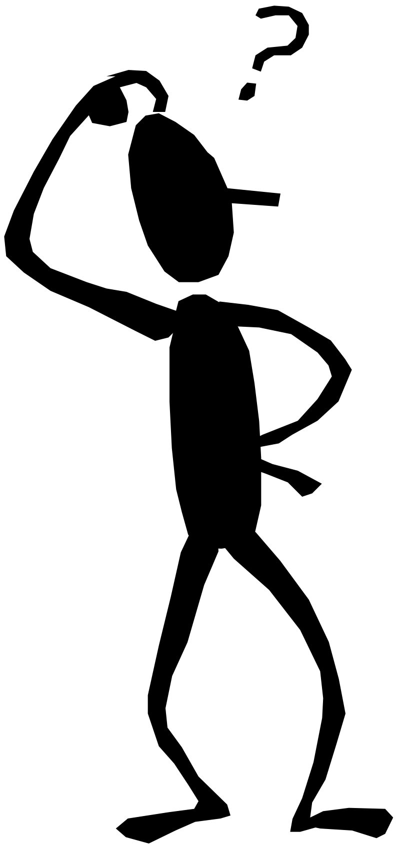
Ask a Hypothetical Question, Get a Valuable Answer?

Joseph Herriges
Catherine Kling
Christopher Azevedo

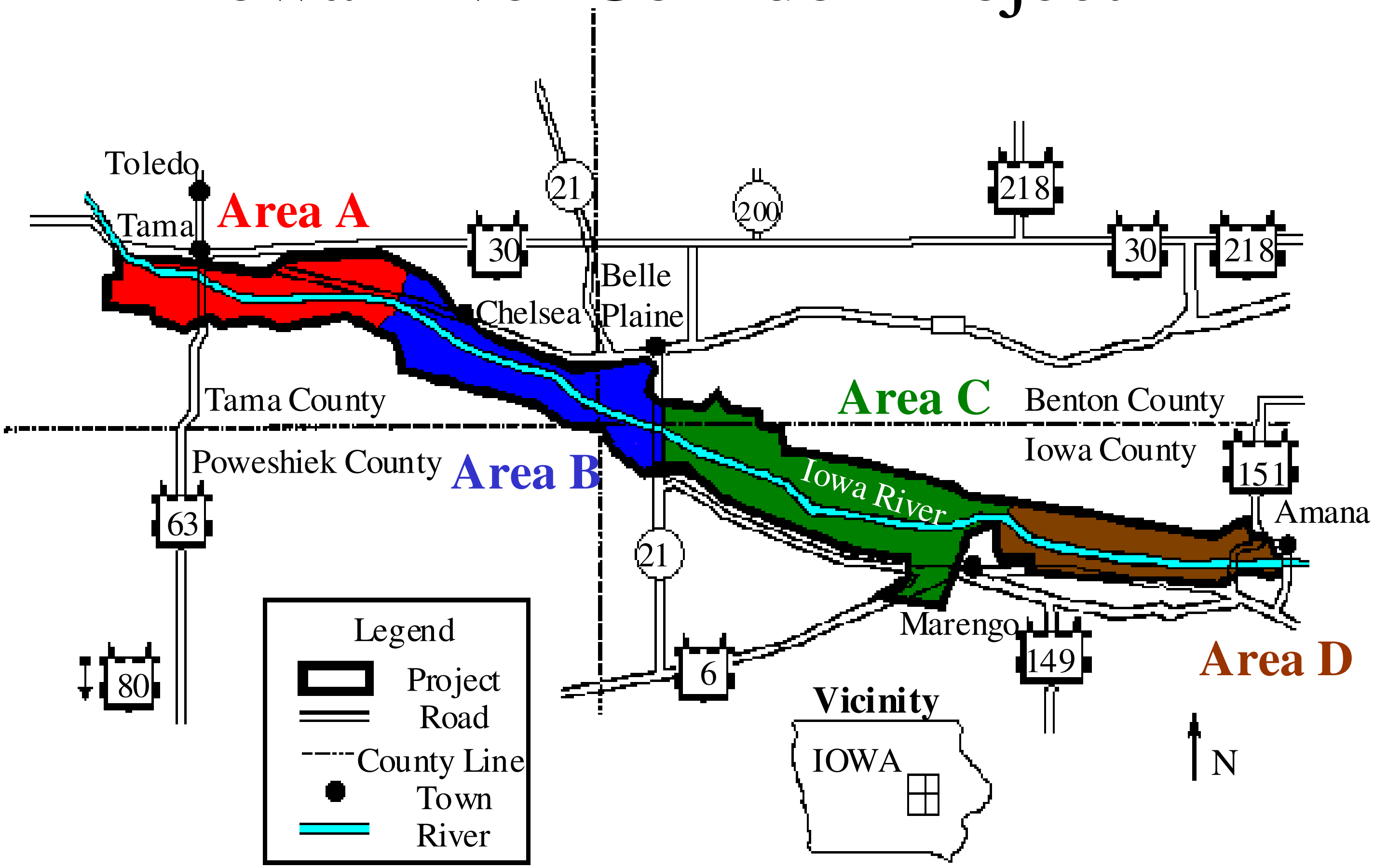


Prairie Pothole Region

Prairie Pothole
Region of North
America

U.S. Fish & Wildlife Service

Iowa River Corridor Project



Sources of Data

1. Revealed Preference (RP): Behavior
   e.g., hedonic price data, travel cost data

2. Stated Preference (SP): Statements
   e.g., contingent valuation, contingent behavior

Previous Work

- Pooling (RP, SP data of same form)
  - Dickie, Fisher, Gerking (1987)
  - Adamowicz, Louviere, and Williams (1994)
  - Layman, Boyce, and Criddle (1996)
  - Englin and Cameron (1996)
- Combining (RP,SP data of different form)
  - Cameron(1992)
  - Larson (1990)
  - Loomis (1997)
  - Huang, Haab, and Whitehead (1997)
  - McConnell, Weninger, and Strand (1999)
  - Cameron, Poe, Ethier, Schulze (1999)

Previous Work (continued)

- Pooling or Combining (2 pieces of SP data)
  - Niklitschek and Leon (1996)
  - Huang, Haab, and Whitehead (1997)

Reasons Cited for Combining or Pooling Data

- Increase Precision of Estimates
- Test Consistency Across RP and SP Data
- Impose discipline of market on SP data, while allowing SP data to “fill in” some information about preferences not captured by RP data (Cameron)

Alternative Interpretations of Consistency Tests (rejection)

1. RP Lovers: View these as validity tests of SP against RP (MWS)
2. SP Lovers: View these as validity tests of RP(?)

