Exploring the RL: Which $\hat{v}_j$ tell us about which variance?

Broadly, the argument to follow shows:

- mixed terms with large a_j mostly provide information about σ_s^2
- mixed terms with small a_j mostly provide information about σ_e^2
- free terms mostly provide information about σ_e^2.

$\Rightarrow$ The few mixed terms with large a_j mostly determine $\hat{\sigma}_s^2$ and the extent of smoothing.

This generalization is oversimplified but useful.

I begin with heuristics and then use case deletion and the modified RL.

We end up explaining two of our puzzles.

Heuristics 1: Arm-waving about the log RL

$$-\frac{n - s_X - s_Z}{2}\log(\sigma_e^2) - \frac{1}{2\sigma_e^2}\mathbf{y}'\mathbf{\Gamma}_c\mathbf{\Gamma}_c'\mathbf{y} \quad -\frac{1}{2}\sum_{j=1}^{s_Z}\left[\log(\sigma_s^2 a_j + \sigma_e^2) + \frac{\hat{v}_j^2}{\sigma_s^2 a_j + \sigma_e^2}\right],$$

$$\text{for } \hat{\mathbf{v}} = (\hat{v}_1, \ldots, \hat{v}_{s_Z})' = \mathbf{P}'\mathbf{\Gamma}_Z'\mathbf{y}.$$

The free terms are a function of σ_e^2 only.

- They're like $n - s_X - s_Z$ mixed terms with $a_j = 0$
- $\Rightarrow$ They can affect $\hat{\sigma}_s^2$ and its approximate SE only indirectly.

The mixed terms are functions of both variances.

- The j^{th} mixed term depends on the variances through $\sigma_s^2 a_j + \sigma_e^2$.
- If a_j is small, it's effectively a function of σ_e^2 only.
- If a_j is large, it's dominated by σ_s^2.

Heuristics 2: 1^{st} derivatives of log RL

Call the j^{th} mixed term m_j.

$$\begin{aligned}
\frac{\partial m_j}{\partial \sigma_e^2} &= -0.5(\sigma_s^2 a_j + \sigma_e^2)^{-1} + 0.5\hat{v}_j^2(\sigma_s^2 a_j + \sigma_e^2)^{-2} \\
\frac{\partial m_j}{\partial \sigma_s^2} &= a_j \frac{\partial m_j}{\partial \sigma_e^2}.
\end{aligned}$$

If a_j is small,

- $\partial m_j / \partial \sigma_s^2$ is much closer to zero than $\partial m_j / \partial \sigma_e^2$.
- Both derivatives are insensitive to changes in σ_s^2.
- $\Rightarrow$ The j^{th} mixed term m_j has little effect on $\hat{\sigma}_s^2$.

If a_j is large,

- $\partial m_j / \partial \sigma_s^2$ is much larger than $\partial m_j / \partial \sigma_e^2$.
- Both derivatives are sensitive to changes in σ_s^2.

Heuristics 3: 2^{nd} derivatives of log RL

$$
\begin{aligned}
\frac{\partial^2 m_j}{\partial(\sigma_e^2)^2} &= 0.5(\sigma_s^2 a_j + \sigma_e^2)^{-2} - \hat{v}_j^2(\sigma_s^2 a_j + \sigma_e^2)^{-3} \\
\frac{\partial^2 m_j}{\partial\sigma_e^2 \partial\sigma_s^2} &= a_j \frac{\partial^2 m_j}{\partial(\sigma_e^2)^2} \\
\frac{\partial^2 m_j}{\partial(\sigma_s^2)^2} &= a_j^2 \frac{\partial^2 m_j}{\partial(\sigma_e^2)^2}
\end{aligned}
$$

Evaluate these at the max-RL estimates, multiply by -1, sum over j
$\Rightarrow$ approx precision$(\hat{\sigma}_e^2, \hat{\sigma}_s^2)$.

If a_j is small,

- m_j makes a negligible contribution to $\partial^2 m_j / \partial(\sigma_s^2)^2$.

If a_j is large,

- m_j makes a much larger contribution to prec$(\hat{\sigma}_s^2)$ than to prec$(\hat{\sigma}_e^2)$.

