ODDITY #2: Introducing spatially-correlated errors makes a fixed effect disappear

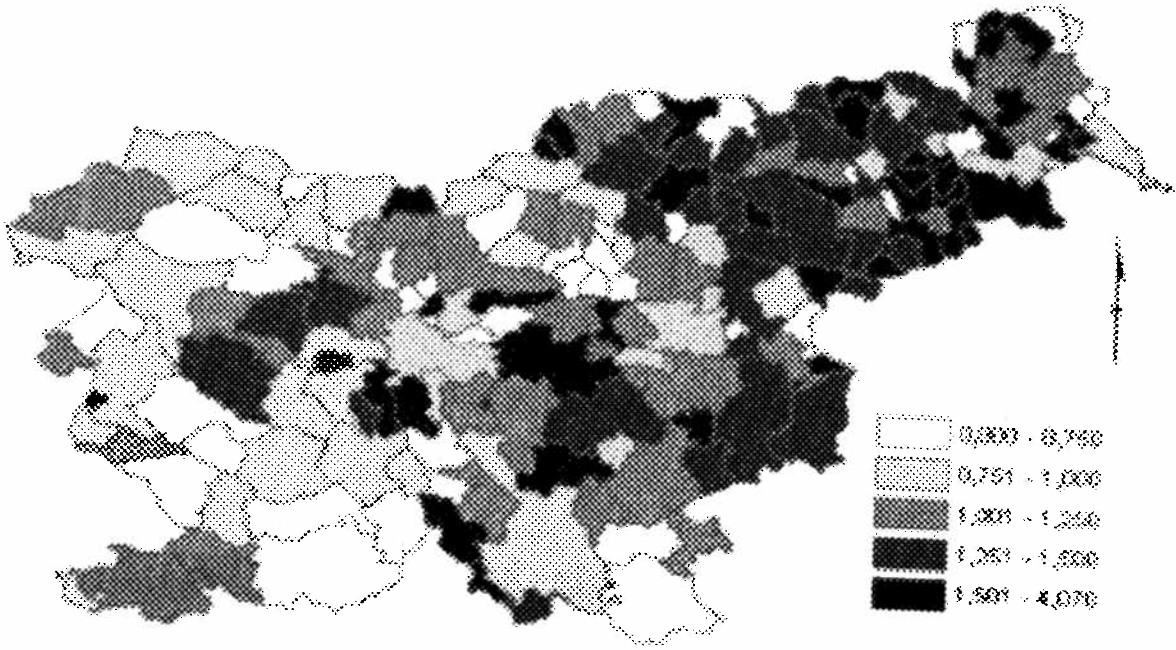
Problem: - Vesna Zadnik, an epidemiologist in Slovenia, is interested in the association of stomach cancer and socioeconomic status in Slovenia (1995-2001)

Dataset: - 192 municipalities partition Slovenia
- For each municipality i, $i=1, \ldots, 192$, we have:
- O_i = Observed # persons with stomach cancer
- E_i = Expected #, indirect standardization
- SEc_i = centered socioeconomic status (SES) score, treated as a continuous measure.

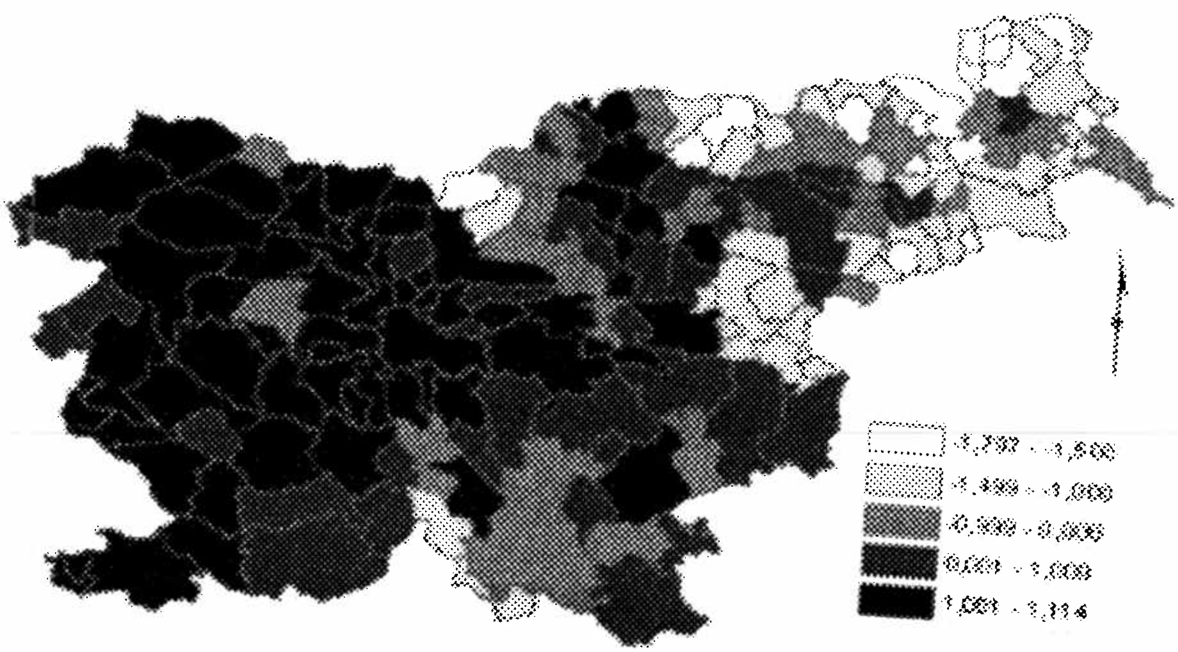
Statistical Question: Is O_i/E_i associated with SEc_i?

Here's a picture of the data:

O_i / E_i



SE_i



VAM T2 add 2 8/11/08

A non-spatial model

- The non-spatial Poisson regression model is

$$O_i \sim \text{Poisson}\left(E_i e^{\alpha + SEc_i\beta}\right),$$

independent across i.

- Under this model, β's posterior median is -0.14 and its 95% interval is (-0.17,-0.10).