   Basis: Randall, mis-measured RP data then biased price coefficient
3. Agnostics: Jointly estimate and constrain parameters to be alike to take advantage of strengths of both

Our Interests

- Value Wetland Use in Iowa using RP and SP Data Jointly
- Test for consistency of RP and SP generally
- Test specific hypotheses concerning sources of bias

Our Interests (more)

- Investigate these issues with two different forms of SP data
  - Dichotomous data: “yes/no” answers
  - Continuous data: “how many?” answers

- Reconsider Interpretation of bias and consistency tests

Model of RP Data

- Standard Demand Function

$$x^{RP} = f(p^{RP}, y^{RP}; \beta^{RP}) + \sigma^{RP} \varepsilon^{RP},$$

- Tobit: Correction for Censoring

$$f(\mathrm{x} \mid \mathrm{x} > 0) = h(\varepsilon) = \frac{(1/\sigma)\phi[(x - f(p, y))/\sigma]}{\Phi[f(p, y)/\sigma]},$$

Model of SP Continuous Data

- Standard Demand Model Again

$$x^{SP} = f(p^{SP}, y^{SP}; \beta^{SP}) + \sigma^{SP} \varepsilon^{SP},$$

- Tobit: Correction for Censoring

- Consumer Surplus at Current Use

$$\hat{c} = g(p, y; \hat{\beta}, \hat{\sigma}).$$

Model of SP Discrete Data

- Only “yes” or “no” response from SP data
- Model Probability of “yes” as:

$$\begin{aligned}\Pr(\text{wtp} = \text{"yes"}) &= \Pr\{f(p^{SP}, y^{SP}; \beta^{SP}) + \sigma^{SP}\varepsilon^{SP} > 0\}\\ &= \Pr\{\varepsilon^{SP} > -f(p^{SP}, y^{SP}; \beta^{SP}) / \sigma^{SP}\}\\ &= \Phi[f(p^{SP}, y^{SP}; \beta^{SP}) / \sigma^{SP}].\end{aligned}$$

