



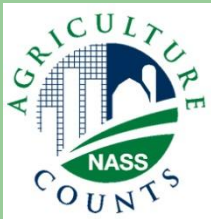
Banff Automated Edit and Imputation Applied to the US Hog Inventory Survey

James Johanson

USDA, National Agricultural
Statistics Service

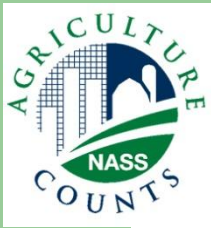


FCSM Conference, November 4, 2013



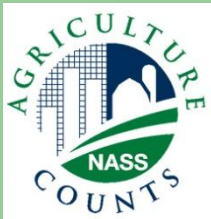
The Agency and Mission

- The National Agricultural Statistics Service (NASS) is a statistical agency located under the U.S. Department of Agriculture (USDA).
- Mission: To provide timely, accurate, and useful statistics in service to U.S. agriculture.



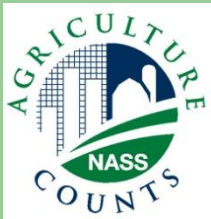
Some Characteristics of E&I in NASS Sample Surveys

- It's labor intensive, especially for large diverse surveys
- It's primarily a manual edit review and updating process (CASIC/Blaise for Hogs)
- It's not necessarily consistent across analysts



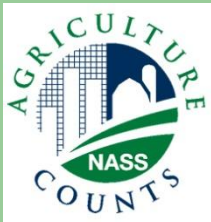
Goals of Automated E&I

- Reduce the time and effort spent manually reviewing or correcting survey questionnaires, without damaging the quality of the resulting data.
- More consistent correction of questionnaires.
- Focus the manual effort on the accuracy of the survey questionnaires that strongly impact the overall results.



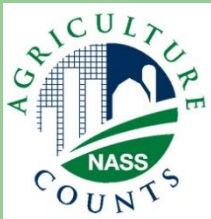
Banff System

- NASS is evaluating Banff (Statistics Canada) to perform the automated statistical data edit and imputation.
- Edit: uses Fellegi-Holt methodology (linear edits define feasible region), which attempts to satisfy all edits by changing the fewest possible values.



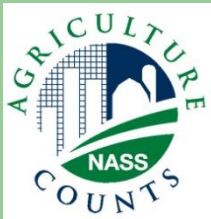
Banff System

- Edits written for commodity values NOT administrative codes.
- Imputation: various alternatives including deterministic, donor, mean based on current data, and previously reported data.
- Interfaces directly with SAS



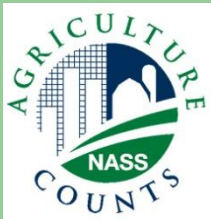
Hog Survey

- Provides detailed inventory of breeding and marketing hogs and the future supply of market hogs.
- Performed on a quarterly basis (December, March, June, September).
- December is the base month and includes all states; top 29 hog states in other quarters.



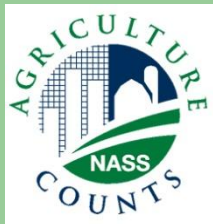
Hog Survey – Banff Testing

- Captured original (raw) survey data (before manual edit) in several top hog producing states.
- After traditional survey processing was complete, Banff was run using the original survey data. Results were compared to the traditional survey results (Control).

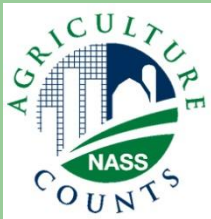


3 Analyses

- **All Data** – Entire survey processed through Banff as one large batch.
- **Daily Batches** – Banff run on each cumulative day batch of data collection.
- **Reviewed** – After daily batches, then used selective editing to review Banff changes in Results Viewer (not part of Banff system).

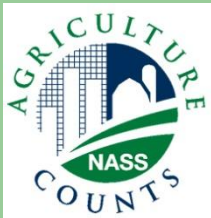


Nothing beats good data collection!!!



