

The Effect of Reporting Mode on Administrative Records:

Are we sacrificing quality for convenience?



2012 Federal CASIC Workshops

Danni Mayclin and Marilyn Worthy

March 29, 2012 | Washington, DC

Overview

- Survey background
- Mode selection
- Data quality evaluation by mode
- Recommendations

Survey Background:

Residential Energy Consumption Survey (RECS)

- Survey of energy characteristics, usage, and costs in U.S. homes
- Conducted periodically since 1978, most recently in 2009
- Two phases of data collection
 - household interview (CAPI)
 - *Energy Supplier Survey (ESS)*
- ESS is a network sample of the companies the household respondents say provide their energy

Survey Background:

Why do we conduct the ESS?

- Combining household characteristics data with energy consumption data allows EIA to estimate fuel and end use consumption
 - ex: average amount of natural gas used for space heating for homes in Virginia
 - unique data product
- Difficult task for household respondents to report their energy consumption and expenditures for 20 months
- Companies (and specifically, their records) are the best source of this data

Survey Background:

Energy Supplier Survey (ESS)

- Target data: usage and cost data for all energy sources used in RECS housing units for 20 months (09/08 – 04/10)
- Companies are heterogeneous groups that range in size, energy sources supplied, record systems, etc.
 - 3% of responding companies reported more than half of ESS data
- ESS is mandatory for companies
- Response rate
 - cases: 90% (17,770/19,647)
 - companies: 90% (1,227/1,363)

Survey Background: 2009 ESS Data Collection Process


- For the first time, 2009 ESS used the internet for data collection and offered respondents their choice of modes
 - Internet data collection and new modes were result of cognitive interviews and pretesting with previous and potential respondents
 - Goals: reduce cost, manage burden for both companies and us
- Mailing with instructions on how to access ESS website
- ESS website provided list of cases as well as mode options

Mode Selection: Options

1. Paper form
 - mail or fax
2. Online form
3. Excel template
4. Other
 - other electronic file
 - non-standard printout

Note: Companies could submit cases in more than one mode

Sample paper form for electricity



U.S. Energy Information Administration

Form EIA-437E
OMB No. 1905-0092, Expires 1/31/13

2009 RECS Electricity Usage Form


Service address:

Account Number:


Please provide information on electricity usage for this service address between
September 2008 and April 2010.

Billing Period	Enter the End Date for each billing period	Enter the Amount used in kWh	With what:			Enter the Total Dollar Amount including taxes <small>(Exclude into fees, merchandise, repairs, and service charges)</small>	Electricity was:		
	MM/DD/YY	XXXX	A=Actual	E=Estimated	R=Read by Customer	\$ XXX.XX	B=Both Sold and Delivered	S=Sold Only	D=Delivered Only
			(select one)				(select one)		
1	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	□□/□□/□□		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	\$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments



Sample online form for electricity



U.S. Energy Information Administration

Form EIA-457E
OMB 1905-0002
Exp. 1/31/2013

2009 RECS Electricity Usage Form

Service address: 123 Main St., Newhaverille, US 98765 Account Number: _____

Please provide information on electricity usage for this service address between **September 2008** and **April 2010**.
Use the Tab key (next column) and arrow key (next button) to navigate through the form. Avoid the Enter key.

Billing Period	Enter the End Date for each billing period MM/DD/YY	Enter the Amount Used in kWh XXXX	kWh were:			Enter the Total Dollar Amount including taxes [Exclude late fees, merchandise, repairs, and service charges] \$XXX.XX	Electricity was:		
			A-Actual	E-Estimated	R-Read by Customer		B-Both Sold and Delivered	S-Sold Only	D-Delivered Only
1	01/01/08	075	<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$ 90	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
2	01/01/08	150	<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$ 90	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
3	02/01/08	200	<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$ 20	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
4		200	<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$ 25	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
5			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
6			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
7			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
8			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
9			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
10			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
11			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
12			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
13			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
14			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
15			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
16			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
17			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
18			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
19			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D
20			<input type="radio"/> A	<input type="radio"/> E	<input type="radio"/> R	\$	<input type="radio"/> B	<input type="radio"/> S	<input type="radio"/> D

Comments:

Supplier ID: 246
EIA Case ID: 31111194L1

Sample Excel template for electricity

ESS Website Electricity Spreadsheet V2_012811.xls [Compatibility Mode] - Microsoft Excel