- Pr(wtp=“no”) = 1-Pr(yes)

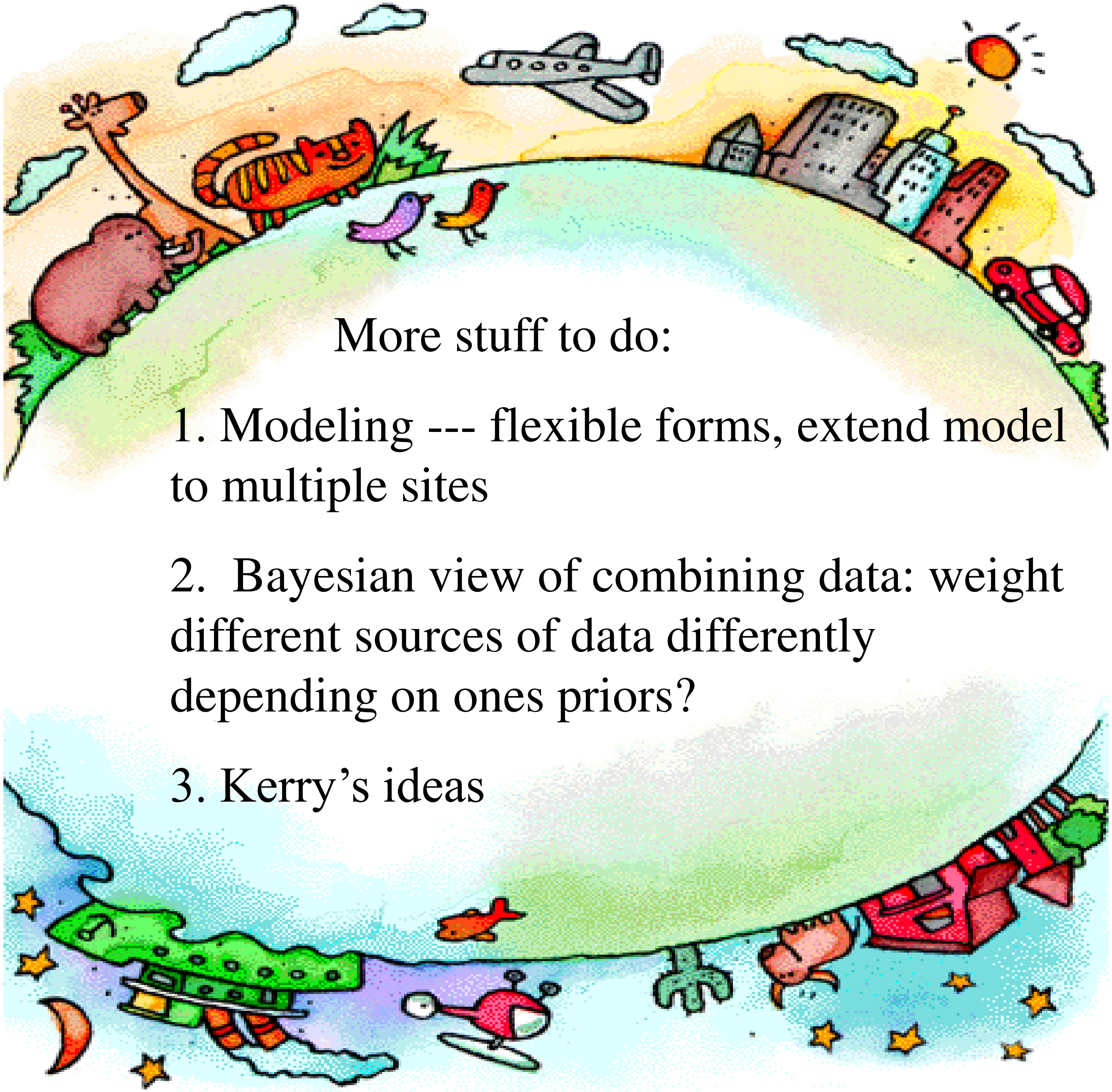
Joint Estimation of SP and RP Data

- Simple sum of RP and SP likelihoods if independent errors, but data from same individuals so correlation likely

- Why not identical errors from same individuals?