- Municipalities with low SE have higher stomach cancer incidence rates.

II 13|7

A spatial model

- A popular Bayesian disease-mapping model is:

$$O_i \sim \text{Poisson}\left(E_i e^{S_i + SEc_i\beta}\right)$$

- S captures spatial clustering; $S_i|S_{(i)}, \tau_S \sim N(\bar{S}_i, \tau_s m_i)$, where $\bar{S}_i$ is the mean of municipality i's m_i neighbors.

- This implies the improper CAR prior

$$p(S|\tau_s) \propto \tau_s^{-(n-G)/2} \exp\left(-\frac{\tau_s}{2} S' QS\right)$$

where G is the number of "islands", $Q_{ii} = m_i$, and $Q_{ij} = -1$ if $i \sim j$ and 0 otherwise.

II B/8

Results

	DIC	p_D	β's median	β's 95% interval
Non-spatial model	1153	2	-0.14	(-0.17, -0.10)
Spatial model	1082	62	-0.02	(-0.10, 0.06)

Adding the spatial random effects causes:

- $\beta's$ variance to increase,
- its center to shift toward zero,
- and its 95% interval to cover zero!

II B/9

Example #2 of "adding RE makes FE go away": Tooth crowns in children

Study:
- Badly-decayed primary teeth ("baby teeth") are often capped with a crown
- Do crown types differ in their failure behavior?
- Question: compare types I, III, IV by time to failure

Dataset:
- 202 children from local pediatric dental practices
- Each child has 1, 2, 3, or 4 crowns in the dataset.
- A given child's crowns are all the same type
- We have covariates (e.g., child's age) that are not relevant for the present purpose.

- Analysis was Cox regression with a random effect (done in R)

- We considered analyses both <u>with</u> and <u>without</u> a random effect for child.

- Parameterization: Type I is the reference
  Types III and IV had indicators.

- Here are the results: SD of child RE ≈ 1.2 $\quad e^{4.8} = 122$

<u>No covariates besides crown type:</u>

		Est	SE	P
No RE	III	0.47	0.20	0.016
	IV	0.13	0.14	0.34
Child RE	III	0.21	0.41	0.61
	IV	0.15	0.26	0.56

!

<u>All covariates included</u>

		Est	SE	P
No RE	III	0.53	0.21	0.01
	IV	0.16	0.15	0.28
Child RE	III	0.23	0.43	0.60
	IV	0.17	0.28	0.55

!

II B / 39   3/9/08

I sent an e-mail to Brian Reich describing this situation – adding an unbalanced one-way random effect makes a FE mostly go away. He replied:

> Boy, that's weird. There's probably a whole new story to tell here. After a quick look at the variance inflation factor in the Slovenia paper, it looks like the only remnants of the model in this equation are the within-kid sample size. Is there any relationship between the number of observations per kid and treatment group?

And the data showed.......

That he's right.

Here are the counts of children with 1, 2, 3, or 4 crowns, for each of the crown types:

# of crowns in data set

Crown type	1	2	3	4
I	6	30	13	33
III	15	13	2	14
IV	1	9	16	50

- A child's number of crowns in the dataset is a surrogate for .... what?
- Whatever it is, it's associated with failure of crowns!

A spatial model

- A popular Bayesian disease-mapping model is:

$$O_i \sim \text{Poisson}\left(E_i e^{S_i + SEc_i\beta}\right)$$

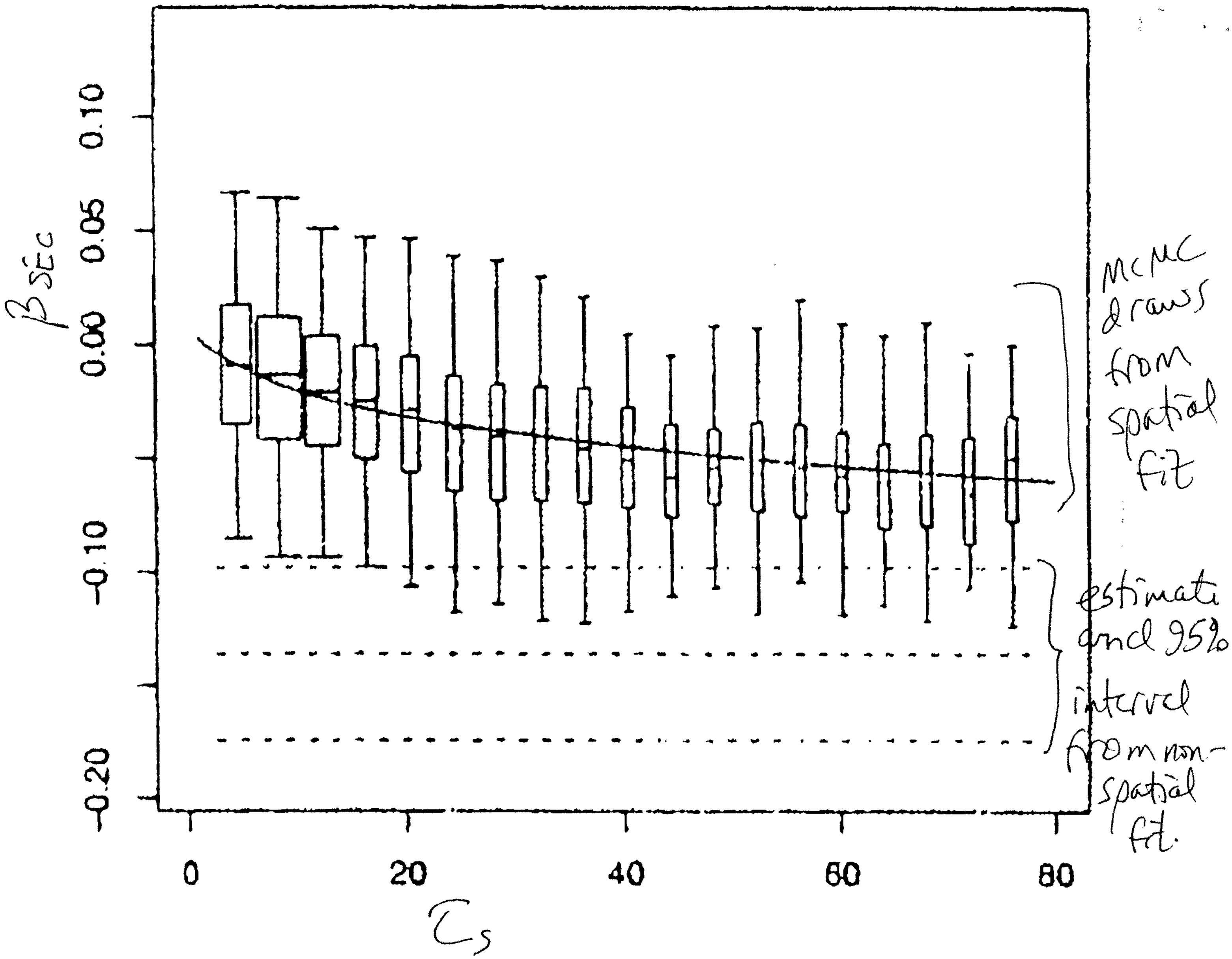
- S captures spatial clustering; $S_i|S_{(i)}, \tau_S \sim N(\bar{S}_i, \tau_s m_i)$, where $\bar{S}_i$ is the mean of municipality i's m_i neighbors.

