

# Findings and Recommendations from Usability Testing for the 2022 Economic Census Item 22 (NAPCS)

---

Prepared for:

2022 Economic Census Stakeholders

Prepared by:

Melissa A. Cidade, Ph.D., Data Collection Methodology and Research Branch

Krysten Mesner, Data Collection Methodology and Research Branch

Office of Economic Planning & Innovation

Economic Programs Directorate

U.S. Census Bureau

February 8, 2021

The Census Bureau has reviewed this report for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied. (Approval ID: CBDRB-FY21-ESMD002-014)

## Table of Contents

Executive Summary.....	3
Research Objectives and Background.....	4
Research Methodology .....	5
Respondent Overview.....	6
Round 1 Overview.....	7
Round 2 Overview.....	10
Findings and Recommendations.....	11
Round 1 Findings and Recommendations .....	11
Finding #1: Respondents were confused by the task. ....	11
Finding #2: The search functionality is unclear to respondents. ....	13
Finding #3: Respondents are not always clear on what to do after entering a product. ....	14
Finding #4: Some search results are not salient to respondents.....	15
Finding #5: Respondents are confused by the looping search results for multiple entries. ....	16
Round 2 Findings and Recommendations .....	17
Finding #6: Search functionality cues help orient respondents. ....	17
Finding #7: Separating navigation buttons from a search button helps orient respondents.....	17
Finding #8: Respondents struggled with the ‘Add selected products/services’ button.....	18
Finding #9: The layout of the search results screen limits respondents’ answers. ....	19
About the Data Collection Methodology and Research (DCMR) Branch.....	20
Appendix A: Interview Protocol .....	21

## Table of Figures

Figure 1: Round 1 Screen 1, Version 1 Screenshot .....	8
Figure 2: Round 1 Screen 1, Version 2 Screenshot .....	9
Figure 3: Round 1 Screen 2 Screenshot .....	9
Figure 4: Round 2 Screen 1 Screenshot .....	11
Figure 5: Screenshot of Round 2 Screen 1 Instructions .....	12
Figure 6: Screenshot of Round 2 NAICS-specify Screen.....	13
Figure 7: Screenshot of Round 2 Search Screen .....	18

## Executive Summary

Over the course of two months in fall 2020, research staff from the Data Collection Methodology and Research Branch conducted 86 half-hour interviews with businesses across sectors and with varying complexity. The purpose of these interviews was to conduct usability testing on Item 22 of the Economic Census (the North American Product Classification System (NAPCS) question). This testing was divided into two rounds, and the major findings from those rounds of research include:

### Round 1 Findings and Recommendations:

- **Finding #1: Respondents were confused by the task.:** Some respondents were confused by the question altogether. We recommend including instructions on the front page as well as tying the pre-populated list to their NAICS code.
- **Finding #2: The search functionality is unclear to respondents.:** Respondents did not necessarily understand that the text capture box on Screen 1 was a search bar. We recommend including additional instructions on searching for additional products and services.
- **Finding #3: Respondents are not always clear on what to do after entering a product.:** Respondents did not intuit to click the 'Save and Continue' button once they entered an additional product or service. We recommend keeping that button for navigation, and instead using the 'Add products/services not listed' button to activate the search.
- **Finding #4: Some search results are not salient to respondents.:** Respondents did not always find an appropriate match from the search results. We recommend additional refinement of the machine learning algorithm driving search results.
- **Finding #5: Respondents are confused by the looping search results for multiple entries.:** Because respondents were 'looped' through each of their write-in responses, but could select as many products or services as they chose from results screens, they often got lost navigating the different lists. We recommend separating out each search independently.

### Round 2 Findings and Recommendations:

- **Finding #6: Search functionality cues help orient respondents.:** In Round 2, we added additional cues for respondents to use the search functionality, including instructions, and found that respondents were more likely to understand the task at hand. We recommend retaining these instructions.
- **Finding #7: Separating navigation buttons from a search button helps orient respondents.:** We found that respondents navigated through the survey more parsimoniously when the navigation buttons were retained for navigation and the 'Add' button automatically executed a search. We recommend retaining this functionality.
- **Finding #8: Respondents struggled with the 'Add selected products/services' button.:** Once respondents made selections from the search results, we found that they did not consistently use the 'Add' button, leading to those selections not being added to their list. We recommend

making this button only available after a search has been executed, and consider adding an error warning if a respondent has made selections but does not click the 'Add' button.

- **Finding #9: The layout of the search results screen limits respondents' answers.:** Some respondents did not notice that the search results screen contained a scroll bar to see more results, which may lead to suboptimal selections based on what they can see. We recommend moving the 'Close' button and the 'Add' button to the very bottom of the screen, and 'floating' the instructions regardless of where a respondent might be on a screen.