Outline

- December 2012 Hogs Quarterly Survey
 - Data flow thru Banff
 - Counts of imputations done
 - Review of estimates



All Hog Questionnaires 8892

Blank 5881
-Non-Response
-Valid Zero

Banff 3011

Pass Banff edits
-Donors 2122

817
Pass Banff edits
after Banff
corrections

Banff tried but
failed 72

Largest Stratum=96 obs

Restore to Original

Review
Impact Records
155 of 721

Manual
correction

Then all must to go thru some administrative edit

December 2012

Quarterly States

Records w/ errors	889
Fields to impute	1864
<hr/>	
Deterministic	180
Donor Imputation	1296
Error Localization	
Deterministic	7
Current (Mean)	219
Error Localization	
Deterministic	6
Previous Report	123
Total Attempted Fixes	1831

Records w/ errors

72

Reduction

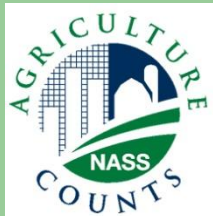
91.9%

Fields to impute

283

84.8%

Red = Record level Black = Item level



Section 1 - Hogs and Pigs Owned

1. On June 1, did this operation (named on label) Own any Hogs or Pigs, regardless of location? (Including hogs or pigs being raised under contract for you by someone else.)

Yes - [Go to Item 2.] No -

Did this operation Own Hogs or Pigs at any time from March 1 through May 31, 2009?

Yes - [Go to Item 7.]
 No - Were any Hogs or Pigs Owned by Someone Else on this operation on June 1?
 Yes - [Go to Item 10a.]
 No - [Go to Section 2 on back]

2. How many Sows and Gilts for Breeding were owned by this operation on June 1? (Including unweaned gilts intended for breeding.)

Number Owned June 1
301 LHOGSILT

a. How many of the [Item 2] Sows and Gilts are Expected to Farrow during June, July, or August?

331 LHGEXP13

b. How many of the [Item 2] Sows and Gilts are Expected to Farrow during September, October, or November?

332 LHGEXP46

3. How many Boars and Young Males for Breeding were owned by this operation on June 1? (Including unweaned boar pigs intended for breeding.)

302 LHOGBOAR

4. Of the Hogs and Pigs for Market and Home Use owned by this operation on June 1, how many were in each of the following four weight groups? (Excluding breeding hogs and pigs reported in [Item 2] or [Item 3].)

a. Under 60 pounds? (Including unweaned pigs intended for market or home use.)

311 LHGUND60

b. 60 - 119 pounds?

312 LHGTO119

c. 120 - 179 pounds?

313 LHGTO179

d. 180 pounds and over? (Including sows and boars no longer used for breeding.)

314 LHGOV180

5. [Add Items 2 + 3 + 4a + 4b + 4c + 4d and verify the total. If necessary, make corrections before continuing.]

300 LHOGTOTL

Then the Total Hogs and Pigs owned by this operation on June 1 was:

Out-Of-State Hogs and Pigs

6. Did this operation own any hogs or pigs in another State on June 1?

Yes - Enter Code 1.
 No - Enter Code 3.

Code
321 LHGMULST

[If Yes, verify this report includes ONLY hogs owned in this State. Complete a separate report for hogs owned in each of the other States.]

Completion Code - Hogs Owned

1 - Incomplete, Owns Hogs	598
2 - Incomplete, Unknown	
3 - No Hogs Owned	MPRESINV

Hogs and Pigs Owned (Continued)

Farrowings and Pig Crop from March 2009 through May 2009

7. Did any Sows or Gilts owned by this operation Farrow during the Last Three Months? (March - May)

Yes - [Complete the following for each month starting with the most recent month]

a. How many Sows and Gilts Owned by this Operation Farrowed during (month)?

Farrowings		
May Sows	April Sows	March Sows
888 LHOGFAR3	891 LHOGFAR2	894 LHOGFAR1

b. How many of the Pigs from these (Item a) litters were:

(i) Owned by this operation on June 1?

(ii) Sold or Slaughtered before June 1?

No - [Continue]

Pig Crop		
May Pig Crop	April Pig Crop	March Pig Crop
889 LHOGPIG3	892 LHOGPIG2	895 LHOGPIG1
890 LHOGSLD3	893 LHOGSLD2	896 LHOGSLD1

Completion Code - Farrowings

1 - Incomplete, Has Farrowings	599
2 - Incomplete, Farrowings Unknown	
3 - No Farrowings	MPRESFAR

Death Loss from March 2009 through May 2009

8. How many Weaned Pigs and Older Hogs Owned by this Operation Died during March, April, and May 2009?

335 LHGDET13

Contract Hog and Pig Production

9. Were any Hogs or Pigs owned by this operation being raised Under Contract by another person or firm on June 1?

Yes - a. How many Producers were raising hogs or pigs for you under contract on June 1?

317 LHGPRODU

b. How many Hogs and Pigs (owned by this operation) were these (Item a) Producers raising for you under contract on June 1?