- This implies the improper CAR prior

$$p(S|\tau_s) \propto \tau_s^{-(n-G)/2} \exp\left(-\frac{\tau_s}{2} S' Q S\right)$$

where G is the number of "islands", $Q_{ii} = m_i$, and $Q_{ij} = -1$ if $i \sim j$ and 0 otherwise.

6/8



A hint of what's going on here:

- Take the MCMC draws from the spatial fit
- Divide them into groups according to draws of τ_s, the <u>precision</u> controlling the CAR smoothing.
- When precision τ_s is small – little smoothness in the CAR – β for SEc is close to zero
- When τ_s is large, β is much farther from zero.

II B/11 3/3/08

Let's consider a simpler model: Normal errors

Assume we're using this model:

$$\underset{N\times 1}{y} = \underset{N\times p}{X}\,\underset{p\times 1}{\beta} + \underset{N\times 1}{S} + \underset{N\times 1}{\varepsilon} \quad \text{iid } N(0, \tfrac{1}{\tau_e})$$

Slovenia: $\log(O_i/E_i) = X\beta + S + H$

Let's make this a bit more explicit:

$$y = X\beta + I_N S + \varepsilon \quad \text{iid } N(0, \tfrac{1}{\tau_e})$$

X: SEc; I_N: implicit regressors; S: $CAR(\tau_s, Q)$

Obvious concern: a regression with N observations and N+p predictors

Usual (brainwashed) thinking: Because we're smoothing/shrinking S, this works (posterior for S is proper)

Still, the potential for collinearity problems is obvious

UNCLEAR: Effect of the constraint $S \sim CAR(\tau_s, Q)$

To clarify the collinearity problems, lets change to a better set of implicit regressors for the CAR effect S

Before: $y = X\beta + I_N S + \varepsilon$ $S \sim CAR(\tau_s, Q)$

After: $y = X\beta + Zb + \varepsilon$ b is Normal with mean 0, precision $\tau_s D$

Where: • $Q = ZDZ'$ is the spectral decomposition
- Z is orthogonal, so $ZZ' = Z'Z = I_N$
- D is diagonal, $d_1 \geq \cdots \geq d_{N-G} > 0$, $d_{N-G+1} = \cdots = d_N = 0$

• $b = Z'S$

Why does this help? because the constraint on b, imposed by the CAR, is simpler:

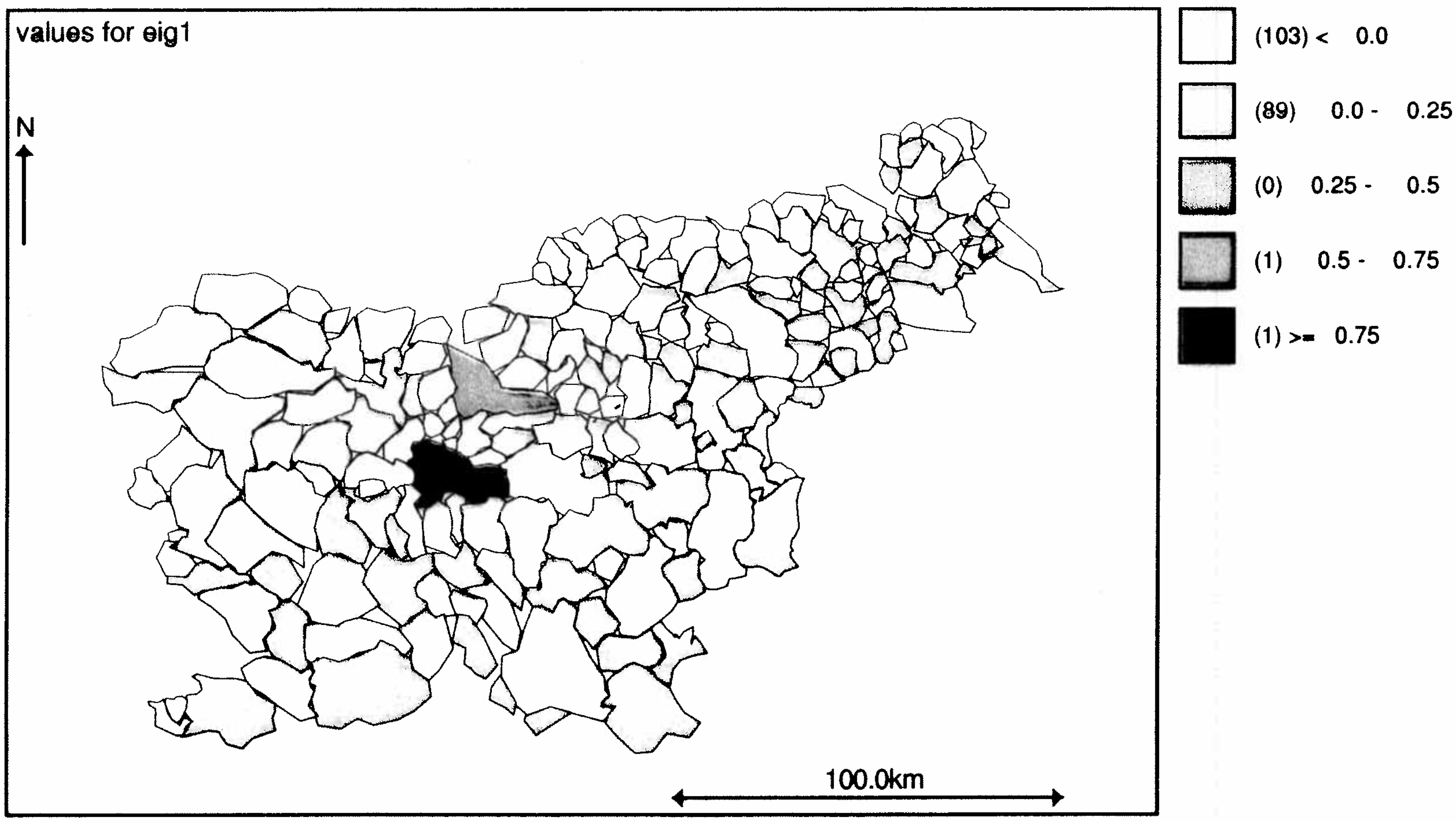
precision $(b) = \tau_s D$, so b's elements are smoothed independently of each other.
- b_1 is smoothed most (d_1 is largest)
- b_{N-G} is smoothed least (d_{N-G} is smallest > 0)
- $b_{N-G+1}, \ldots, b_N$ aren't smoothed at all.