  - Sources of RP error
    - Recall concerning # of visits
    - Errors in optimization
    - Random Preferences
    - Omitted variables
  - Sources of SP error (previous three plus)
    - Inaccurate comprehension of survey wtp question
    - Phone vs. mail vs. in person survey
    - Inaccurate comprehension of other survey details (payment vehicle, time table for payment, etc.)

Joint Estimation of SP and RP Data (Correlated)

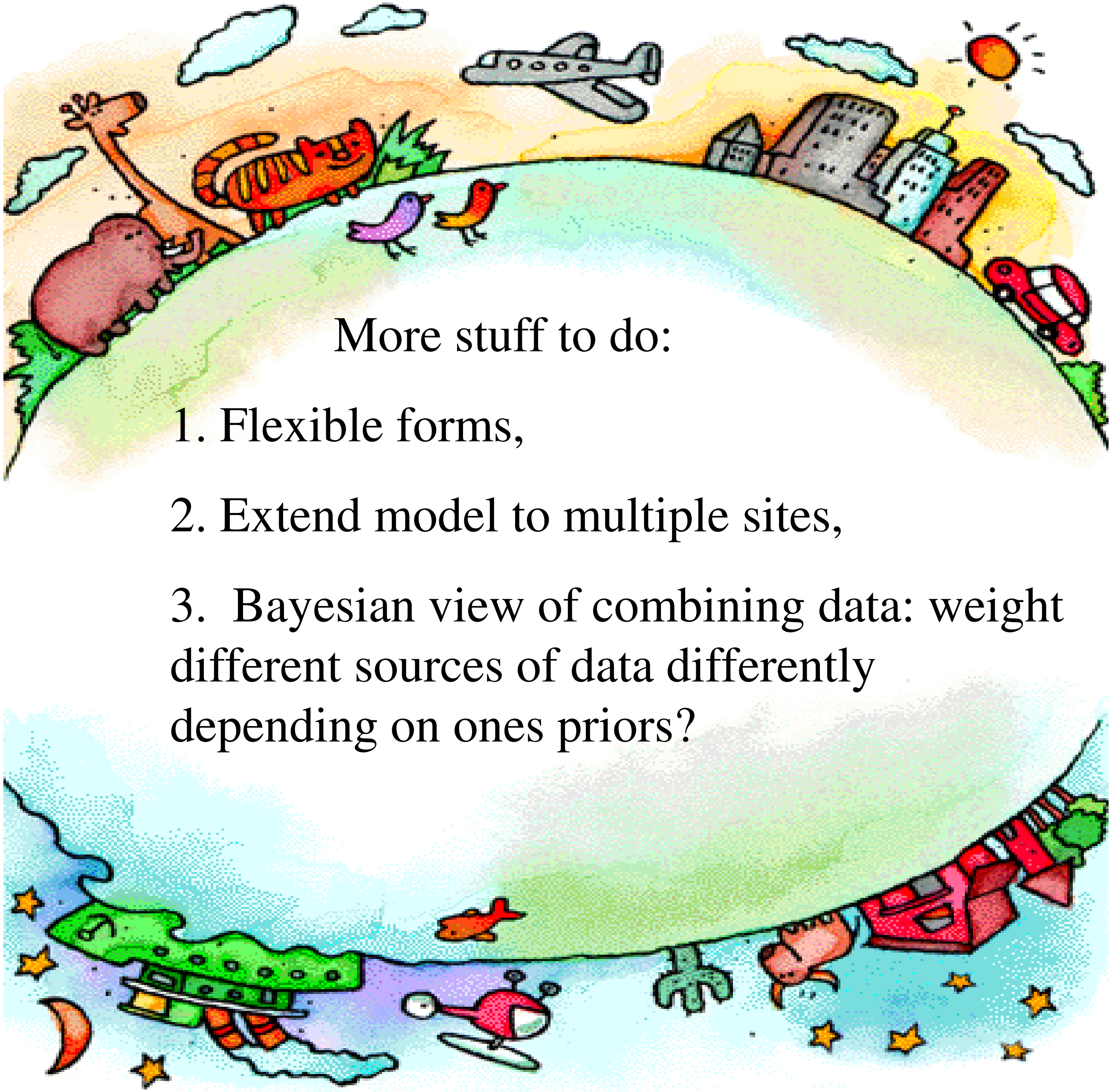


More stuff to do:

1. Modeling --- flexible forms, extend model to multiple sites

2. Bayesian view of combining data: weight different sources of data differently depending on ones priors?

3. Kerry's ideas



More stuff to do:

1. Flexible forms,

2. Extend model to multiple sites,

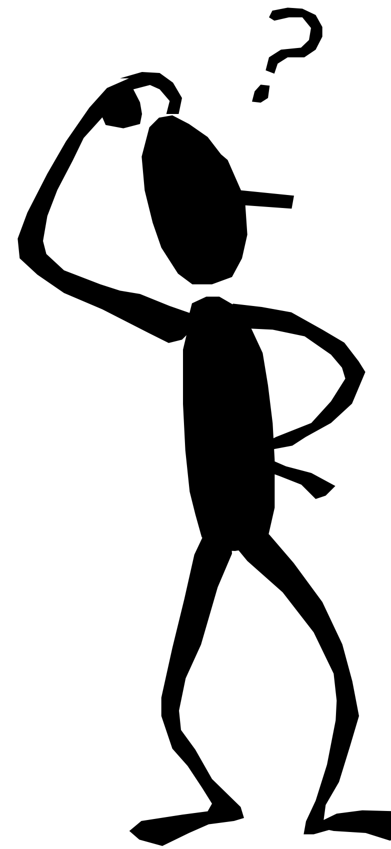
3. Bayesian view of combining data: weight different sources of data differently depending on ones priors?

Consistency Tests

1. Parameter values are identical across RP and SP data: test equality of all coefficients