Examining the RL using case-deletion diagnostics for GLMs

The RL is the likelihood from a gamma-errors GLM, with

- s_Z "observations" for the s_Z mixed terms
- 1 "observations" for the free terms

We'll use GLM machinery to see the effect of deleting each "observation"

- Deleting the j^{th} mixed term $\equiv$ making $(\mathbf{\Gamma}_Z \mathbf{P})_j$ a fixed effect instead of a random effect, which we might actually do.
- Deleting the free terms $\equiv$ reducing the dataset to the mixed-term $\hat{v}_j$.

The usual case-deletion diagnostic uses a linear approximation to the GLM to approximate the effect of deleting a case.

We'll consider the effects of case deletion on $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ separately, rather than on the 2-vector of parameters jointly.

Let $\mathbf{W}$ be the diagonal weight matrix evaluated at the full-data estimates $(\tilde{\sigma}_s^2, \tilde{\sigma}_e^2)$, with entries $\text{var}(\hat{v}_j^2)^{-1}$.

$$\begin{aligned} W_j &= \frac{1}{2}(\tilde{\sigma}_e^2 + \tilde{\sigma}_s^2 a_j)^{-2}, \quad j = 1, \ldots, s_Z \\ W_{s_Z+1} &= \frac{n - s_X - s_Z}{2}(\tilde{\sigma}_e^2)^{-2}. \end{aligned}$$

The GLM's design matrix $\mathbf{B}$ has rows $\mathbf{B}_j = [1 \quad a_j]$; $a_{s_Z+1} = 0$ for free terms.

Apply the usual case-deletion formula to the approximate linear model

$$\mathbf{W}^{1/2}\hat{\mathbf{v}}^2 = \mathbf{W}^{1/2}\mathbf{B}\left[\begin{array}{c} \sigma_e^2 \\ \sigma_s^2 \end{array}\right] + \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{I}_{s_Z+1})$$

The one-step approx change in $(\hat{\sigma}_e^2, \hat{\sigma}_s^2)$ from deleting $\hat{v}_j^2$ is the 2-vector

$$\begin{aligned} & W_j^{1/2}(\hat{v}_j^2 - \tilde{\sigma}_s^2 a_j - \tilde{\sigma}_e^2) \\ \times \quad & (-W_j^{1/2}(\mathbf{B}'\mathbf{W}\mathbf{B})^{-1}\mathbf{B}_j')/(1 - W_j\mathbf{B}_j(\mathbf{B}'\mathbf{W}\mathbf{B})^{-1}\mathbf{B}_j'); \end{aligned}$$

The one-step approx change in $(\hat{\sigma}_e^2, \hat{\sigma}_s^2)$ from deleting $\hat{v}_j^2$ is the 2-vector

$$W_j^{1/2}(\hat{v}_j^2 - \tilde{\sigma}_s^2 a_j - \tilde{\sigma}_e^2) \quad (1)$$

$$\times \quad (-W_j^{1/2}(\mathbf{B}'\mathbf{W}\mathbf{B})^{-1}\mathbf{B}_j')/(1 - W_j\mathbf{B}_j(\mathbf{B}'\mathbf{W}\mathbf{B})^{-1}\mathbf{B}_j'); \quad (2)$$

- (1) is the standardized (Pearson) residual for case j.
- (2) is a 2-vector multiplier of the standardized residual, with one row each for $\hat{\sigma}_e^2$ and $\hat{\sigma}_s^2$.

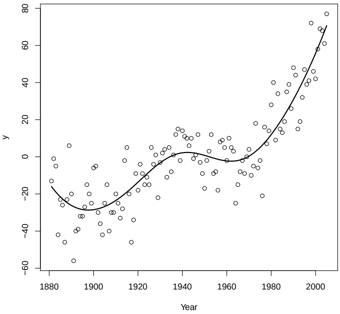
H2013 Sec 16.1.2 examines (2) in painful detail, which I'll spare you.