AND the columns of Z are interpretable

VAM J3add2 8/18/08 (replaces II B/13)

Interpreting Q's eigenvectors: largest d_i

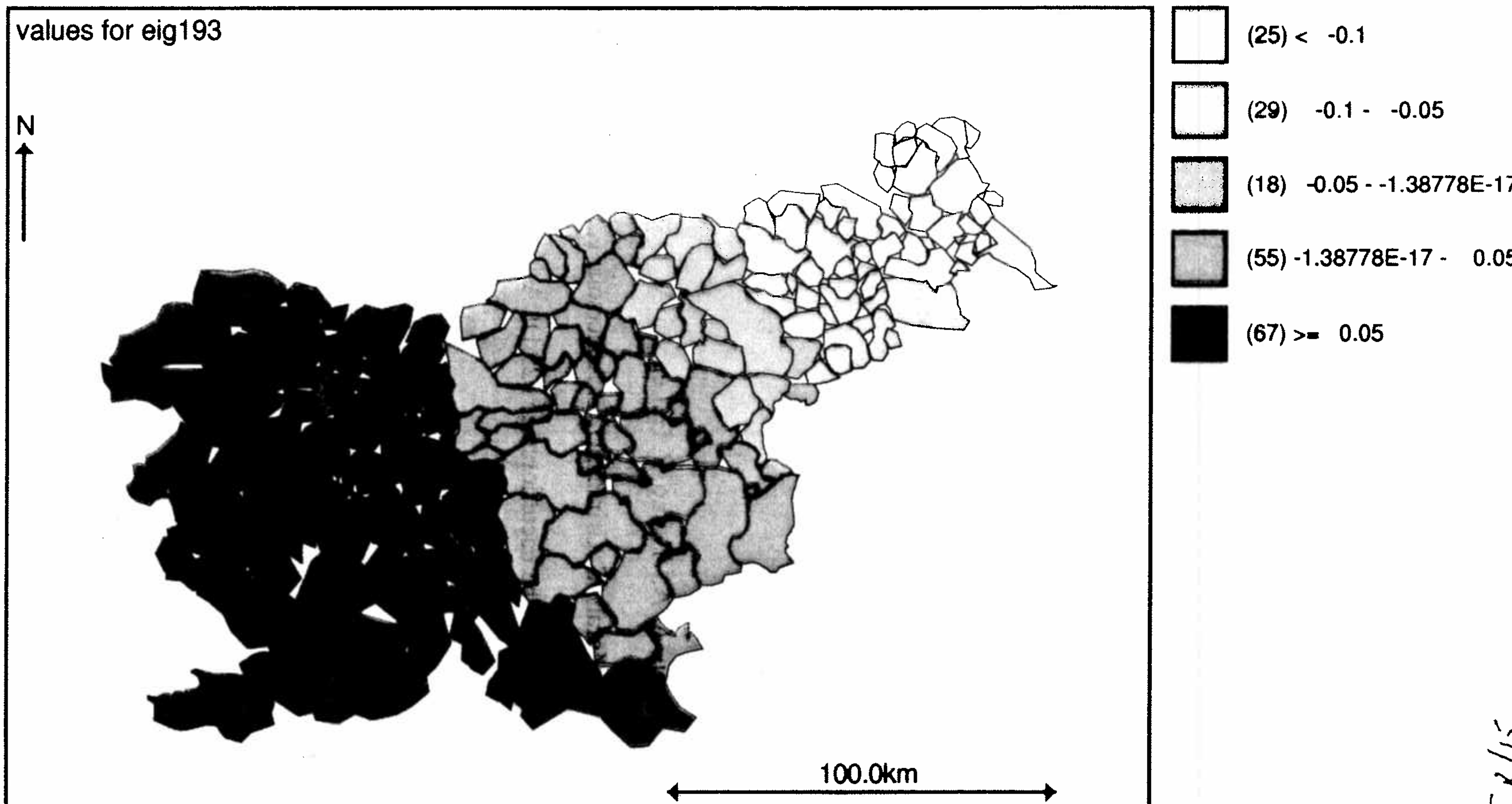
The eigenvector corresponding to $d_1 = 14.46$ compares the two regions with the most neighbors and their neighbors' average.



II 13/14

Interpreting Q's eigenvectors: smallest d_i

The eigenvector corresponding to $d_{N-G} = 0.03$ measures the east/west gradient



II B/15

What does this mean in the Slovenia problem?