Home Insert Page Layout Formulas Data Review View JMP Acrobat

Paste Clipboard Font Alignment Number Styles Cells Editing

B47

	A	B	C	D	E	F	G	H	I	J	K
	EIA_Case_ID	Service_Address	Service_Unit_Number	Service_City	Service_State	Service_Zip	Account_Number	End_Date	KWH	A_E_R	Cost
1	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
2	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
3	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
4	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
5	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
6	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
7	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
8	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
9	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
10	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
11	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
12	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
13	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
14	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
15	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
16	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
17	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
18	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
19	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
20	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
21	100-EL1	100 Lake St	101	Anytown	US	10000	10000				
22	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
23	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
24	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
25	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
26	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
27	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
28	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
29	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
30	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
31	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
32	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
33	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
34	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
35	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
36	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
37	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
38	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
39	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
40	200-EL1	200 Beech Ave		Anytown	US	10000	20000				
41	200-EL1	200 Beech Ave		Anytown	US	10000	20000				

Variable Descriptions Service Address List Reporting Template

Editing Overview

Post-data collection edits

- Included comments, missing data, outliers, or inconsistent data patterns
- Cases that had edit failures were manually reviewed

Tools in deciding whether to make changes

- ESS respondent comments
- Scanned energy bills collected during the household interview
- Data from the household survey, such as housing unit type, main heating fuel, square footage of the housing unit, move-in date, and respondent comments

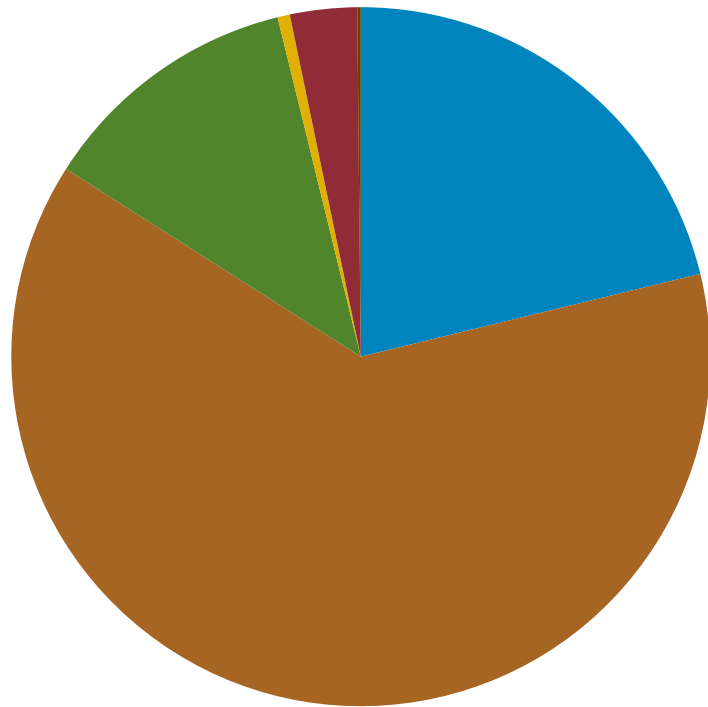
Purpose of this analysis

To evaluate:

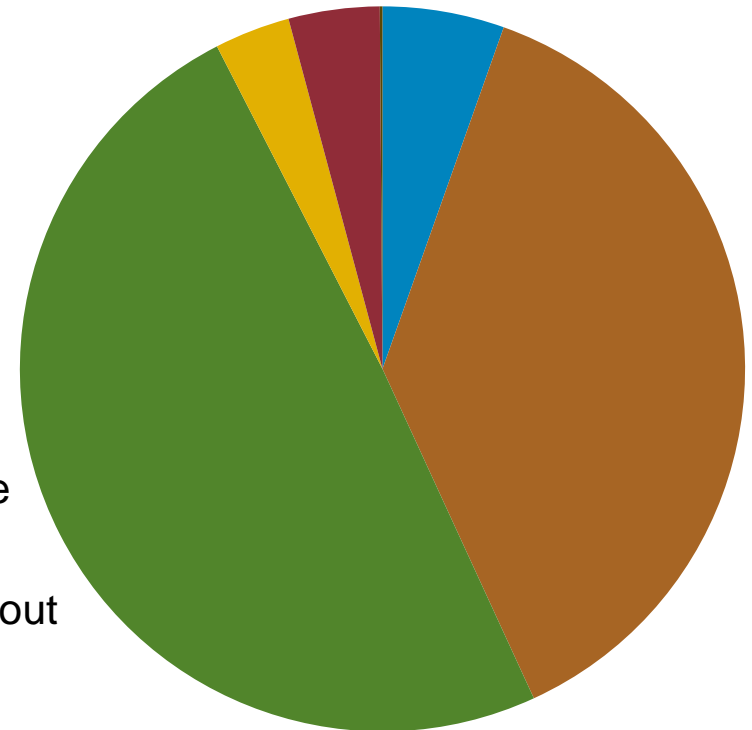
- Why do companies choose certain response modes?
- Does the mode selected affect the quality of the reported administrative data?
- What modes should be considered for future ESS cycles to balance costs and quality?

