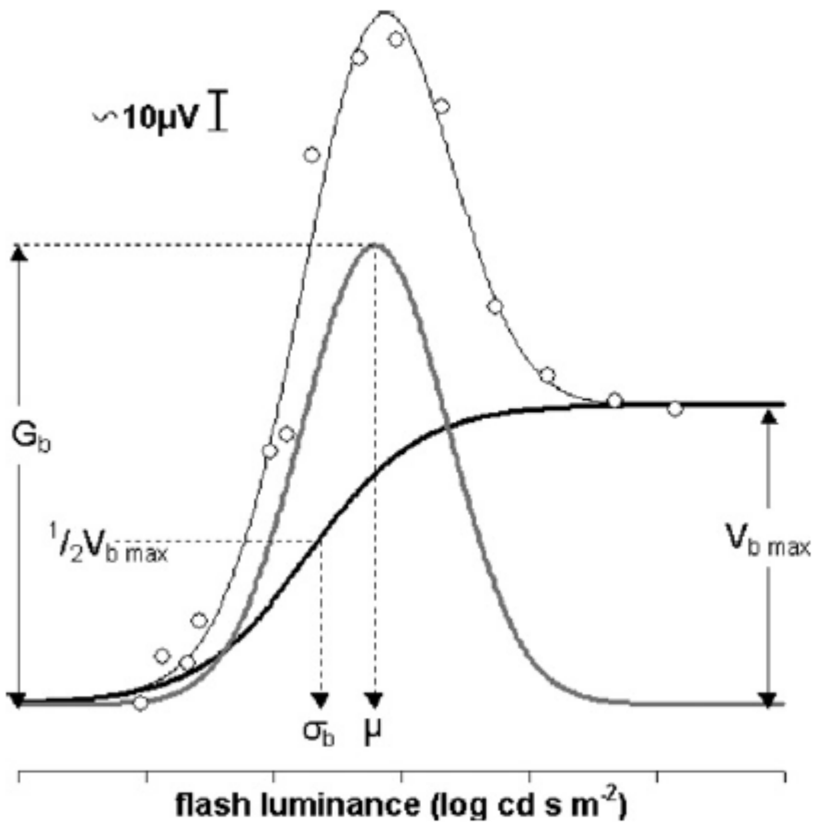
- They know what the data will look like (roughly)



SWINBURNE UNIVERSITY OF TECHNOLOGY

Non-linear regression

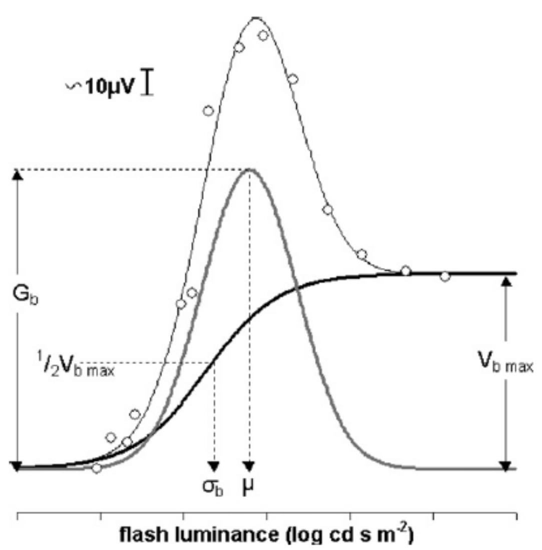
- They also know what the wave is made up of a normal density curve and a logistic curve (a cumulative distribution function)



SWINBURNE UNIVERSITY OF TECHNOLOGY

Non-linear regression

- There is a theoretical framework for this decomposition
- "On-response amplitude" (b-waves) follows the logistic growth function
- "Off-response amplitude" (d-waves) follows Gaussian density function