Key facts re (2): As j increases from 1 to $s_z + 1$

- (2)'s entry for $\hat{\sigma}_e^2$ decreases from positive to negative.
- (2)'s entry for $\hat{\sigma}_s^2$ increases from negative to positive.
- They have opposite signs for *most* j, both negative for *some* j.

Example: Penalized spline fit to the GMST data

$n = 125$, truncated quadratic basis, 30 knots at years $1880 + \{4, \ldots, 120\}$



This penalized spline can be understood as

- a quadratic regression with unshrunk coefficients
- + regression on higher-order polynomials with shrunken coefficients
- and the extent of shrinkage increases with the polynomial order.

Following are the case-deletion diagnostics and their components.

			scaled	Multiplier for	
j	a_j	$\hat{v}_j^2$	resid	σ_e^2	σ_s^2
1	3.6(1)	2856	-0.65	0.81	-727
2	3.2(0)	5248	0.48	0.63	-666
3	5.6(-1)	815	0.14	0.05	-475
4	1.5(-1)	409	0.31	-0.69	-252
5	4.9(-2)	8	-0.68	-1.23	-116
6	2.0(-2)	335	0.74	-1.51	-51.0
7	8.8(-3)	6	-0.68	-1.64	-22.5
8	4.3(-3)	497	1.65	-1.71	-9.67
9	2.3(-3)	15	-0.63	-1.73	-3.58
10	1.3(-3)	248	0.49	-1.75	-0.51
11	7.7(-4)	461	1.53	-1.76	1.13
12	4.8(-4)	119	-0.13	-1.76	2.04
16	9.5(-5)	429	1.38	-1.77	3.23
19	3.6(-5)	8	-0.67	-1.77	3.41
26	7.0(-6)	588	2.15	-1.77	3.51
29	5.1(-6)	78	-0.33	-1.77	3.51
30	3.7(-6)	365	1.07	-1.77	3.52
31	0	137	-0.37	-78.31	156.15

				Estimate's change from			
			scaled	$\tilde{\sigma}_e^2 = 145.3$		$\tilde{\sigma}_s^2 = 947.7$	
j	a_j	$\hat{v}_j^2$	resid	1-step	exact	1-step	exact
1	3.6(1)	2856	-0.65	-0.53	-0.45	471.1	448.2
2	3.2(0)	5248	0.48	0.30	0.92	-318.1	-453.8
3	5.6(-1)	815	0.14	0.01	0.03	-68.2	-74.7
4	1.5(-1)	409	0.31	-0.21	-0.21	-77.5	-79.0
5	4.9(-2)	8	-0.68	0.83	0.76	78.4	94.7
6	2.0(-2)	335	0.74	-1.12	-1.11	-37.7	-38.2
7	8.8(-3)	6	-0.68	1.12	1.12	15.3	16.3
8	4.3(-3)	497	1.65	-2.81	-2.81	-15.9	-15.8
9	2.3(-3)	15	-0.63	1.10	1.10	2.3	1.8
10	1.3(-3)	248	0.49	-0.86	-0.86	-0.2	0.2
11	7.7(-4)	461	1.53	-2.68	-2.69	1.7	3.3
12	4.8(-4)	119	-0.13	0.23	0.23	-0.3	-0.4
16	9.5(-5)	429	1.38	-2.44	-2.44	4.4	6.3
19	3.6(-5)	8	-0.67	1.18	1.19	-2.3	-3.1
26	7.0(-6)	588	2.15	-3.81	-3.82	7.5	10.6
29	5.1(-6)	78	-0.33	0.58	0.58	-1.1	-1.6
30	3.7(-6)	365	1.07	-1.90	-1.90	3.8	5.2
31	0	137	-0.37	28.86	29.06	-57.5	-66.0

Points about the previous two pages of tables

The one-step approximation is pretty good.

This supports the heuristics on earlier slides.

- The multiplier for σ_e^2 starts $+$ and $\downarrow$ to $-$.
- The multiplier for σ_s^2 starts $-$ and $\uparrow$ to $+$.
- Both multipliers are negative for $j = 4, \ldots, 10$.