Most companies chose to use the online form, but about half of all cases were submitted by Excel template

ESS companies



ESS cases

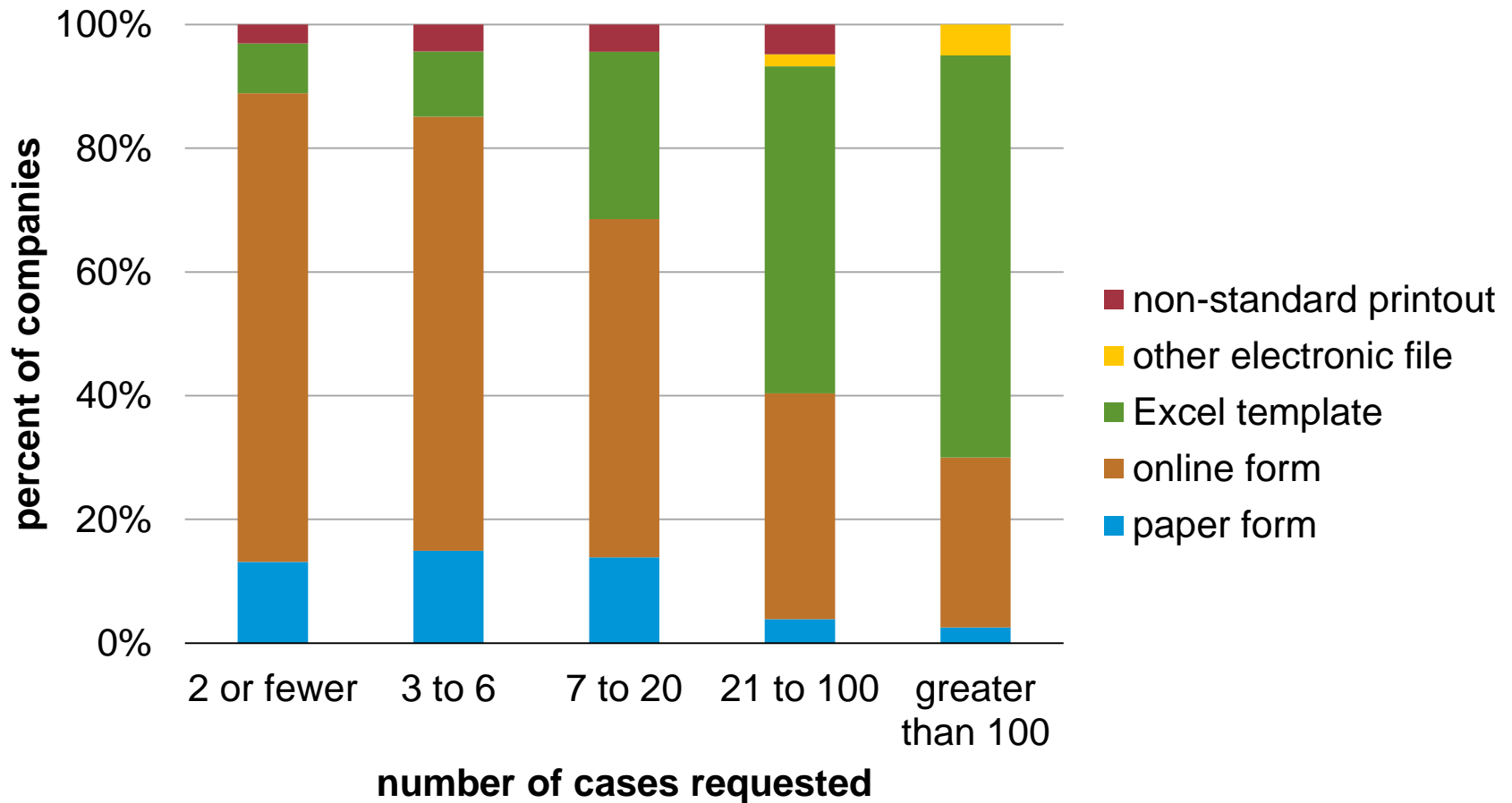


- paper form
- online form
- Excel template
- other electronic file
- non-standard printout

What affects mode selection?