SWINBURNE UNIVERSITY OF TECHNOLOGY

Mixed model analysis

- The data – two groups – each person contributes 9 observations

	id_num	freq	center	gender	group	eye	vert	b_	agec	irisc
1	1	114	2	0	0	1	4	25.2	-2.854173	-.1162393
2	1	119	2	0	0	1	4	16.1	-2.854173	-.1162393
3	1	367	2	0	0	1	4	8.3	-2.854173	-.1162393
4	1	398	2	0	0	1	4	28.9	-2.854173	-.1162393
5	1	602	2	0	0	1	4	32.3	-2.854173	-.1162393
6	1	799	2	0	0	1	4	32.7	-2.854173	-.1162393
7	1	949	2	0	0	1	4	30.2	-2.854173	-.1162393
8	1	1114	2	0	0	1	4	23.2	-2.854173	-.1162393
9	1	1204	2	0	0	1	4	23.5	-2.854173	-.1162393
10	2	114	2	0	0	0	4	22.2	-2.854173	-.0562393

Mixed model analysis

```
mixed b_  i.group##(c.freq##c.freq##c.freq##c.freq) ///
i.eye i.gender agec || center: || id_num:
```

b_	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.group	1.311512	.8834097	1.48	0.138	-.4199387	3.042964
freq	.0328881	.0009217	35.68	0.000	.0310816	.0346946
c.freq#c.freq	-8.70e-07	2.09e-06	-0.42	0.677	-4.96e-06	3.22e-06
c.freq#c.freq#c.freq	-5.38e-08	5.31e-09	-10.13	0.000	-6.42e-08	-4.34e-08
c.freq#c.freq#c.freq#c.freq	2.96e-11	3.07e-12	9.62	0.000	2.36e-11	3.56e-11
group#c.freq						
1	.0055715	.0013011	4.28	0.000	.0030213	.0081216
group#c.freq#c.freq						
1	2.69e-06	2.93e-06	0.92	0.359	-3.05e-06	8.42e-06
group#c.freq#c.freq#c.freq						
1	-2.50e-08	7.48e-09	-3.34	0.001	-3.96e-08	-1.03e-08
group#c.freq#c.freq#c.freq#c.freq						
1	1.65e-11	4.33e-12	3.81	0.000	8.01e-12	2.50e-11
1.eye	-.3604371	.7547737	-0.48	0.633	-1.839766	1.118892
1.gender	2.677328	.8008358	3.34	0.001	1.107718	4.246937
agec	-.1780289	.0912172	-1.95	0.051	-.3568114	.0007537
_cons	18.93146	1.442409	13.12	0.000	16.10439	21.75853