333 LHGCONTR

[Verify that these hogs and pigs ARE included in the (Item 5) Total, then continue.]

No - [Continue]

10. Were any Hogs or Pigs Owned by Someone Else on this operation on June 1?

Yes - a. How many Hogs and Pigs Owned by someone else were on this operation on June 1?

322 LHGUNCNT

b. Who owns the hogs and pigs?

c. Is this hog owner a contractor?

Yes No

[Verify that these hogs and pigs ARE NOT included in Item 5 on the previous page.]

[Complete a separate questionnaire for these Item a hogs.]

No - [Continue]

(Owner's Name, Address, & Phone Number)

Name _____

Address _____

City, St., Zip _____

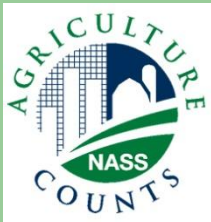
Phone _____

[Complete Section 2 only if the operation shown on the label DOES NOT own hogs or raise hogs under contract, otherwise go to Section 3.]

Completion Code - Hog Presence

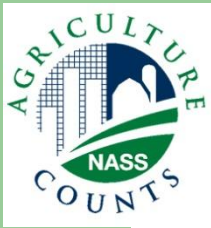
1 - Incomplete, Hogs Present	499
2 - Incomplete, Unknown Presence	
3 - Valid Zero	MPRESHOG

Variable	Corrections	Percent
Death Loss	421	26.76
Litter Rate 3 mo ago	213	13.54
Pigs born and kept 3 mo ago	173	11.00
Total hog inventory	141	8.96
Sow that farrowed 3 mo ago	119	7.57
Litter Rate 2 mo ago	112	7.12
Litter Rate 1 mo ago	97	6.17
Pigs born and kept 2 mo ago	64	4.07
Pigs born and kept 1 mo ago	63	4.01
Expected farrowings next quarter	56	3.56
Expected farrowings in 2 quarters	38	2.42
Hogs under 50 pounds	16	1.02
Sows and gilts for breeding	15	0.95
Hogs raised under contract	9	0.57
Hogs between 50 and 119 pounds	8	0.51
Sow that farrowed 2 mo ago	8	0.51
Hogs over 180 pounds	6	0.38
Sow that farrowed 1 mo ago	6	0.38
Hogs between 120 and 179 pounds	4	0.25
Number of contract producers	2	0.13
Boars for breeding	2	0.13
TOTAL	1573	100.01



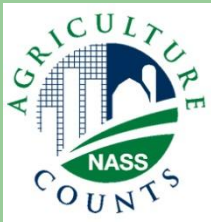
Largest Stratum Omitted

- Above
 - One large batch (All)
 - Included largest stratum
- Below
 - Daily batches
 - Omits largest stratum
- Why Batches?
 - Editing ongoing during data collection



Manual Review with Selective Editing Results Viewer

- December 2012 Quarterly
 - Daily batches (followed by)
 - Reviewed top 20% of each state
 - At least one in each state



Selective Editing

- An item-level score is calculated for specified items on the questionnaire.
- Item-level score = $w * |\text{original} - \text{Banff}| / T$
 - weight and total from the previous survey is utilized
- Record-level score = the largest item-level score for the record (Unit Score).



Quarterly Hogs and Pigs

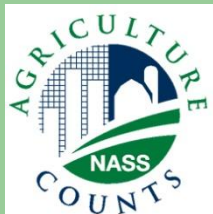
Click on a column to sort table

POID & SUBTRACT	STRATUM	EXPANSION FACTOR	FINAL WEIGHT	REVIEWED	UNIT SCORE
	92	1	1.2570733864	Yes	1.2535967655
	88	1.448	3.864285423	Yes	0.0370953143
	92	1	1.2570733864	Yes	0.0180552794
	80	60.6	92.552727273	Yes	0.0112122975
	88	1.448	3.864285423	Yes	0.0104488568
	82	5.2173913043	8.3094198378	Yes	0.0100959382
	98	1	1	Yes	0.0089792653
	80	60.6	92.552727273	No	0.0049056018
	90	1	2.047826087	No	0.0042372719
	98	1	1	No	0.0034666655
	98	1	1	Yes	0.0029438228
	98	1	1	No	0.0027717969
	98	1	1	No	0.0026167314
	98	1	1	No	0.0026167314
	92	1	1.2570733864	Yes	0.0026010811
	98	1	1	No	0.0025876566