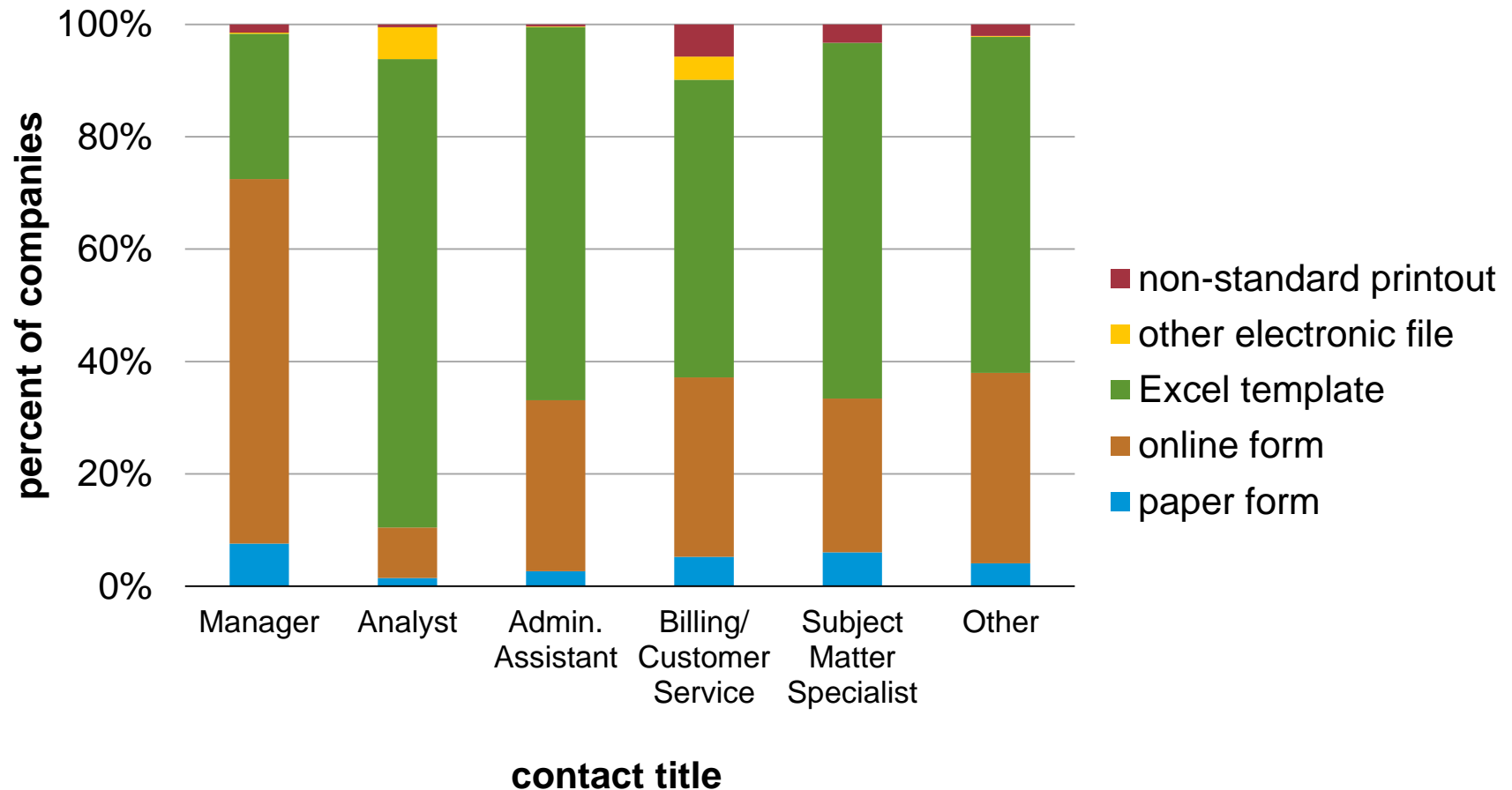
- Size: number of cases requested
- Access: contact's position in company affects access to records
- Size, access, and other variables are confounding variables, as companies selected their mode

Size: Companies with fewer cases preferred online forms, while companies with more cases chose Excel template



*Chart limited to companies that had a choice of all reporting mode options.

Access: Managers were most likely to choose online forms while analysts were most likely to use the Excel template



*Chart limited to companies that had a choice of all reporting mode options.

How did we evaluate quality?