- $G = 1$ island in the spatial map (i.e, it's a connected map).
  Thus, b_N corresponds to an intercept, which is not shrunk toward zero

- b_{N-G}, which is constrained <u>least</u> by the CAR model (has the smallest positive d_i) will create the biggest collinearity problem.

<u>BUT</u> • b_{N-G} is the coefficient of Z_{N-G}, which is the NE to SW gradient in Slovenia

<u>AND</u> • SE_c has a strong NE to SW gradient (as does stomach cancer incidence)

Thus, we have a big potential here for a collinearity problem, particularly if τ_s (precision of the CAR) is small, i.e. we're not smoothing much (and it is)

Marginal posterior of β (FE):

Without the spatial (CAR) effect: $y \sim N(X\beta,\ \tau_e I_N)$

$$E(\beta \mid \tau_e, y) = \hat{\beta}_{OLM} = (X'X)^{-1}X'y \quad \text{(flat prior on } \beta\text{)}$$

$$Var(\beta \mid \tau_e, y)^{-1} = \tau_e (X'X)$$

With the canonical CAR added

$$y \sim N(X\beta + Zb,\ \tau_e I_N) \quad b \sim N(0, \tau_s D)$$

$$\begin{aligned}
E(\beta \mid \tau_e, \tau_s, y) &= E\left(E\{\beta \mid b, \tau_e, \tau_s, y\}\right) \\
&= E\left((X'X)^{-1}X'(y - Zb) \mid \tau_s, \tau_e, y\right) \\
&= (X'X)^{-1}X'(y - Z\hat{b}) \qquad \boxed{\hat{b} = E(b \mid \tau_e, \tau_s, y)} \\
&= \hat{\beta}_{OLM} - (X'X)^{-1}X'Z\,E(b \mid \tau_e, \tau_s, y)
\end{aligned}$$

$\boxed{\text{NOT on } \beta!}$

$$Var(\beta \mid \tau_e, \tau_s, y)^{-1} = \tau_e X'X - \tau_e X'Z\left(I_N + \frac{\tau_s}{\tau_e} D\right)^{-1} Z'X$$

(To prove: integrate b out of $(\beta, b) \mid \tau_e, \tau_s, y$)

The variance inflation factor

The increase in β's variance due to adding the CAR effects is

$$\begin{aligned} VIF &= \frac{\beta\text{'s variance under the CAR model}}{\beta\text{'s variance under the OLM}} \\ &= \left(1 - \sum_{j=1}^{n} \frac{\rho_j^2}{1 + rd_j}\right)^{-1} \end{aligned}$$

- $r = \tau_s/\tau_e$
- ρ_j is the correlation between X and Z's i^{th} column
- $\sum_{j=1}^{N} \rho_j^2 = \text{Var}(X) = 1$

II B/17

The variance inflation factor (cont.)

$$
\begin{aligned}
VIF &= \frac{\beta\text{'s variance under the CAR model}}{\beta\text{'s variance under the OLM}} \\
&= \left(1 - \sum_{j=1}^{n} \frac{\rho_j^2}{1 + rd_j}\right)^{-1}
\end{aligned}
$$

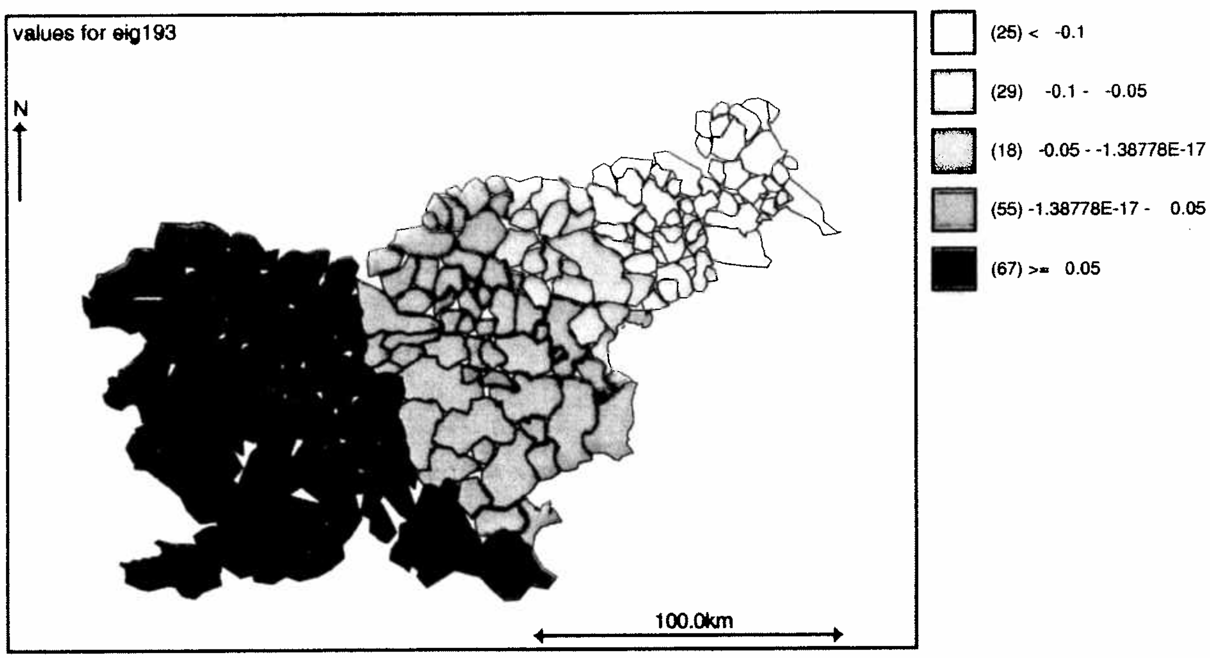
- $1 < VIF < \infty$ for all $r = \tau_s / \tau_e > 0$
- $VIF \to \infty$ as $r \to 0^+$; $VIF \to 1$ as $r \to \infty$
- VIF is large if there is little smoothing and X is highly correlated with eigenvectors associated with small d_i.