## Research Objectives and Background

The Economic Census is a mandatory survey conducted by the Census Bureau every five years under the authority of Title 13, United States Code (U.S.C.), Section 131. The survey collects data electronically from nearly 4 million establishments (including large, medium and small companies representing all U.S. locations and industries) on a range of operational and financial topics. Data from the survey are used as the official five-year measure of American business and the economy. As part of the Economic Census, the section of the survey that collects detailed information about the revenue that businesses generated by the goods and services they provide using the North American Product Classification System (NAPCS).<sup>1</sup>

Historically, the Census Bureau requested revenue details mainly for goods and services associated with an establishments' primary North American Classification System (NAICS) classification, thus focusing primarily on where goods were produced rather than which goods were produced. However, beginning with the 2017 Economic Census, the Census Bureau introduced a new way of collecting and disseminating the product revenue information through the NAPCS framework. This framework focuses on the demand-side economy, providing a summary of where goods and services are sold, in the primary industry in which they are produced.

Under the NAPCS framework, businesses were asked to explicitly report detailed information on all revenue generating goods and services they offer, even if those goods and services are not typical of that business' industry. Although some goods and services might make up a fraction of the business' revenue, respondents should still report these products, and provide a description of them and their associated revenue.

Item 22 in the 2017 Economic Census - the NAPCS survey question - first presented respondents with a list of products typical in their industry. After selecting products, respondents were then asked to report revenues associated with each of the selected products on a subsequent screen. The survey included additional open-ended response fields where the respondent could write in other revenue generating goods or services that were not pre-listed. This two-stage design takes advantage of embedded

---

<sup>1</sup> Further information regarding the Economic Census can be found on the Census Bureau's website at: <https://www.census.gov/programs-surveys/economic-census.html>. For more information about NAPCS, please visit <https://www.census.gov/eos/www/napcs/>.

automation – basing subsequent questions on previous responses – which is available due to the flexibilities of the electronic mode of the instrument.

This two-stage approach to asking about NAPCS in 2017 still resulted in a substantial number of write in responses, as respondents reported items that they did not believe were captured in the pre-listed descriptions. Each of these write-in responses, in turn, is sent through extensive processing relying on an outsized amount of resources to capture the product or service lines. As such, it is best to reduce the number of write-in responses to reduce the amount of post-hoc data handling and cleaning. To assist in the selection of products/services, and to potentially reduce the number of write-in responses, the upcoming 2022 NAPCS survey item plans to incorporate machine learning functionality. Additionally, the 2022 NAPCS survey item now includes the open-ended response fields to capture items not pre-listed at the first stage of the two question series.

This design proposal differs substantively from prior reporting strategies and therefore requires an evaluation to assess its feasibility prior to implementing in the 2022 Economic Census instrument. Additionally, the incorporation of machine learning and its design elements warrant input from respondents. This report outlines the results of that testing along with recommendations for implementation of the machine learning-based functionality.

Data gathered from testing helps to inform user-centered design decisions for the design of the 2022 NAPCS survey items. This will ensure that respondents can easily report the NAPCS codes most appropriate for them and minimize write-in responses as much as possible. Additional objectives for the evaluation of the online NAPCS survey item in the 2022 Economic Census instrument include the following:

- **Evaluate the performance** of NAPCS in terms of item non-response attributed to its design, functionality and ease of use;
- **Inform the design** of the machine learning functionality so that it meets the needs and expectations of respondents, allowing them to make appropriate selections as frequently as possible; and
- **Provide recommendations** for improvements to the design of the NAPCS survey items that can potentially enhance data quality and respondents' reporting experience.

## Research Methodology

For this testing, we used a usability testing framework to better understand the ways that respondents interact with the updated Item 22 format, including machine learning. Usability for web-based surveys

evolved out of Human Factors research (most notably industrial sociology, with its focus on efficiency and productivity), and came into its own as a field of interest in the early 2000s.<sup>2</sup>

Usability aims to improve data quality by reducing errors and preventing item or unit nonresponse due to respondent burden. Using the framework laid out by the International Organization for Standardization (IOS), the five key components of usability testing include:

1. The product – in this case, the redesigned Item 22 is the ‘product’ of interest for this testing.
2. The specified users of the product – here, we are interested in companies that have been or could be sampled to participate in the 2022 Economic Census.
3. The goals of the users – while the Census Bureau’s ultimate goal may be to reduce write-in responses and increase data quality, the goal of the users – the respondents – must be front and center, chiefly, to complete the survey as quickly and easily as possible.
4. The context of use – the unit of analysis in economic surveys is not the individual or the household, but rather the company; as such, testing should take place in the context of the company, that is, during normal business hours as part of the routine work of the respondent. Usability does not need to reproduce all of the possible contexts in which a respondent answers a survey, but “to the extent that a product can be tested in the primary and most likely context, the better the product will be.”<sup>3</sup>
5. Metrics of evaluation – this includes the ways in which the effectiveness of the design will be assessed, and should be defined before testing begins.