1) Completeness

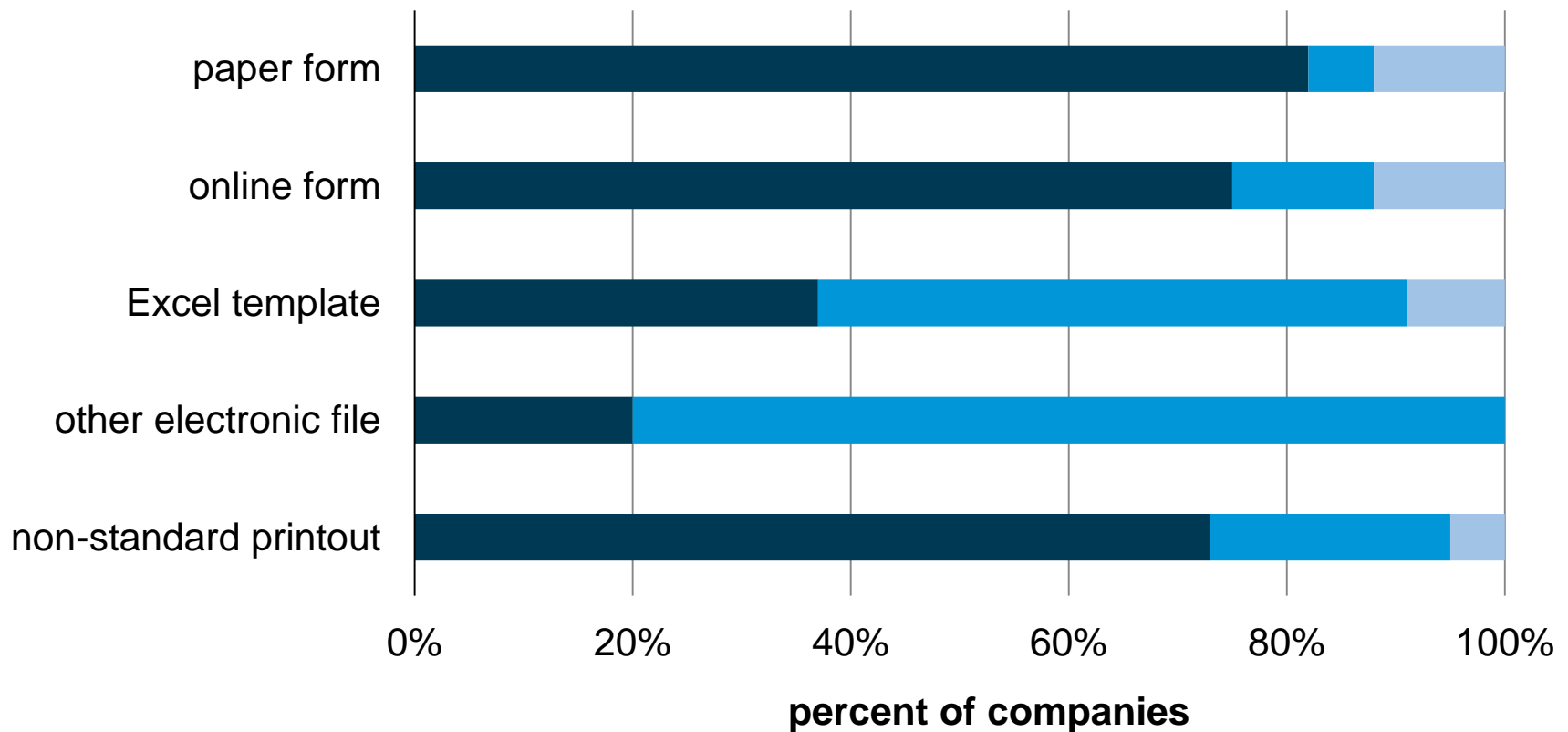
- Percentage of requested cases submitted (unit non-response)
 - Did companies submit all of the cases we requested from them?
- Completeness of submitted cases (item non-response)
 - For the cases that were submitted, did we receive data for the full time period (20 months) requested?