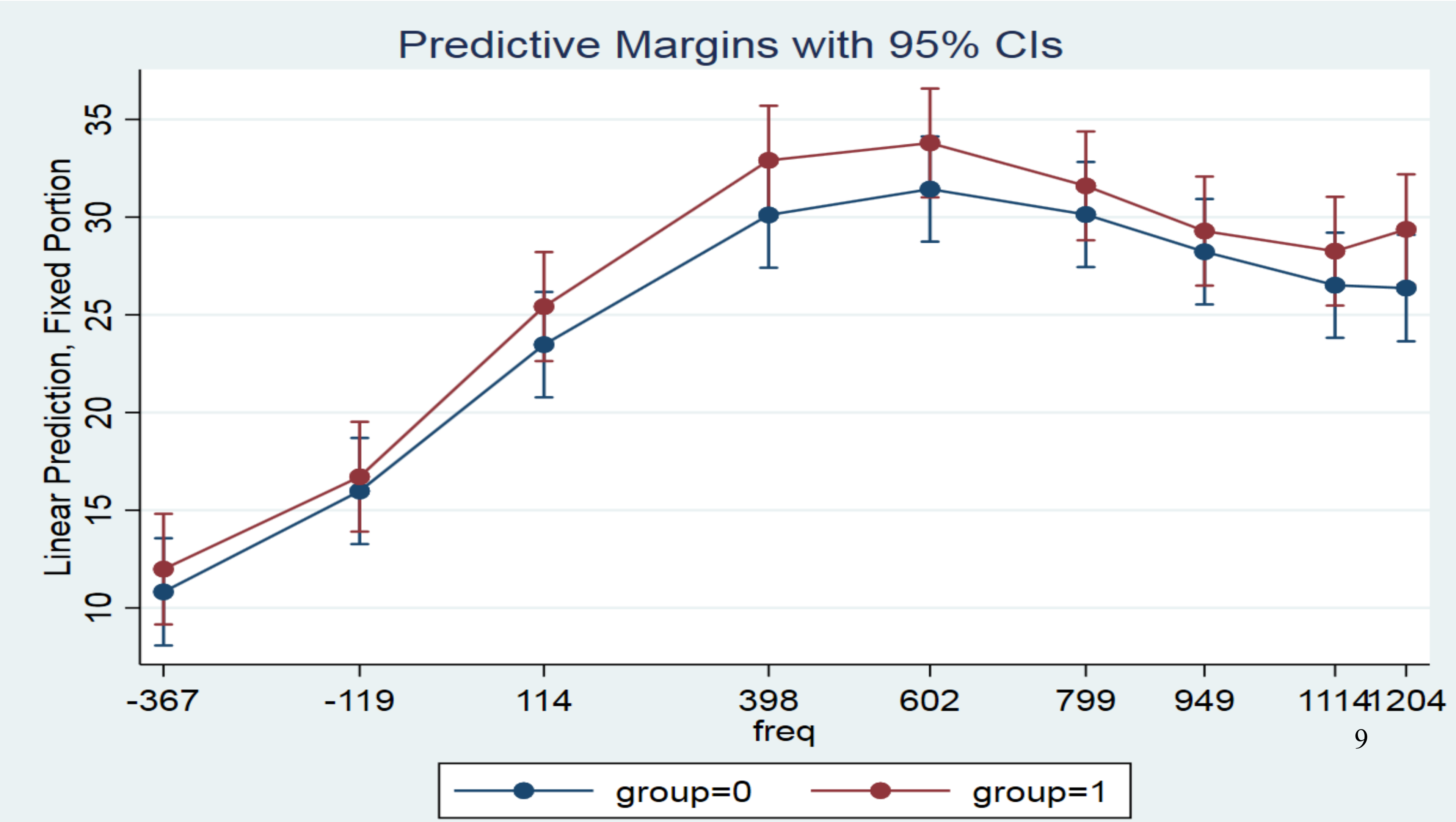
Mixed model analysis

```
margins, at(freq = (-367 -119 114 398 602 799 949 1114 1204) ///
group = (0 1)) noestimcheck
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
_at						
1	10.82212	1.40051	7.73	0.000	8.077169	13.56707
2	15.98201	1.38605	11.53	0.000	13.2654	18.69862
3	23.47454	1.376188	17.06	0.000	20.77726	26.17182
4	30.11157	1.377717	21.86	0.000	27.41129	32.81184
5	31.43626	1.373177	22.89	0.000	28.74488	34.12764
6	30.13359	1.372741	21.95	0.000	27.44307	32.82411
7	28.22864	1.376894	20.50	0.000	25.52997	30.9273
8	26.5165	1.372404	19.32	0.000	23.82664	29.20637
9	26.36978	1.390502	18.96	0.000	23.64444	29.09511
10	11.98417	1.443883	8.30	0.000	9.154214	14.81413
11	16.71393	1.432853	11.66	0.000	13.90559	19.52227
12	25.4219	1.423534	17.86	0.000	22.63182	28.21197
13	32.90578	1.425326	23.09	0.000	30.11219	35.69937
14	33.79462	1.420914	23.78	0.000	31.00968	36.57957
15	31.59931	1.420277	22.25	0.000	28.81562	34.383
16	29.28806	1.424139	20.57	0.000	26.49679	32.07932
17	28.25842	1.41979	19.90	0.000	25.47568	31.04116
18	29.37414	1.437015	20.44	0.000	26.55764	32.19064

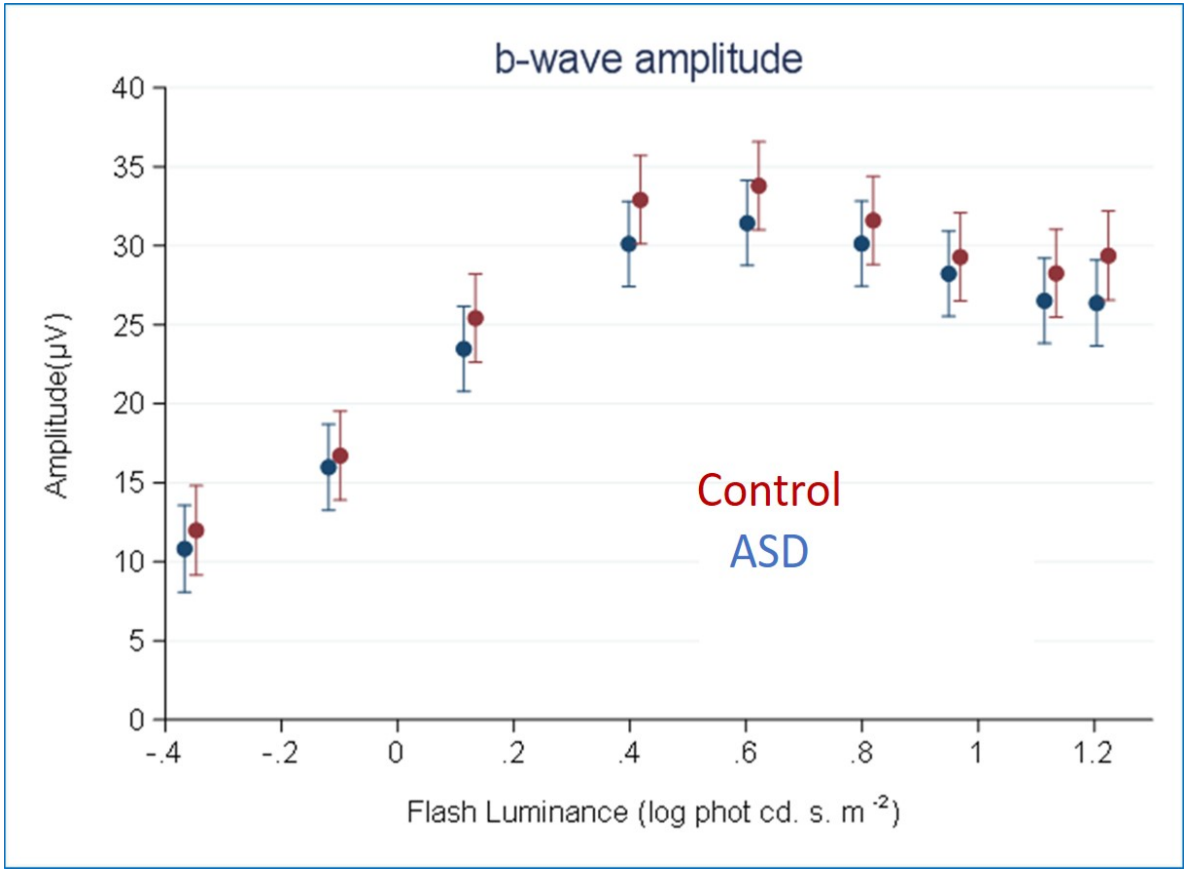
Mixed model analysis