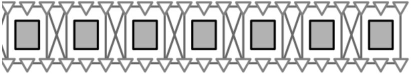
$\Rightarrow$ The influence on

- σ_s^2 is in the "right" direction and largest for small j.
- σ_e^2 is in the "right" direction and largest for large j.

Note: Pearson residuals are necessarily $>$ $\approx$ -0.7.

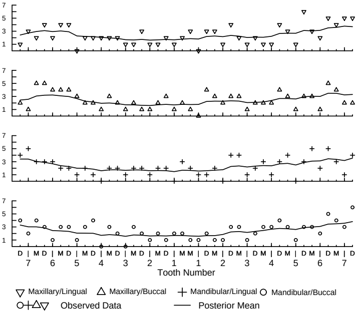
Example: Simple ICAR model for periodontal data

ICAR with these neighbor pairs for $n = 168$ sites.



Priors: Flat on the two island (arch) means; σ_e^2 and $\sigma_s^2 \sim$ IG(0.01,0.01).

Posterior medians: σ_e^2 1.25, σ_s^2 0.25, σ_e^2/σ_s^2 4.0 – very smooth fit.



j	a_j	$\hat{v}_j^2$	scaled resid	Multiplier for σ_e^2	σ_s^2
1	149.0	2.9	-0.7	0.028	-0.057
2	149.0	4.9	-0.6	0.028	-0.057
3	37.3	34.3	1.6	0.023	-0.050
4	37.3	28.2	1.2	0.023	-0.050
5	16.6	0.2	-0.7	0.017	-0.041
6	16.6	1.0	-0.6	0.017	-0.041
7	9.4	2.4	-0.2	0.012	-0.033
8	9.4	3.2	-0.1	0.012	-0.033
9	6.1	3.6	0.2	0.007	-0.027
10	6.1	0.7	-0.5	0.007	-0.027
16	2.4	6.2	1.6	-0.003	-0.013
17	1.9	4.3	1.0	-0.005	-0.011
30	0.8	4.1	1.3	-0.010	-0.003
31	0.8	7.9	3.1	-0.010	-0.003
83	0.3	35.3	17.8	-0.014	0.002
84	0.3	45.1	22.9	-0.014	0.002
94	0.2	4.4	1.7	-0.015	0.003
115	0.2	4.2	1.5	-0.015	0.003

j	a_j	$\hat{v}_j^2$	scaled resid	Multiplier for σ_e^2	σ_s^2
128	0.2	1.8	0.3	-0.015	0.003
129	0.2	6.5	2.8	-0.015	0.003
147	0.2	0.5	-0.5	-0.015	0.003
148	0.2	0.0	-0.7	-0.015	0.003
149	0.2	1.0	-0.1	-0.015	0.003
150	0.2	0.1	-0.7	-0.015	0.003
151	0.2	0.9	-0.2	-0.015	0.003
152	0.2	0.1	-0.7	-0.015	0.003
153	0.2	0.0	-0.7	-0.015	0.003
154	0.2	0.0	-0.7	-0.015	0.003
155	0.2	0.2	-0.6	-0.015	0.003
156	0.2	0.6	-0.4	-0.015	0.003
157	0.2	0.9	-0.2	-0.015	0.003
158	0.2	0.1	-0.7	-0.015	0.003
159	0.2	5.4	2.2	-0.015	0.003
164	0.2	2.7	0.7	-0.015	0.003
165	0.2	0.2	-0.6	-0.015	0.003
166	0.2	1.0	-0.2	-0.015	0.003