2) Correctness

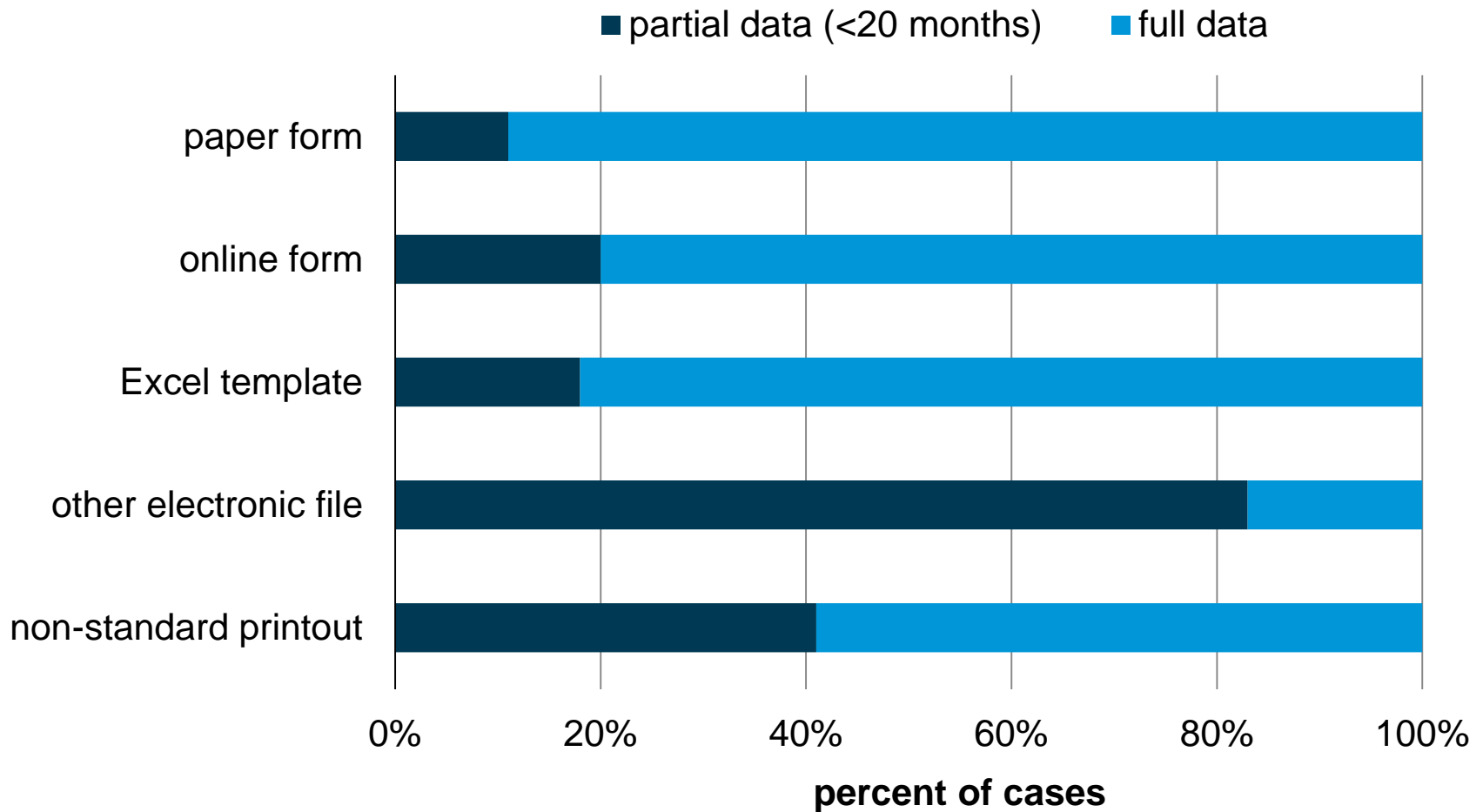
- Percentage of cases with edit failures
- Percentage of cases with data changes made during editing
 - Direct reflection of errors made by companies

Completeness: Most companies that chose to use the Excel template or other electronic file did not submit data for all cases

■ submitted 100% of cases ■ submitted 75% - 99% of cases ■ submitted less than 75% of cases