```
marginsplot
```



Mixed model analysis

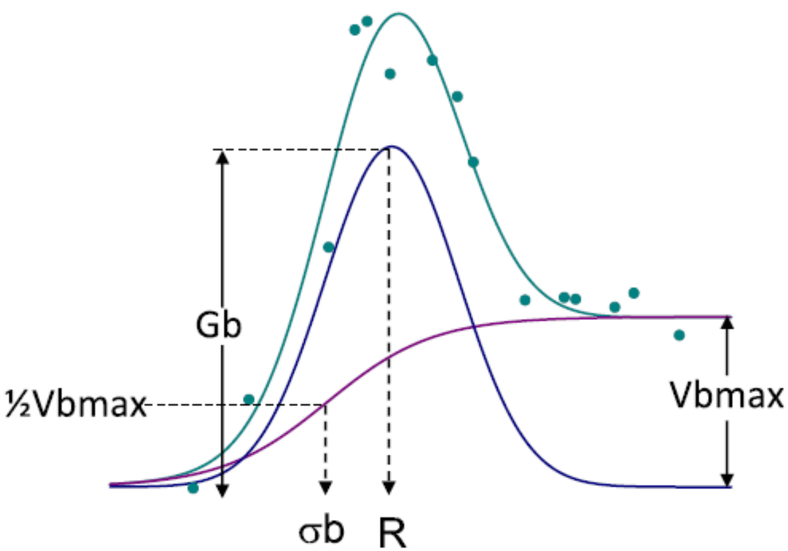
The authors ended up wanting this picture:



SWINBURNE UNIVERSITY OF TECHNOLOGY

Non-linear regression

- The authors wanted to estimate:



$$Vb = Gb\left[\left(\frac{I}{R}\right)^{\frac{\ln\left(\frac{R}{I}\right)}{B^2}}\right] + \frac{Vb\max I}{I + \sigma b}$$

Gaussian — Logistic Growth

Gb : maximal Gaussian amplitude

R : flash luminance at which Gb occurs

B : width of the Gaussian where σ_G is its SD

Vbmax : maximal saturated logistic growth

Non-linear regression first attempt

- The data – one group

	id_num	freq	group	b	id
1	1	.13	0	8.3	1
2	1	.6	0	16.1	2
3	1	1	0	25.2	3
4	1	1.3	0	28.9	4
5	1	2	0	32.3	5
6	1	4.4	0	32.7	6
7	1	7.1	0	30.2	7
8	1	9.6	0	23.2	8
9	1	16	0	23.5	9
10	2	.13	0	13.3	10
11	2	.6	0	14.9	11
12	2	1	0	22.2	12