				Estimate's change from			
			scaled	$\tilde{\sigma}_e^2 = 1.26$		$\tilde{\sigma}_s^2 = 0.25$	
j	a_j	$\hat{v}_j^2$	resid	1-step	exact	1-step	exact
1	149.0	2.9	-0.7	-0.018	-0.018	0.037	0.037
2	149.0	4.9	-0.6	-0.017	-0.017	0.035	0.035
3	37.3	34.3	1.6	0.036	0.048	-0.078	-0.094
4	37.3	28.2	1.2	0.027	0.033	-0.058	-0.067
5	16.6	0.2	-0.7	-0.012	-0.013	0.028	0.030
6	16.6	1.0	-0.6	-0.010	-0.011	0.024	0.026
7	9.4	2.4	-0.2	-0.003	-0.003	0.008	0.009
8	9.4	3.2	-0.1	-0.001	-0.001	0.003	0.003
9	6.1	3.6	0.2	0.001	0.002	-0.006	-0.006
10	6.1	0.7	-0.5	-0.004	-0.005	0.014	0.016
16	2.4	6.2	1.6	-0.004	-0.005	-0.022	-0.021
17	1.9	4.3	1.0	-0.005	-0.005	-0.011	-0.011
30	0.8	4.1	1.3	-0.013	-0.013	-0.004	-0.004
31	0.8	7.9	3.1	-0.031	-0.031	-0.009	-0.009
83	0.3	35.3	17.8	-0.242	-0.253	0.028	0.043
84	0.3	45.1	22.9	-0.312	-0.329	0.036	0.060
94	0.2	4.4	1.7	-0.024	-0.025	0.005	0.006
115	0.2	4.2	1.5	-0.022	-0.023	0.005	0.005

			scaled	Estimate's change from			
				$\tilde{\sigma}_e^2 = 1.26$		$\tilde{\sigma}_s^2 = 0.25$	
j	a_j	$\hat{v}_j^2$	resid	1-step	exact	1-step	exact
128	0.2	1.8	0.3	-0.004	-0.004	0.001	0.001
129	0.2	6.5	2.8	-0.041	-0.042	0.009	0.010
147	0.2	0.5	-0.5	0.007	0.007	-0.001	-0.002
148	0.2	0.0	-0.7	0.010	0.011	-0.002	-0.002
149	0.2	1.0	-0.1	0.002	0.002	0.000	0.000
150	0.2	0.1	-0.7	0.010	0.010	-0.002	-0.002
151	0.2	0.9	-0.2	0.003	0.003	-0.001	-0.001
152	0.2	0.1	-0.7	0.010	0.010	-0.002	-0.002
153	0.2	0.0	-0.7	0.010	0.011	-0.002	-0.002
154	0.2	0.0	-0.7	0.010	0.010	-0.002	-0.002
155	0.2	0.2	-0.6	0.009	0.009	-0.002	-0.002
156	0.2	0.6	-0.4	0.006	0.006	-0.001	-0.001
157	0.2	0.9	-0.2	0.004	0.004	-0.001	-0.001
158	0.2	0.1	-0.7	0.010	0.010	-0.002	-0.002
159	0.2	5.4	2.2	-0.033	-0.034	0.007	0.008
164	0.2	2.7	0.7	-0.011	-0.011	0.002	0.003
165	0.2	0.2	-0.6	0.009	0.009	-0.002	-0.002
166	0.2	1.0	-0.2	0.003	0.003	-0.001	-0.001

Points about the previous four pages of tables

No free terms in this example, only mixed terms.

Broadly supports the heuristics; the a_j decline more slowly than for the spline, $\Rightarrow$ $\hat{v}_j^2$ not as strongly differentiated in their effects.

Again, the influence on

- σ_s^2 is in the "right" direction and largest for small j.
- σ_e^2 is in the "right" direction and largest for large j.

$j = 3, 4$ are influential for σ_s^2:

- Canonical predictors are quadratic (bigger PD in back).
- Largest multiplier; large-ish scaled residuals $\Rightarrow$ largest influence.

$j = 83, 84$ are influential for σ_e^2:

- Canonical predictor is the contrast of direct vs. interproximal sites.
- Huge + scaled residuals; near-maximal multiplier $\Rightarrow$ big influence.
- These also have the second-largest influence on σ_s^2.

Puzzle #1 explained: Those contrasts tell us about σ_e^2

When $j = 83, 84$ were made fixed effects, the posterior medians did this:

- $\sigma_e^2 \downarrow$ 1.25 to 0.63; $\sigma_s^2 \uparrow$ 0.25 to 0.41
- ratio σ_e^2/σ_s^2: $\downarrow$ 4.0 to 1.6
- DF in the fit: $\uparrow$ 23.5 to 45.6 (2 are for the new FEs).