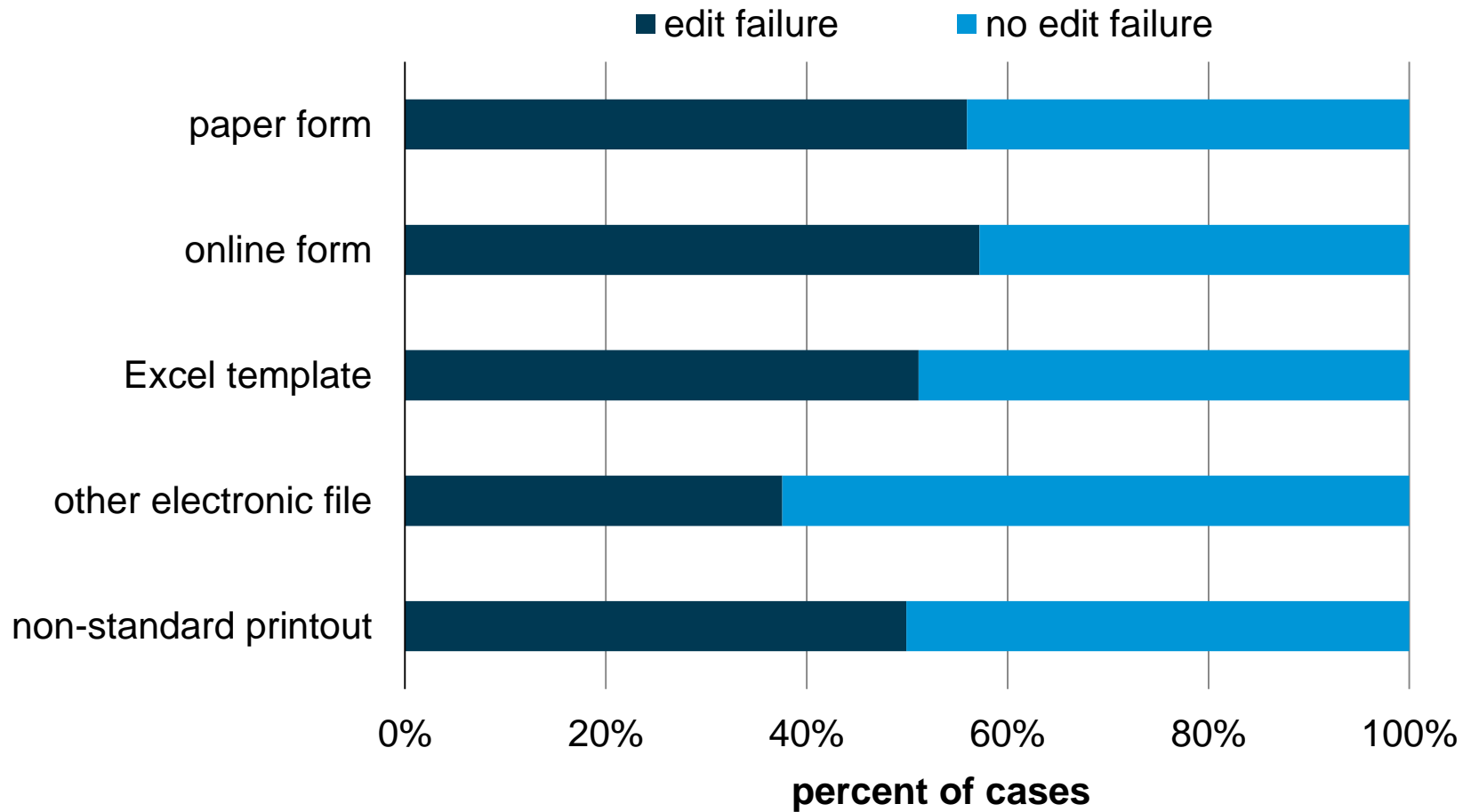
Completeness: Partial data submissions were most common with other electronic files and least common with paper forms



Completeness: “Takeaways”

- Companies using **paper forms** were best at submitting all cases and submitting data that covered the full time period.
- Companies using **Excel template** or **other electronic file** were most likely to “miss” cases.
- **Other electronic files** and **non-standard printouts** had lots of partial data.

Correctness: Individual forms higher case-level edit failure rate than other modes