```
nl (b = {Gb=1}*(freq /{R=1})^(ln({R}/freq)/{B=1}) + ///
{vbmax=1 }*freq /(freq + {sigmab = 1})) , iter(200)
```

Non-linear regression – first attempt

Wrong starting values:

```
Iteration 199:  residual SS =  234090.6
```

Source	SS	df	MS
Model	-45279.098	4	-11319.7744
Residual	234090.63	1534	152.601456
Total	188811.54	1538	122.764327

Number of obs	=	1,539
R-squared	=	-0.2398
Adj R-squared	=	-0.2430
Root MSE	=	12.3532
Res. dev.	=	12100.31

b	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
/Gb	4.592236	123.1835	0.04	0.970	-237.0337	246.2182
/R	1156274	2.85e+09	0.00	1.000	-5.59e+09	5.59e+09
/B	1102.619	216977.5	0.01	0.996	-424501.3	426706.6
/vbmax	41.14959	11.45293	3.59	0.000	18.68453	63.61464
/sigmab	1.280869	.6824065	1.88	0.061	-.057679	2.619417

Parameter R taken as constant term in model & ANOVA table

convergence not achieved

Non-linear regression – SigmaPlot - $1250

f = Gb*((x/R)^(ln(R/x)/b^2))+Vbmax*(x/(x+sigmab))

R	Rsqr	Adj Rsqr	Standard Error of Estimate
0.5540	0.3069	0.3051	9.2365

	Coefficient	Std. Error	t	P
R	2.3149	0.1249	18.5337	<0.0001
b	0.8695	0.0949	9.1604	<0.0001
Gb	8.8577	0.7988	11.0888	<0.0001
sigmab	0.2681	0.0297	9.0330	<0.0001
Vbmax	26.8083	0.6488	41.3170	<0.0001

Constraints

sigmab>0.1
R>1
Vbmax<26

SWINBURNE UNIVERSITY OF TECHNOLOGY

Non-linear regression – Stata

```
nl (b = {Gb=10}*(freq /{R=1})^(ln({R}/freq)/{B=0.1}) ///
+ {vbmax=30 }*freq /(freq + {sigmab = 0.1})) if group==0, iter(100)
```

```
Iteration 62:  residual SS =  130869.8
```

Source	SS	df	MS
Model	1028099.5	5	205619.894
Residual	130869.82	1534	85.3127892
Total	1158969.3	1539	753.066465

Number of obs =	1,539
R-squared =	0.8871
Adj R-squared =	0.8867
Root MSE =	9.236492
Res. dev. =	11205.38

b	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
/Gb	8.857744	.7988016	11.09	0.000	7.290885	10.4246
/R	2.314917	.1249036	18.53	0.000	2.069917	2.559917
/B	.7561121	.1650842	4.58	0.000	.4322975	1.079927
/vbmax	26.80831	.6488434	41.32	0.000	25.53559	28.08102
/sigmab	.2680745	.0296774	9.03	0.000	.2098619	.3262871

Both groups, one analysis - nlsur

Zellner, A. An efficient method of estimating seemingly…
American Statistician Journal, 1962