Select Variables

- | | | |
|---|---|---|
| <input type="checkbox"/> All | <input type="checkbox"/> Breeding Stock | <input type="checkbox"/> Hogs Raised By Others |
| <input checked="" type="checkbox"/> Unit Score Only | <input type="checkbox"/> Market Hogs | <input type="checkbox"/> Contractees |
| <input type="checkbox"/> Total Hogs and Pigs | <input type="checkbox"/> Total Deaths | <input type="checkbox"/> Hogs Raised For Others |
| <input type="checkbox"/> Total Farrowings | <input type="checkbox"/> Pig Crop Still Owned | <input type="checkbox"/> Pig Crop Sold |

Review



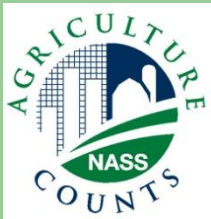
SignEdit System

Survey
Month
Year
Reporting Unit (921)
Response Code (

Reviewed
Operation Name
Person Name
Stratum
Unit Score
National Rank

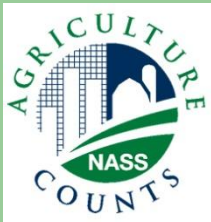
ItemCode	MasterVarName	VarNameDescription	previous	original	banff	current
301	LHOGGILT	SOVS GILTS AND YOUNG GILTS FOR BREEDING	.	-1	-1	0
331	LHGEXP13	SOVS EXPECTED TO FARROW 1 QUARTER FROM NOW	.	-1	-1	0
332	LHGEXP46	SOVS EXPECTED TO FARROW 2 QUARTERS FROM NOW	.	-1	-1	0
302	LHOGBOAR	BOARS AND YOUNG MALES FOR BREEDING	.	-1	-1	0
315	LHGUND50	MARKET HOGS AND PIGS UNDER 50 POUNDS	.	-1	-1	0
316	LHG50119	MARKET HOGS AND PIGS 50 TO 119 POUNDS	.	-1	-1	0
313	LHGTO179	MARKET HOGS AND PIGS 120 TO 179 POUNDS	.	-1	-1	0
314	LHGOV180	MARKET HOGS AND PIGS OVER 180 POUNDS	.	-1	-1	0
300	LHOGTOTL	TOTAL HOGS AND PIGS	.	-6	-6	0
888	LHOGFAR3	SOVS FARROWED 1 MONTH AGO	.	-1	-1	0
889	LHOGPIG3	PIGS ON HAND 1 MONTH AGO	.	-1	-1	0
890	LHOGSLD3	PIGS SOLD OR SLAUGHTERED 1 MONTH AGO	.	-1	-1	0
891	LHOGFAR2	SOVS FARROWED 2 MONTHS AGO	.	-1	-1	0
892	LHOGPIG2	PIGS ON HAND 2 MONTHS AGO	.	-1	-1	0
893	LHOGSLD2	PIGS SOLD OR SLAUGHTERED 2 MONTHS AGO	.	-1	-1	0
894	LHOGFAR1	SOVS FARROWED 3 MONTHS AGO	.	-1	-1	0
895	LHOGPIG1	PIGS ON HAND 3 MONTHS AGO	.	-1	-1	0
896	LHOGSLD1	PIGS SOLD OR SLAUGHTERED 3 MONTHS AGO	.	0	0	0
335	LHGDET13	HOG DEATHS 1 QUARTER AGO	.	-1	0	0
333	LHGCONTR	HOGS UNDER CONTRACT FOR ANOTHER	.	0	0	0
317	LHGPRODU	NUMBER OF HOG CONTRACTEES	.	0	0	0
322	LHGUNCNT	HOGS OWNED BY SOMEONE ELSE	.	0	0	0
598	MPRESINV	HOG INVENTORY COMPLETION CODE	.	.	.	1
599	MPRESFAR	HOG FARROWINGS COMPLETION CODE	.	.	.	1
499	MPRESHOG	HOG COMPLETION CODE	.	.	.	1
492	LHOGINTS	HOGS INTENTIONS	.	.	.	0
488	LHGFACIL	HOG FACILITIES PRESENCE	.	.	.	0

Adds transparency to the automated imputation



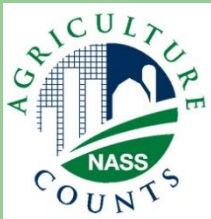
Data sent to Summary

- Estimates
 - Looked at 12 key estimates
- Standard Errors
 - Used to calculate confidence intervals
- Coefficients of Variation (CVs)



Estimates Definitions

- Test1 – Does the Control estimate fall within the Banff 95% confidence interval?
- Test2 – Does the Banff estimate fall within the Control 95% confidence interval?
- Percent relative change in the estimate from the Control to the Banff. (>10% and >5%)