II B|18

VIF for the Slovenia example

SEc satisfies both of these criteria:

- there is little smoothing; $p_D = 62$.
- the correlation between *SEc* and the eigenvector with smallest d_i (below) is 0.72.



IIb/19

- Reich, Hodges, & Zadnik (2006) has lots of other stuff
- Like other collinearity effects, adding the CAR can make your fixed effect (β) get larger or smaller in magnitude, and can make it change sign

---

<u>A rationalization I've heard:</u>

- The CAR is representing a missing spatial effect, and this collinearity effect is thus appropriate.

<u>Response:</u>

- This collinearity effect happens even if there <u>is</u> no missing spatial effect
- The implicit regressors, implied by the CAR model/prior, are determined <u>solely</u> by the spatial map and will coincide with genuine missing predictors only by accident.

ANOTHER Approach to the result in the Slovenia problem (i.e., adding a RE "zaps" a FE)

- The approach we've taken so far will seem foreign to many spatial specialists, who tend to think of CAR models and other models as describing covariance of errors, not a linear regression with constrained coefficient estimates.
- I'll now show how you can start with this spatial-analysis perspective and get the same explanation of
  - change in $\hat{\beta}$
  - Variance inflation of $\hat{\beta}$

(This can be done in at least two different ways.)

II B/31 3/9/08

Again, let's consider the simpler case of Normal errors

start again with

$$\overset{N\times 1}{y} = X\beta + S + \varepsilon \qquad S \sim CAR(\tau_s, Q)$$
$$\varepsilon \text{ iid } N(0, \tfrac{1}{\tau_e})$$

To think of this in terms of spatial correlation (instead of collinearity), we must re-write it as

$$y = X\beta + \xi$$ where the error ξ has covariance Σ capturing the spatial correlation in S

- This involves a little cleverness because the precision matrix of S can't be inverted to give its covariance matrix and simply added to $\frac{1}{\tau_e} I_N$
- Omitting the details, it's not hard to show that

$$\Sigma = \frac{1}{\tau_e}\left[I_N - (I_N + rQ)^{-1}\right]^{-1} \qquad r = \frac{\tau_s}{\tau_e} = \frac{\sigma_e^2}{\sigma_s^2}$$

Unfortunately, this is obscure because Q is not simple.

VAM T3 add 5 8/18/08 (replaces IIB/32)

Recap: $y = X\beta + \xi$, $\text{cov}(\xi) = \Sigma$ (*)

$$\Sigma = \frac{1}{\tau_e}\left[I_N - (I_N + rQ)^{-1}\right]^{-1} \qquad r = \frac{\tau_s}{\tau_e} = \frac{\sigma_e^2}{\sigma_s^2}$$

To simplify this, consider again

$Q = ZDZ'$, the spectral decomposition,

where: Z is orthogonal

D is diagonal

$$\Sigma = \frac{1}{\tau_e} Z \underbrace{\left[I_N - (I_N + rD)^{-1}\right]^{-1}}_{\text{diagonal: } \frac{1+rd_i}{rd_i} \text{ are the diagonal elements.}} Z'$$

<u>Last step</u>: pre-multiply (*) above by Z':

$$Z'y = Z'X\beta + \xi^*, \quad \text{cov}(\xi^*) = \frac{1}{\tau_e}\,\text{diag}\left(\frac{1+rd_i}{rd_i}\right)$$

our data & model are now:

$$(z_1'y) = (z_1'X)\beta + \xi_1^* \quad \text{var}(\xi_1^*) = \frac{1}{\tau_e}\left(\frac{1+rd_1}{rd_1}\right)$$

$$\vdots$$

$$(z_{N-G}'y) = (z_{N-G}'X)\beta + \xi_{N-G}^* \quad \text{var}(\xi_{N-G}^*) = \frac{1}{\tau_e}\left(\frac{1+rd_{N-G}}{rd_{N-G}}\right)$$

If <u>r is small</u> (little smoothing overall) and <u>d_i is small</u>

$z_i'y$ has little influence on $\hat{\beta}$

What does this mean for Slovenia?

- Z_{N-G} describes the NE to SW gradient in Slovenia
  - $Z'_{N-G}y$ is the info in stomach cancer incidence about the NE to SW gradient
  - $Z'_{N-G}X$ is the info in SES about the NE to SW gradient
- r is small (little smoothing)
- d_i is small

Thus, the CAR "error covariance" throws away the info in the data about the association between stomach cancer and SES

Implication • This effect can happen for <u>any</u> spatial covariance embodying the idea that "near regions are more similar than far regions".

We can now consider two interpretations of this "Slovenia effect"

(1) In the model $y = X\beta + \xi$, $\xi = S + \varepsilon$,

the error ξ is correlated with the regressor X.

- Well Known (to econometricians): If X and ξ are correlated, the GLS estimate $\hat{\beta}_{GLS}$ is biased – Thus, this is a bias problem (follows from the alternative interpretation)

(2) It doesn't make much sense here to conceive of S in $y = X\beta + S + \varepsilon$ as a draw from a probability distribution, if that implies that you could make another draw and get a different S.

- Instead, it does less violence to the subject matter to think of S like we (I?) think of penalized splines: as a fixed but unknown thing that we want to estimate smoothly, for which we adopt a random-effects form for convenience.

- Then this is a collinearity problem.

So what do we do about this?

One view (not mine)

- The analysis without the random effect is wrong.
  - Don't even consider it
  - The change from adding the RE to the analysis is not an interesting quantity.

Well, but...