2. Parameter values are identical but errors have different variances (heteroskedasticity): test equality of all coefficients except variances

RP and SP^C Joint Models

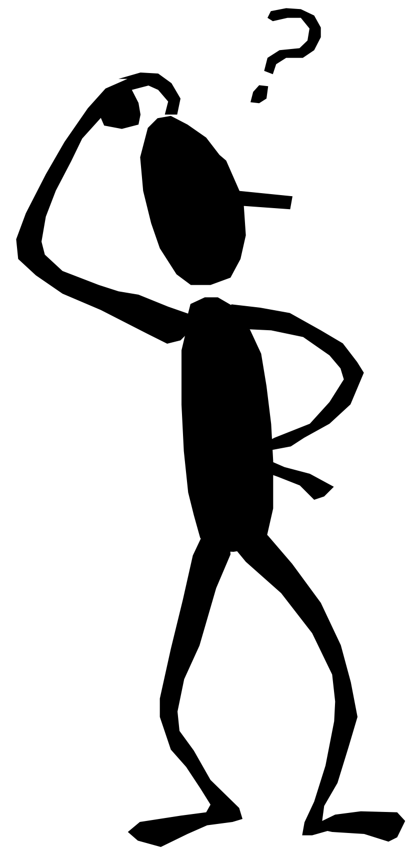
	Independent	Correlated	Fully Consistent	Hetero-skedasticity
Constant^{RP}	20.65 (8.21)	19.12 (7.98)	14.83 (7.92.)	16.11 (7.92)
$k^{SP}_{constant}$	0.50 (-2.54)**	0.61 (-2.45)**	1.00	1.00
Price^{RP}	-0.82 (-8.82)	-0.76 (-9.11)	-0.55 (-14.87)	-0.61 (-11.22)
k^{SP}_{price}	0.58 (-3.54)**	0.60 (-3.97)**	1.00	1.00
Income^{RP}	0.14 (3.84)	0.13 (3.65)	0.11 (3.24)	0.12 (3.41)
k^{SP}_{income}	0.75 (-0.61)**	0.64 (-1.20)**	1.00	1.00
σ^{RP}	13.28 (18.42)	13.43 (18.13)	13.64 (19.77)	13.19 (18.56)
k^{SP}_{σ}	1.12 (1.02)**	1.01 (0.11)**	1.00	1.15 (1.53)**
ρ		0.63 (13.27)	0.64 (13.48)	0.64 (13.41)
-log $\mathscr{L}$	1180.80	1136.92	1142.91	1141.55
CS^{RP}	114.14	122.35	169.45	153.27
CS^{SP}	197.93	203.92		
p-values (Likelihood ratio tests)			0.02 (reject @5%)	0.03 (reject)