```
nlsur

(b0 = {Gb0=10}*(freq0 /{R0=1})^(ln({R0}/freq0)/{Bsquare0=0.1})
+ {vbmax0=30 }*freq0 /(freq0 + {sigmab0 = 0.1}))

(b1 = {Gb1=10}*(freq1 /{R1=1})^(ln({R1}/freq1)/{Bsquare1=0.1})
+ {vbmax1=30 }*freq1 /(freq1 + {sigmab1 = 0.1}))


lincom [Gb0]_cons - [Gb1]_cons
lincom [R0]_cons - [R1]_cons
lincom [Bsquare0]_cons - [Bsquare1]_cons
lincom [vbmax0]_cons - [vbmax1]_cons
lincom [sigmab0]_cons - [sigmab1]_cons
```

This example

	id	id_num0	freq0	b0	id_num1	freq1	b1
1	1	1	.13	8.3	178	.13	9.6
2	2	1	.6	16.1	178	.6	12.5
3	3	1	1	25.2	178	1	22.4
4	4	1	1.3	28.9	178	1.3	28.4
5	5	1	2	32.3	178	2	38.5
6	6	1	4.4	32.7	178	4.4	33.5
7	7	1	7.1	30.2	178	7.1	31.6
8	8	1	9.6	23.2	178	9.6	31.3
9	9	1	16	23.5	178	16	33.9
10	10	2	.13	13.3	179	.13	11.2

Both groups, one analysis - nlsur

FGNLS regression

	Equation	Obs	Parms	RMSE	R-sq	Constant
1	b0	1,539	5	9.221563	0.8871*	(none)
2	b1	1,539	5	8.287691	0.9205*	(none)

* Uncentered R-sq

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/Gb0	8.976546	.7960763	11.28	0.000	7.416265	10.53683
/R0	2.321304	.1227957	18.90	0.000	2.080628	2.561979
/Bsquare0	.7546038	.1619968	4.66	0.000	.4370959	1.072112
/vbmax0	26.7926	.6467582	41.43	0.000	25.52497	28.06022
/sigmab0	.2688664	.0296231	9.08	0.000	.2108062	.3269267
/Gb1	11.00446	1.047834	10.50	0.000	8.950747	13.05818
/R1	1.55251	.0268428	57.84	0.000	1.499899	1.605121
/Bsquare1	.1395632	.0302051	4.62	0.000	.0803623	.198764
/vbmax1	31.22634	.3649384	85.57	0.000	30.51108	31.94161
/sigmab1	.3008137	.0229593	13.10	0.000	.2558142	.3458132

SWIN BUR NE
SWINBURNE UNIVERSITY OF TECHNOLOGY

Both groups, one analysis - nlsur

```
. lincom [Gb0]_cons - [Gb1]_cons

 ( 1)  [Gb0]_cons - [Gb1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	-2.027919	1.319276	-1.54	0.124	-4.613653	.5578155

```
. lincom [R0]_cons - [R1]_cons

 ( 1)  [R0]_cons - [R1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	.7687934	.1259884	6.10	0.000	.5218606	1.015726

Robust standard errors

nlsur ….., vce(robust)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
/Gb0	8.976546	.8689462	10.33	0.000	7.273443	10.67965
/R0	2.321304	.1271539	18.26	0.000	2.072087	2.57052
/Bsquare0	.7546038	.1676105	4.50	0.000	.4260932	1.083114
/vbmax0	26.7926	.6779278	39.52	0.000	25.46388	28.12131
/sigmab0	.2688664	.0213295	12.61	0.000	.2270614	.3106715
/Gb1	11.00446	1.10713	9.94	0.000	8.83453	13.1744
/R1	1.55251	.0294145	52.78	0.000	1.494859	1.610161
/Bsquare1	.1395632	.0290729	4.80	0.000	.0825814	.1965449
/vbmax1	31.22634	.3791374	82.36	0.000	30.48325	31.96944
/sigmab1	.3008137	.0175374	17.15	0.000	.2664411	.3351863

Both groups, one analysis - nlsur