These contrasts, for direct vs. interprox sites,

- have small a_j (0.33) $\Rightarrow$ mostly inform about σ_e^2.
- have huge positive scaled residuals $\Rightarrow$ when changed to FEs, $\hat{\sigma}_e^2 \downarrow$.

It seems the increase in $\hat{\sigma}_s^2$ is mostly secondary to the reduction in $\hat{\sigma}_e^2$.

	Estimate's change from			
v_j changed	$\tilde{\sigma}_e^2 = 1.26$		$\tilde{\sigma}_s^2 = 0.25$	
to fixed effects	1-step	exact	1-step	exact
only $j = 83$	-0.242	-0.253	0.028	0.043
only $j = 84$	-0.312	-0.329	0.036	0.060
$j = 83$ and 84	-0.558	-0.609	0.065	0.147

Examining the RL using the modified RL

$$-\frac{1}{2}\sum_{j=1}^{s_Z}\left[\log(\sigma_s^2 a_j + \sigma_e^2) + \frac{\hat{v}_j^2}{\sigma_s^2 a_j + \sigma_e^2}\right] \quad -\frac{n - s_X - s_Z}{2}\log(\sigma_e^2) - \frac{1}{2\sigma_e^2}\mathbf{y}'\mathbf{\Gamma}_c\mathbf{\Gamma}_c'\mathbf{y}$$

$$\text{for } \hat{\mathbf{v}} = (\hat{v}_1, \ldots, \hat{v}_{s_Z})' = \mathbf{P}'\mathbf{\Gamma}_Z'\mathbf{y}.$$

I considered modified RLs that include only $j = 1, \ldots, m$ for various m.
($j = s_Z + 1$ is the free terms for $\hat{\sigma}_e^2$)

This is like case deletion but

- the calculations are exact
- they include the approximate SEs for $\hat{\sigma}_e^2$ and $\hat{\sigma}_s^2$.

(Approx SE: $-1\times$ the hessian, invert it, sqrt of the diagonal.)

The main point I want you to take from this

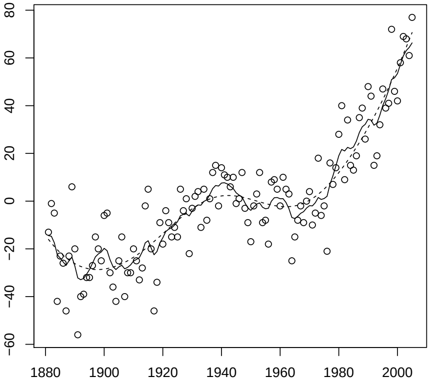
The approximate SEs (posterior SDs) for σ_s^2, σ_e^2 in these modified RLs support the heuristics:

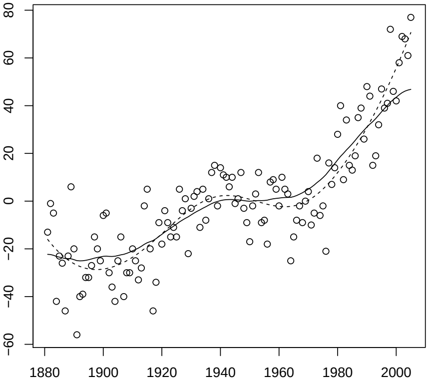
- σ_s^2: The approx SE declines quickly for small m (as small j are added), but not at all for later m.
- σ_e^2: The approx SE declines for all m (as all j are added), with a big drop when the free terms are added (if there are any).

Puzzle #3 explained (different smooths of GSMT data)

Left: Max RL for penalized spline (dash, 6.7 DF), ICAR (solid, 26.5 DF)

Right: Penalized spline (dash, 6.7 DF); ICAR with 6.7 DF (solid)



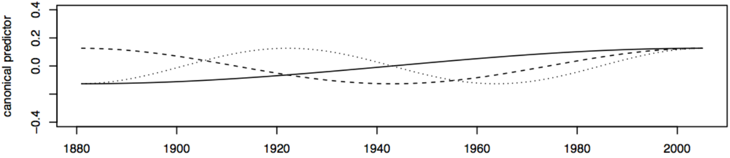


If you force the ICAR fit to have 6.7 DF, it allocates them stupidly.