The t-statistics are in parentheses below the coefficient estimates.
Double asterisks indicate tests for difference from 1.00.

RP and SP^D Joint Models

	Independent	Correlated	Fully Consistent	Hetero-skedasticity
$Constant^{RP}$	20.65 (8.21)	19.59 (7.92)	15.77 (6.76)	18.80 (7.86)
$k^{SP}_{constant}$	0.77 (-0.97)**	0.80 (-1.01)**	1.00	1.00
$Price^{RP}$	-0.82 (-8.82)	-0.78 (-8.85)	-0.58 (-11.17)	-0.74 (-8.80)
k^{SP}_{price}	0.72 (-2.09)**	0.72 (-2.45) **	1.00	1.00
$Income^{RP}$	0.14 (3.84)	0.14 (3.66)	0.12 (3.49)	0.13 (3.66)
k^{SP}_{income}	0.90 (-0.23)**	0.82 (-0.48)**	1.00	1.00
σ^{RP}	13.28 (18.42)	13.51 (19.79)	13.47 (20.15)	13.38 (20.10)
k^{SP}_{σ}	1.12 not estimated	1.01 not estimated	1.00	1.53 (2.04)**
ρ		0.56 (9.72)	0.57 (9.84)	0.57 (9.87)
$-\log \mathscr{L}$	926.25	900.74	904.34	901.57
CS^{RP}	113.94	118.86	159.56	125.56
CS^{SP}		144.21		
p-values (Likelihood ratio tests)			0.07 (accept)	0.43 (accept)



The t-statistics are in parentheses below the coefficient estimates.
Double asterisks indicate tests for difference from 1.00

Tests of Bias Stories

- Hypothesis: When respondents answer SP questions, they ignore their budget constraint: Test equivalence of all parameters except income coefficient (and variance)

  Premise: RP is accurate, test SP against it (RP Lovers)

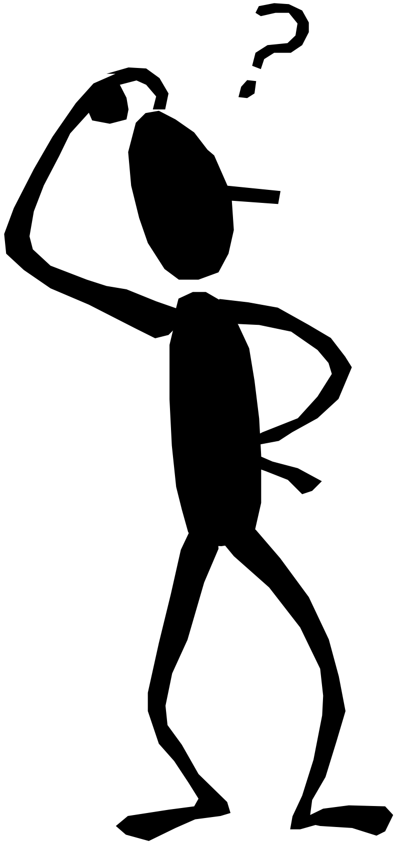
- Hypothesis: Price term in RP data is mis-measured: Test equivalence of all parameters except price coefficient (and variance)