```
. lincom [Gb0]_cons - [Gb1]_cons

 ( 1)  [Gb0]_cons - [Gb1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
(1)	-2.027919	1.319276	-1.54	0.124	-4.613653	.5578155

```
. ( 1)  [Gb0]_cons - [Gb1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
(1)	-2.027919	1.452963	-1.40	0.163	-4.875675	.8198372

Seemingly unrelated regression

Provides joint estimates from several regression models

Estimates are more efficient

- accounts for correlated errors
    - Greater correlation increases errors
- Multicollinearity between independent variables increases efficiency
- SE's are smaller

Seemingly unrelated regression – rarely used

Chen. C, et.al. Altered metabolite levels and correlations… *(Metabolomics)* 2017.

n = 158, 113 response variables, 15 covariates

SUR doesn't account for multiple comparisons.

(Benjamini-Hochberg false discovery algorithm)

Outcomes don't need to be the same in kind

Xuecai, Xu, et.al.   Accident severity and traffic signs… *(Accident analysis and prevention)* 2018.

This example

	id	id_num0	freq0	b0	id_num1	freq1	b1
1	1	1	.13	8.3	178	.13	9.6
2	2	1	.6	16.1	178	.6	12.5
3	3	1	1	25.2	178	1	22.4
4	4	1	1.3	28.9	178	1.3	28.4
5	5	1	2	32.3	178	2	38.5
6	6	1	4.4	32.7	178	4.4	33.5
7	7	1	7.1	30.2	178	7.1	31.6
8	8	1	9.6	23.2	178	9.6	31.3
9	9	1	16	23.5	178	16	33.9
10	10	2	.13	13.3	179	.13	11.2

This example

Joint model – nlsur

	Coef.	Std. Err.	z
/Gb0	8.976546	.7960763	11.28
/R0	2.321304	.1227957	18.90
/Bsquare0	.7546038	.1619968	4.66
/vbmax0	26.7926	.6467582	41.43
/sigmab0	.2688664	.0296231	9.08

Separate model – nl

b	Coef.	Std. Err.	t
/Gb	8.857744	.7988016	11.09
/R	2.314917	.1249036	18.53
/B	.7561121	.1650842	4.58
/vbmax	26.80831	.6488434	41.32
/sigmab	.2680745	.0296774	9.03

Questions or comment?

clustered standard errors

```
nl

(b = (group)*({Gb0=10}*(freq /{R0=1})^(ln({R0}/freq)/{Bsquare0=0.1})+
{vbmax0=30 }*freq /(freq + {sigmab0 = 0.1})) +

(1-group)*({Gb1=10}*(freq /{R1=1})^(ln({R1}/freq)/{Bsquare1=0.1})+
{vbmax1=30 }*freq /(freq + {sigmab1 = 0.1})) )

, vce(cluster id_num)
```

b	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
/Gb0	9.363485	.	.	.	.	.
/R0	.2792664	.	.	.	.	.
/Bsquare0	.0066939	.	.	.	.	.
/vbmax0	32.58207	.6637381	49.09	0.000	31.27666	33.88749
/sigmab0	.2236393	.008941	25.01	0.000	.2060544	.2412243
/Gb1	9.314159	.6486444	14.36	0.000	8.038428	10.58989
/R1	1.583685	.0143639	110.25	0.000	1.555434	1.611935
/Bsquare1	.1345217	.0194137	6.93	0.000	.0963395	.1727038
/vbmax1	28.69313	.7084113	40.50	0.000	27.29985	30.08641
/sigmab1	.2979273	.0136946	21.76	0.000	.2709932	.3248614

Seemingly unrelated non-linear regression

The code:

```
        nlsur (b0 = {Gb0=10}*(freq0 /{R0=1})^(ln({R0}/freq0)/{Bsquare0=0.1})
 + {vbmax0=30 }*freq0 /(freq0 + {sigmab0 = 0.1})) ///
        (b1 = {Gb1=10}*(freq1 /{R1=1})^(ln({R1}/freq1)/{Bsquare1=0.1}) + {
vbmax1=30 }*freq1 /(freq1 + {sigmab1 = 0.1})), variables(id_num0 freq0 b0)
```

SWINBURNE UNIVERSITY OF TECHNOLOGY

Seemingly unrelated regression

Seemingly unrelated regression models are so called because they appear to be joint estimates from several regression models, each with its own error term. The regressions are related because the (contemporaneous) errors associated with the dependent variables may be correlated. Chapter 5 of Cameron and Trivedi (2010) contains a discussion of the seemingly unrelated regression model and the feasible generalized least-squares estimator underlying it.

Example 1

When we fit models with the same set of right-hand-side variables, the seemingly unrelated regression results (in terms of coefficients and standard errors) are the same as fitting the models separately (using, say, `regress`). The same is true when the models are nested. Even in such cases, `sureg` is useful when we want to perform joint tests. For instance, let us assume that we think

$$\texttt{price} = \beta_0 + \beta_1 \texttt{foreign} + \beta_2 \texttt{length} + u_1$$

$$\texttt{weight} = \gamma_0 + \gamma_1 \texttt{foreign} + \gamma_2 \texttt{length} + u_2$$

Mixed model analysis

How do you model this curve?

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + h_k(x)(x-a)^k,$$

and $\lim_{x \to a} h_k(x) = 0$. This is called the Peano form of the remainder.

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots.$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

Mixed model analysis

Every function can be modelled as accurately as required by a polynomial curve.

```
        mixed b_  i.group##(c.freq##c.freq##c.freq##c.freq)
i.eye i.gender agec || center: || id_num:,
```

Why did I know to stop at $freq^4$?

– can run the model with freq up to the 5th power and do a LR test

– can run the model with freq up to the 5th power and check the highest terms – they will be non-significant .

SWINBURNE UNIVERSITY OF TECHNOLOGY

Seemingly unrelated non-linear regression

The output:

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/Gb0	8.976546	.7960763	11.28	0.000	7.416265	10.53683
/R0	2.321304	.1227957	18.90	0.000	2.080628	2.561979
/Bsquare0	.7546038	.1619968	4.66	0.000	.4370959	1.072112
/vbmax0	26.7926	.6467582	41.43	0.000	25.52497	28.06022
/sigmab0	.2688664	.0296231	9.08	0.000	.2108062	.3269267
/Gb1	11.00446	1.047834	10.50	0.000	8.950747	13.05818
/R1	1.55251	.0268428	57.84	0.000	1.499899	1.605121
/Bsquare1	.1395632	.0302051	4.62	0.000	.0803623	.198764
/vbmax1	31.22634	.3649384	85.57	0.000	30.51108	31.94161
/sigmab1	.3008137	.0229593	13.10	0.000	.2558142	.3458132

SWINBURNE UNIVERSITY OF TECHNOLOGY

Seemingly unrelated non-linear regression

The code:

```
        nlsur (b0 = {Gb0=10}*(freq0 /{R0=1})^(ln({R0}/freq0)/{Bsquare0=0.1})
 + {vbmax0=30 }*freq0 /(freq0 + {sigmab0 = 0.1})) ///
        (b1 = {Gb1=10}*(freq1 /{R1=1})^(ln({R1}/freq1)/{Bsquare1=0.1}) + {
vbmax1=30 }*freq1 /(freq1 + {sigmab1 = 0.1})), variables(id_num0 freq0 b0)
```

SWINBURNE UNIVERSITY OF TECHNOLOGY

Seemingly unrelated non-linear regression

Comparing the estimates:

```
lincom [Gb0]_cons - [Gb1]_cons
lincom [R0]_cons - [R1]_cons
lincom [Bsquare0]_cons - [Bsquare1]_cons
lincom [vbmax0]_cons - [vbmax1]_cons
lincom [sigmab0]_cons - [sigmab1]_cons
```

```
. lincom [Gb0]_cons - [Gb1]_cons

 ( 1)  [Gb0]_cons - [Gb1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	-2.027919	1.319276	-1.54	0.124	-4.613653	.5578155

```
. lincom [R0]_cons - [R1]_cons

 ( 1)  [R0]_cons - [R1]_cons = 0
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	.7687934	.1259884	6.10	0.000	.5218606	1.015726