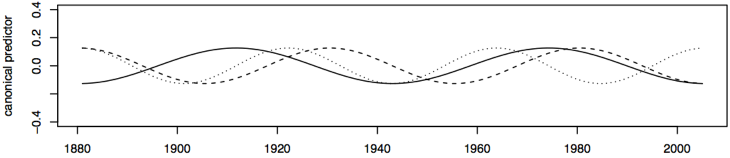
Let's apply our tools to this model & dataset.

Canonical predictors are ≈ linear, quadratic, and then like the p-spline's.

For $a_1 = 1583$ (solid), $a_2 = 396$ (dashed), $a_3 = 176$ (dotted)



For $a_4 = 99$ (solid), $a_5 = 63$ (dashed), $a_6 = 44$ (dotted)



Note: $a_1/a_6 \sim 36$, cf the spline (1,841) and the perio ICAR (35).

The case-delection diagnostics have a few tantalizing hints.

j	a_j	$\hat{v}_j^2$	scaled resid	Multiplier σ_e^2	σ_s^2	Exact change from $\tilde{\sigma}_e^2 = 119.9$	$\tilde{\sigma}_s^2 = 21.8$
1	1583.2	52872	0.4	2.4	-3.0	1.2	-1.4
2	395.9	2657	-0.5	2.3	-2.9	-1.4	1.7
3	176.0	7388	0.6	2.2	-2.9	1.9	-2.2
4	99.0	7732	1.7	2.1	-2.8	5.4	-6.0
5	63.4	1	-0.7	2.0	-2.7	-1.8	2.2
6	44.1	15	-0.7	1.8	-2.6	-1.7	2.1
7	32.4	57	-0.7	1.6	-2.4	-1.5	1.9
8	24.8	146	-0.6	1.4	-2.3	-1.1	1.5
9	19.6	4	-0.7	1.3	-2.2	-1.3	1.8
10	15.9	494	0.0	1.1	-2.0	0.1	-0.1
11	13.2	240	-0.3	0.9	-1.9	-0.4	0.6
12	11.1	1078	1.4	0.7	-1.8	1.3	-2.7
28	2.1	926	3.2	-1.1	-0.5	-3.4	-1.6
30	1.8	906	3.3	-1.2	-0.4	-3.8	-1.3
47	0.8	1214	5.5	-1.7	0.0	-9.9	0.2
59	0.5	964	4.5	-1.9	0.1	-9.0	0.9
69	0.4	1690	8.5	-2.0	0.2	-18.8	2.9

If we change $j = 1, 2$ ($\approx$ linear and quadratic) into fixed effects,
this model still wildly overfits:

- $\hat{\sigma}_e^2 = 119.5$, $\hat{\sigma}_s^2 = 22.2$, DF = 26.8 DF (3 FE, 23.8 RE)

BUT NOW, the only real difference between this model and the p-spline is the a_j, which decline much faster in j for the p-spline than the ICAR.

Recall: the DF in the fitted coefficient v_j of canonical predictor j:
$a_j/(a_j + \sigma_e^2/\sigma_s^2)$

$\Rightarrow$ DF in v_j declines more quickly in j for the spline, which explains

- our mystery
- the specific sense in which the ICAR allocates its DF stupidly.

(Next slide)

DF allocated to each FE and canonical predictor under these models

Polynomial order	Spline DF = 6.7	ICAR DF = 26.5	ICAR DF = 6.7	j
Intercept	1	1	1	FE
Linear	1	0.997	0.94	1
Quadratic	1	0.99	0.80	2
Cubic	0.996	0.97	0.63	3
Quartic	0.95	0.95	0.49	4
Quintic	0.79	0.92	0.38	5
(etc.)	0.49	0.89	0.30	6
	0.24	0.85	0.24	7
	0.11	0.82	0.20	8
	0.05	0.78	0.16	9
	0.03	0.74	0.14	10
	0.0002	0.42	0.04	20
	0.00002	0.25	0.018	30
		0.12	0.007	50
		0.05	0.003	100
		0.04	0.002	124