  Premise: SP is accurate, test RP against it (SP Lovers)

- Hypothesis: Both of the above
  - Premise: Both potentially inaccurate (Agnostics)

RP and SP^C Joint Models: Hypothesis Tests

	Correlated	Price Hypothesis	Income Hypothesis	Joint Price and Income
$Constant^{RP}$	19.12 (7.98)	17.24 (7.44)	15.99 (7.44)	17.26 (7.67)
$k^{SP}_{constant}$	0.61 (-2.45)**	1.00	1.00	1.00
$Price^{RP}$	-0.76 (-9.11)	-0.67 (-11.35)	-0.61 (-9.86)	-0.70 (-9.05)
k^{SP}_{price}	0.60 (-3.97)**	0.92 (-1.37)**	1.00	0.80** (-2.48)
$Income^{RP}$	0.13 (3.65)	0.12 (3.53)	0.12 (3.41)	0.14 (3.82)
k^{SP}_{income}	0.64 (-1.20)**	1.00	0.96 (-0.16)**	0.46 (-1.80)**
σ^{RP}	13.43 (18.13)	13.32 (18.22)	13.18 (18.51)	13.38 (18.17)
k^{SP}_{σ}	1.01 (0.11)**	1.10 (1.02)**	1.15 (1.52)**	1.06 (0.61)**
ρ	0.63 (13.27)	0.63 (13.22)	0.64 (13.21)	0.64 (13.38)
$-\log \mathcal{L}$	1136.92	1140.82	1141.54	1139.38
CS^{RP}	122.35	139.02	154.44	132.56
CS^{SP}	203.92	151.42		166.74
p- values for likelihood tests		0.02 (reject @5%)	0.01 (reject)	.03 (reject)



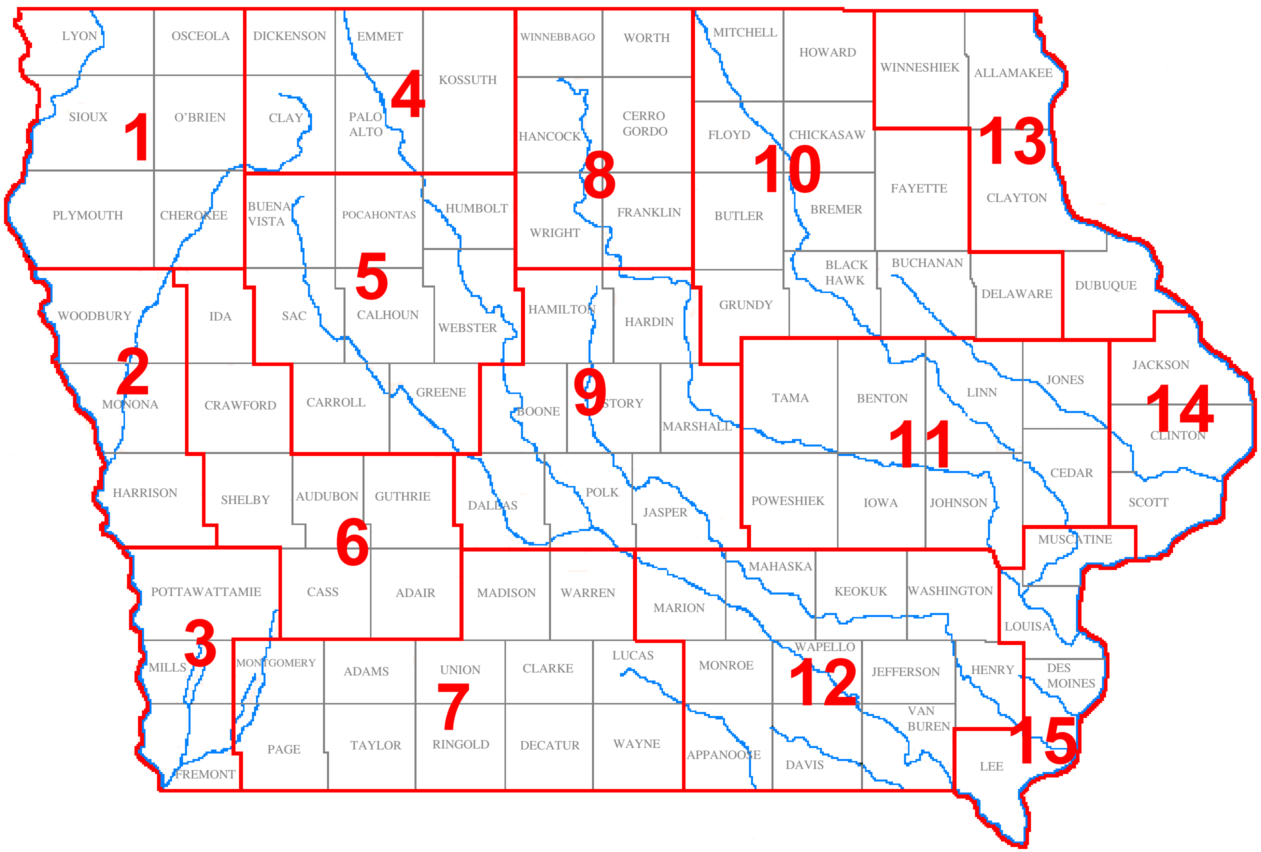
The t-statistics are in parentheses below the coefficient estimates.
Double asterisks indicate tests for difference from 1.00.