348 Estimtes (29 states * 12)

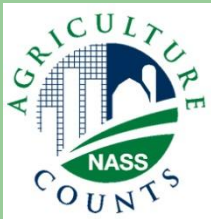
Analysis	Test1	Test2	>10%	>5%
All	7	9	3	8
Batches	8	11	1	8
Reviewed	8	10	1	3

Investigate State Level Estimtes

State	Variable Description	Reviewed			Batch	All	
		U+	Test1	Test2	% Rel Change		
1	Expected farrowings in 2 quarters	9	NO	YES	-1.34	-1.34	1.34
8	Pigs born last quarter	13	YES	NO	0.32	0.32	0.37
8	Farrowings last quarter	13	YES	NO	0.30	0.30	0.38
17	Death loss	353	NO	NO	3.28	3.66	3.41
26	Hogs over 180 pounds	52	YES	NO	7.11	7.11	7.11
26	Death loss	62	NO	NO	5.03	5.03	10.65
27	Death loss	452	NO	NO	-2.52	-2.52	-2.18
29	Death loss	120	NO	NO	0.80	0.80	0.47
30	Death loss	43	YES	NO	3.63	3.63	4.67
40	Expected farrowings next quarter	18	YES	NO	0.15	0.15	0.00
48	Boars for breeding	31	YES	YES	23.23	23.23	23.23
48	Market hogs under 50 pounds	25	YES	YES	-0.33	5.25	5.25
48	Pigs born last quarter	23	YES	YES	2.68	6.16	6.33
48	Farrowings last quarter	23	YES	YES	1.61	5.06	5.06
51	Death loss	9	NO	YES	-0.66	-0.66	-0.66
51	Pigs born last quarter	7	YES	YES	0.33	6.02	6.99
51	Farrowings last quarter	7	YES	YES	2.37	7.33	12.28
56	Pigs born last quarter	5	NO	YES	0.10	0.10	0.10
56	Farrowings last quarter	5	NO	NO	0.12	0.12	0.07

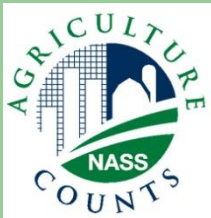
National Estimates

Variable Description	Reviewed			Batch	All	
	U+	Test1	Test2	% Rel Difference		
Sows for breeding	1958	YES	YES	0.01	0.04	0.05
Boars for breeding	1647	YES	YES	2.42	25.22	24.62
Expected farrowings next quarter	1774	YES	YES	0.08	0.13	-0.01
Expected farrowings in 2 quarters	1703	YES	YES	-0.06	-0.02	0.11
Hogs under 50 pounds	2365	YES	YES	0.09	0.05	-0.02
Hogs between 50-119 pounds	2324	YES	YES	0.01	0.03	0.03
Hogs between 120-179 pounds	2189	YES	YES	0.01	0.01	-0.02
Hogs over 180 pounds	2444	YES	YES	0.12	0.17	0.18
Total hog inventory	3301	YES	YES	0.06	0.10	0.08
Farrowings last quarter	1715	YES	YES	0.24	0.41	0.35
Pigs born last quarter	1715	YES	YES	0.21	0.27	0.32
Death loss	2904	YES	YES	0.24	0.46	0.38



1 person in 1 week?

- Significance Editing Review – 9 hrs 15 mins
 - 721 Cleaned by Banff, put in viewer
 - 155 reviewed
 - 108 were changed
- For reference, South Dakota's sample size was 400.



All Hog Questionnaires 8892

Blank 5881
-Non-Response
-Valid Zero

Banff 3011

Pass Banff edits
-Donors 2122

817
Pass Banff edits
after Banff
corrections

Banff tried but
failed 72

Largest Stratum=96 obs

Review
Impact Records
155 of 721

Restore to Original

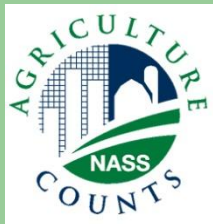
Manual
correction

Then all must to go thru some administrative edit



Goals of Automated E&I

- Reduce the time and effort spent manually reviewing/correcting survey questionnaires, without damaging the quality of the resulting data.
- More consistent correcting of questionnaires.
- Focus the manual effort on the accuracy of the survey questionnaires that strongly impact the overall results.



Thank you!

James_Johanson@nass.usda.gov

<http://www.nass.usda.gov/>