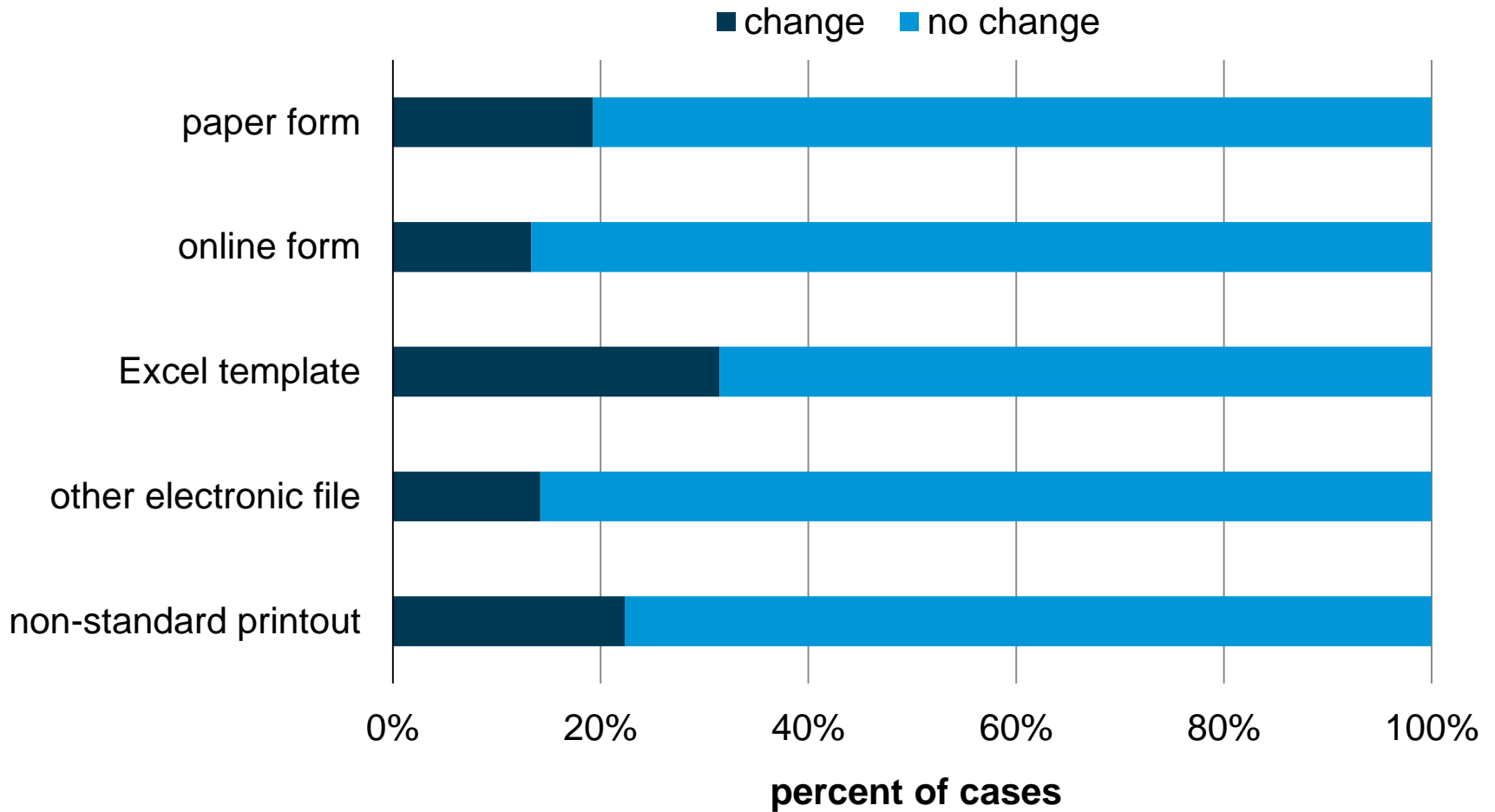
Correctness: The highest edit failure rate occurred when the largest companies submitted online forms

case-level edit failure rate by mode

	paper form	online form	Excel template	other elec. file	non-std. printout
2 or fewer	59	35	S	S	S
3 to 6	56	46	54	S	S
7 to 20	69	46	54	S	49
21 to 100	56	48	46	S	70
greater than 100	37	73	54	33	34

S = less than 75 cases or less than 5 companies

Correctness: Cases submitted via Excel template required the most data changes during editing



Correctness: Companies with a medium-sized caseload were most error-prone, as shown by the frequency of data changes required

percent of cases with data changes by mode

	paper form	online form	Excel template	other elec. file	non-std. printout
2 or fewer	17	7	S	S	S
3 to 6	17	11	23	S	S
7 to 20	28	10	39	S	41
21 to 100	18	19	34	S	36
greater than 100	15	14	30	7	15

S = less than 75 cases in category or less than 5 companies

Correctness: “Takeaways”

- Edit failures were most common for cases submitted via **online form**, increasing our editing burden. However, these edit failures did not correspond to many data changes.
 - May have been cleaner because certain data values were not accepted when submitting online.
- Cases submitted via **Excel template** had the highest level of data changes, indicating those cases had the most error.
- Appear to be interactions between data request size, mode, and quality.

Conclusions

- **Paper forms** had the least missing cases, the most complete data, but needed some data changes. They also require keying, which adds to our cost and burden.
- **Online forms** required the least data changes of any mode, and had a high level of completeness. They were chosen by the most companies.
- **Excel template** submissions had more missing cases, less complete data, and needed more data changes than other modes. However, this was the most common mode for companies with the largest burden (in terms of cases).
- **Other electronic files** were the least complete, but didn't need many data changes.
- **Non-standard printouts** had low levels of completeness and correctness.

Considerations for future ESS

- Our homework
 - Evaluate editing process in attempt to decrease number of false positives
 - Evaluate editing process by mode; are same edits needed for each mode?
- Steer companies to modes based on size, access, or other variables
- Suggestion: When companies submit an Excel template, ask a few key questions to help identify errors prior to editing.
 - e.g. Are taxes included in your cost figures?

Future research

- Qualitative research to better understand response mode decision process
- Sensitivity analysis to determine whether a different sample would produce different results

For more information

RECS Website <http://www.eia.gov/consumption/residential/>

Danni Mayclin Danielle.Mayclin@eia.gov

Marilyn Worthy Marilyn.Worthy@eia.gov