Iowa Wetlands Survey

Lowell Washburn

IOWA STATE UNIVERSITY

Wetland Usage Information Gathered by Zones



Respondents were asked:

- The number of trips made to each zone (traditional RP data)
- Would they still take any trips if cost of access were higher? (SP data, discrete)
- How their number of trips would change with increased costs (SP data, continuous)

Alternative Interpretation of Results

1. RP Lovers

- Results prove SP biased (RP/SPc reject consistency)

- RP/SPd fails to reject, but that is due to low information content; shows how insidious SP data can be! Can trick SP lovers into feeling confident when shouldn't.

- Conclusion: Use RP data to estimate welfare measures, could jointly estimate to get efficiency gains, but probably not worth the trouble

Alternative Interpretation of Results

2. SP Lovers

- Results prove RP biased (RP/SPc reject consistency)
- RP/SPd fails to reject, but that is due to low information content
- Conclusion: Use SP data to estimate coefficients and compute welfare, could jointly estimate to get efficiency gains, but probably not worth the trouble

Alternative Interpretation of Results

2a. SP Lovers (less faithful sect)

- RP/SPd doesn’t reject consistency because easier to answer SPd, results indicate SPd may be more accurate

- Conclude: Combine RP/SPd to do welfare estimates

- More research is needed to understand how we should write SP questions for most reliable information.

Alternative Interpretation of Results

3. Agnostics

- These results support the idea that there are problems with both kinds of data: SP and RP, the fact that they are inconsistent indicates that there is bias in one or both

- Conclude: Combine RP/SPc to do welfare estimates, use all information available and throw it in the likelihood function, hope that whatever bias is present in one method is countered by the other and get some efficiency gains.

- More research is needed on all of these issues!

Table II: Summary Statistics

Variable	Mean	Std Dev
X	8.23	10.91
Percent of X positive	66.55%	
X^{new}	2.68	6.27
Percent of X^{new} Positive	27.34%	
P	27.79	12.30
P^{new}	54.24	18.17
Y	40826	25.32

Table I: Bid distribution for the Stated Preference Data

Bid	Number of Surveys
$5	1200
$10	1200
$15	1200
$20	600
$30	600
$40	600
$50	600

Table III: Coefficient and Consumer Surplus Estimates: Independent Models

	RP	SP^D	SP^C
Constant	20.65 (8.21)	15.87 (3.56)	10.39 (2.70)
Price	-0.82 (-8.82)	-0.59 (-6.73)	-0.47 (-5.72)
Income	0.14 (3.84)	0.13 (2.31)	0.11 (.08)
σ	13.28 (18.42)	14.92 (not estimated)	14.92 (10.80)
-log(L)	791.35	134.90	389.45
CS	113.94		197.93

The t-statistics are in parentheses below the coefficient estimates