- It is not always so clear that an analysis including the RE is necessary
  - Example: Spoont pilot data (PTSD treatment adherence in VA)

- If the object in adding a RE is to "get the standard errors right", it is possible to do so *and* avoid the collinearity effect

Using this technical trick...
(which, I admit, makes me really nervous)

A trick to avoid the collinearity effect (Reich et al 2006 . sections 3, sec 4 to pp. 1120)

Normal case: Instead of fitting

$$y = X\beta + S + \varepsilon \qquad \begin{array}{l} S \sim CAR(\tau_s, Q) \\ \varepsilon \sim \text{iid } N(0, \frac{1}{\tau_e}) \end{array}$$

Fit: $y = X\beta + P^c S + \varepsilon \qquad P^c = I - X(X'X)^{-1}X'$

$S \sim CAR(\tau_s, Q)$

$\varepsilon \sim \text{iid } N(0, \frac{1}{\tau_e})$

- In other words, restrict the spatial smoothing to the part of the fitted-value space orthogonal to the fixed effects X

- For the normal case, the conditional posterior of β given (τ_e, τ_s) is the same as for the model without S
  - BUT adding $P^c S$ makes τ_e smaller, a posteriori, producing the desired effect.

VAM T3 add 10 8/18/08 replaces II B/29

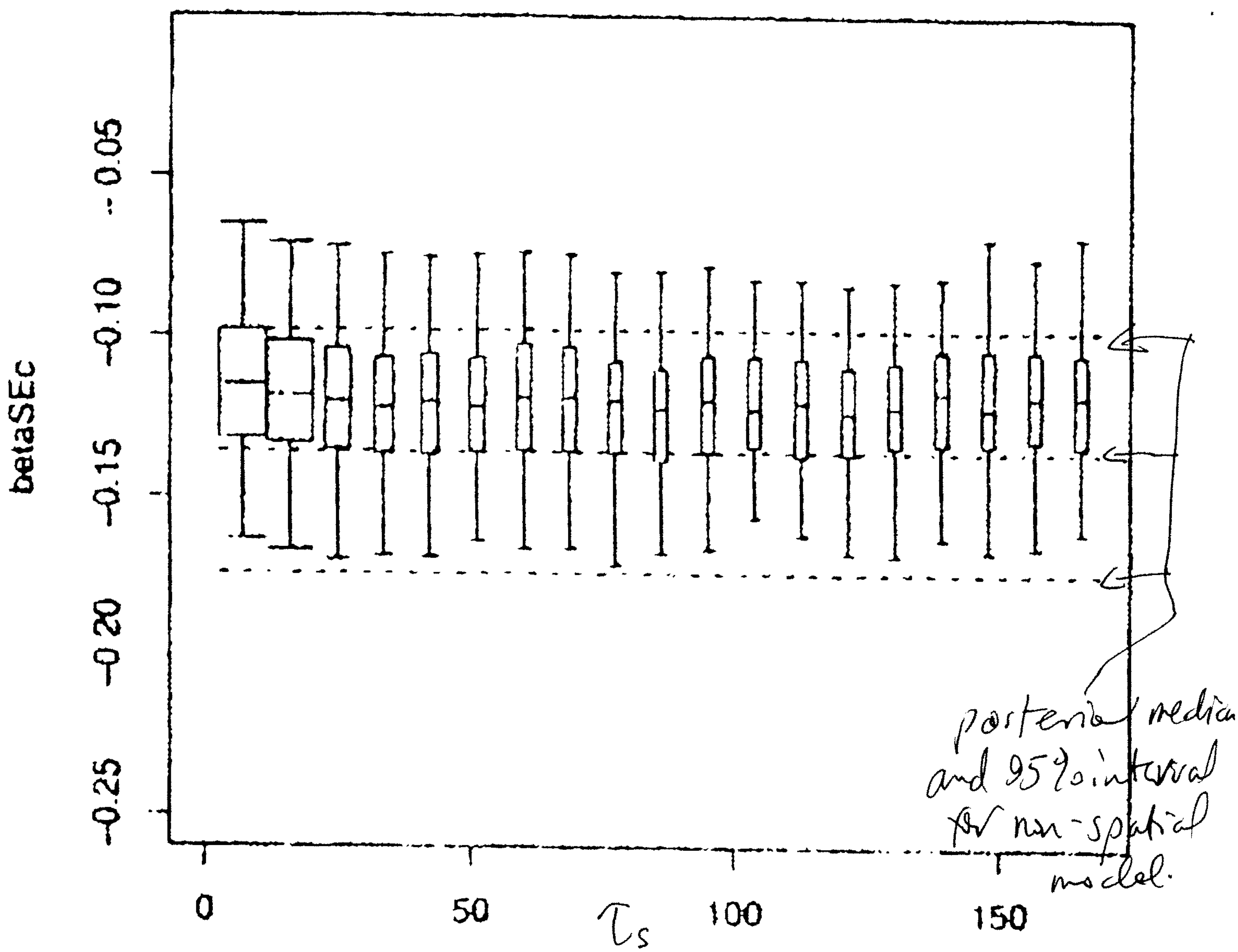
This idea is generally applicable, and thus attractive and <u>dangerous</u>

RHZ (2006), p.1201: "The primary objective in a spatial regression may be to estimate the fixed effects; in such a situation, the CAR random effects are added merely to account for spatial correlation in the residuals when computing the posterior variance of the fixed effects, or to improve predictions. With these objectives in mind, it may be reasonable to proceed by removing the combinations of CAR random effects that are collinear with the fixed effects"

~~~~ = weasel words!

- If the RE does represent a true (implicit) missing effect, e.g. if cluster size really *is* informative, this idea arbitrarily credits all the disputed variation in y to the fixed effect X.

VAMT3 add 11, comes after IIB/30
~~~~



- Doing this for the Slovenia data produces MCMC draws for β and τ_s as above
    - This is for the Poisson model discussed in Reich et al (2006), so the previous slide's theory is only approximate
    - Still, the posterior for β is now essentially independent of the smoothing.

- How about the "crowns in children" xample?
- We know hardly anything about this; I'll show some first-cut thoughts

(after reminding you what the crowns-in-kids phenomenon was)

- Analysis was Cox regression with a random effect (done in R)

- We considered analyses both <u>with</u> and <u>without</u> a random effect for child.

- Parameterization: Type I is the reference
  Types III and IV had indicators.

- Here are the results: SD of child RE ≈ 1.2

$e^{4.8} = 122$

No covariates besides crown type:

		Est	SE	P
No RE	III	0.47	0.20	0.016
	IV	0.13	0.14	0.34
Child RE	III	0.21	0.41	0.61
	IV	0.15	0.26	0.56

!

All covariates included

		Est	SE	P
No RE	III	0.53	0.21	0.01
	IV	0.16	0.15	0.28
Child RE	III	0.23	0.43	0.60
	IV	0.17	0.28	0.55

!

II B/39 3/9/08

I sent an e-mail to Brian Reich describing this situation – adding an unbalanced one-way random effect makes a FE mostly go away. He replied:

> [illegible], that's weird. There's probably a whole new story to tell here. After a quick look at the variance inflation factor in the Slovenia paper, it looks like the only elements of the model in this equation are the within-kid sample sizes. Is there any relationship between the number of observations per kid and treatment group?

And the data showed.......

II B / 40 3/9/08

That he's right.

Here are the counts of children with 1, 2, 3, or 4 crowns, for each of the crown types:

Crown type	# of crowns in data set: 1	2	3	4
I	6	30	13	33
III	15	13	2	14
IV	1	9	16	50

- A child's number of crowns in the dataset is a surrogate for .... what?
- Whatever it is, it's associated with failure of crowns!

Re: "crowns in kids" example: For the case of clustered data, very little appears to be known

- Based on some preliminary forays, I have a rough understanding of this and some (vaguely specified) conjectures
- These are based on comparing 2 groups, using these models:

Dumb: $y_{ik} = \mu_i + \varepsilon_{ik}$ $\quad i = 1, 2$ $\quad k = 1, \cdots, n_i$ $\quad \varepsilon_{ik} \sim \text{iid } N(0, \sigma_e^2)$ (t-test)

R.E.: $y_{ijk} = \mu_i + c_{ij} + \varepsilon_{ijk}$ $\quad i = 1, 2$ groups
$\quad j = 1, \cdots, m_i$ clusters in grp i
$\quad k = 1, \cdots, n_{ij}$ $\quad \sum_j n_{ij} = n_i$

$c_{ij} \sim N(0, \sigma_s^2)$
$\varepsilon_{ijk} \sim N(0, \sigma_e^2)$

Big idea here:

Dumb analysis compares n_1 individuals in Group 1 vs. n_2 individuals in Group 2

RE analysis compares m_1 clusters in Group 1 vs. m_2 clusters in Group 2

NOTE: - this is a very simple case!

Consider this situation:

	Group 1	Group 2
One Big cluster ($n=50$)	$n_{1,1}=50$	$n_{2,1}=50$
25 little clusters ($n=2$ each)	$n_{1,2}=2$ … $n_{1,26}=2$	$n_{2,2}=2$ … $n_{2,26}=2$

Example: Dumb analysis no effect, RE analysis big effect

True mean $\mu_i + c_{ij}$

	cluster BIG	cluster SMALL
Group 1	0	10
Group 2	10	0

$\sigma_e^2 = 25$

Fake data:

	Group estimates 1	Group estimates 2	SE diff'c	P difference
Dumb	4.6	4.6	0.69	0.93
RE	8.6	0.05	0.89	<0.0001

Example: Adding RE makes effect go away

True mean $\mu_i + c_{ij}$

	cluster BIG	cluster SMALL
Group 1	0	0
Group 2	10	0

$\sigma_e^2 = 25$

Fake data:

	Group estimates 1	Group estimates 2	SE diff'c	P difference
Dumb	−0.8	3.9	0.62	<0.0001
RE	−0.7	−0.1	0.87	0.63

UAM B add 13 8/2/08

What's going on here?

- In each group, the big cluster's n is the same size as all of the little clusters combined.
- In the Dumb test, the big cluster's weight is the same as all the little clusters combined.
- In the RE test, the big cluster is just 1 compared to the 25 little clusters, so the big cluster gets little weight
  - True, the big cluster's RE is smoothed less toward the group's overall estimate
  - But it counts for less

Example: Adding RE makes effect appear (same as on previous page)

True mean μ_{itcij}

		cluster BIG	SMALL
Group	1	0	10
	2	10	0

Group estimates

Fake data		1	2	
	Dumb	4.6	4.6	← BIG & SMALL clusters have $n=50$ each
	RE	8.6	0.05	← 1 big cluster, 25 little clusters; little clusters predominate

Tentative generalization:

- If the two groups have the same distribution of clustersites
- The Dumb and RE effects differ if and only if y_{ijk} shows an interaction of group and cluster site

Now consider this situation:

	Group 1	Group 2
Big clusters ($n=50$ each)	$n_{1,1}=50$, $n_{1,2}=50$, $n_{1,3}=50$	$n_{2,1}=50$
Little clusters ($n=2$ each)	25 clusters: $n_{1,4}=2$ … $n_{1,28}=2$	75 clusters: $n_{2,2}=2$ … $n_{2,76}=2$

Now you don't need y_{ijk} to show a group-by-cluster site interaction

True mean $\mu_i + c_{ij}$

Group	cluster BIG	cluster SMALL	True avg (by unit)
1	10	0	7.5
2	10	0	2.5

Fake data with $\sigma_e^2 = 25$

	Group estimates 1	Group estimates 2	SE difference	P difference
Dumb	7.1	2.5	0.7	<0.0001
RE	0.9	0.3	1.0	0.54

The (many) small clusters swamp the (few) large clusters, slightly less so in Group 1, which has more big clusters, fewer small ones.

VAM TJ add 15 8/21/08

What's happening in the "Kids 'n' crowns data"?

n kids:

Kid has this many crowns

Crown type:	1	2	3	4
I	6	30	13	33
III	15	17	5	21
IV	2	9	16	49

Note: Yes, this differs a bit from the earlier table.

So cluster size is associated with treatment group: Crown types I and IV have more big clusters.

How does the outcome relate to cluster size?

fraction of crowns failing

Kid has this many crowns

	1	2	3	4
overall % failing	57%	36%	43%	35%
% failing T I	67	28	21	36
% failing T III	60	38	47	37
% failing T IV	0	61	63	34

only 2 kids, 1 crown each

A crude test shows crown-type by #crowns interaction is significant (FWIW)

VAMT3 add #17 8/21/08