In their work on usability and survey research, Geisen and Romano Bergstrom<sup>4</sup> define three measures of evaluation – effectiveness, efficiency, and satisfaction. Effectiveness can be measured in terms of whether or not users are successfully able to complete specified tasks. Efficiency is the number of steps it takes a respondent to complete a task. Satisfaction is often a self-rated measure or qualitative comment elicited during the testing that demonstrates the respondents’ perceived ease of use and level of frustration with the product of interest.

With this in mind, we conducted the EC22 Item 22 usability testing to investigate the effectiveness and efficiency of the new design, while also aiming to reduce errors and prevent nonresponse and/or text write in responses.

## Respondent Overview

For this research, interviews lasted no more than 30 minutes. We recruited respondents by emailing a link to access an online survey platform whereby respondents could schedule an appointment for the

---

<sup>2</sup> Couper, Mick. 2000. “Usability Evaluation of Computer-Assisted Survey Instruments.” *Social Science Computer Review*, 18(4): 384-396.

<sup>3</sup> Geisen, Emily and Jennifer Romano Bergstrom. 2017. *Usability for Survey Research*. Elsevier, Inc: Cambridge, MA: 8.

<sup>4</sup> Ibid.

interview by selecting a day and time that worked best for their schedule. Once a respondent selected an appointment, they received an automated reminder email on the morning of their interview appointment that confirmed their appointment and encouraged them to fill out the online consent form. About thirty minutes before the scheduled appointment, interviewers sent one final email to respondents with the call-in information and a link to the online testing instrument.

Respondents, interviewers, and observers all met through Skype for Business. During the interview, researchers asked respondents to share their screens so that we could watch them fill out the survey and take notes on the ease of completion. To maintain comparability between interviewers and among interviews, researchers used a protocol to guide the interviews (see Appendix A).

To facilitate iterative research design, this project was split into two rounds of recruitment and interviewing. After the first round of research, key stakeholders gathered to debrief about their observations of the interviews and to work through changes for the second round.

Please note: throughout this report, respondent pronouns have been randomly assigned. All screenshots represent fictitious data.

### Round 1 Overview

For Round 1, researchers conducted a total of 50 interviews between August 2020 and October 2020. For Round 1, the sample frame was all multi-unit firms sampled for the 2017 Economic Census. In this case, we were interested in representation from different sizes of companies – small, medium, and large. We used an internal indicator variable for firm size to recruit across sector and size, resulting in a wide array of firms.

*Table 1: Distribution of Responding Firms for Round 1*

	<b>Total Interviews</b>
Service	9
Wholesale	6
Retail	19
Manufacturing	11
Mining/Construction*	5
<b>Total</b>	<b>50</b>
*Industries collapsed to meet disclosure requirements.	

Because we wanted to induce respondents to interact with the machine learning aspect of this research, we presented them with an initial listing of products and services specifically chosen to be outside of their industry. Early in the research, we noted that the cognitive burden of interacting with an unfamiliar list derailed the usability focus of this research. For about half of the interviews in this round, we downloaded the appropriate list of products and services and sent it via PDF in the confirmation

email before the respondent interview. During the interview, but before the respondent interacted with the online tool, we had them open the appropriate list and review it to better understand what it was that we were testing.

For the Round 1 interviewing, we tested two versions of the instrument, each with two screens. We randomly assigned respondents to be presented with either Version 1 or Version 2. Version 1, Screen 1 had respondents click an ‘Add products/services not listed’ button, and a write-up box appeared on screen. Respondents would enter their product/service into that box, and click a ‘Save and Continue’ button to proceed to Screen 2 – the search screen.

Figure 1: Round 1 Screen 1, Version 1 Screenshot

https://ext21.respond.qa.census.gov/research/app2#/

**DETAILS OF SALES, SHIPMENTS, RECEIPTS, OR REVENUE**

Select ALL that apply.

Description	Select	Product Code
<b>Manufacturing of:</b>		
1. Ship modification, conversion, overhaul, and other manufacturing services, military	<input type="checkbox"/>	2053950000
2. Other apparel contract manufacturing services	<input type="checkbox"/>	2052085000
3. Catalog and directory printing (letterpress)	<input type="checkbox"/>	2052825000
<b>All other sales, shipments, receipts, or revenue from this establishment</b>		
4. Pre-burial services for human remains	<input type="checkbox"/>	7005200000
5. Residential mortgage loans	<input type="checkbox"/>	7000300000

+ Add product/service not listed

Back Save and Continue

Version 2 Screen 1 for Round 1 presented respondents with the other/specify text capture boxes more like the 2017 Economic Census. For this version, respondents would enter their product or service in the text box and click the ‘Save and Continue’ button to advance to the second screen.



## Cleared for Public Release

Figure 2: Round 1 Screen 1, Version 2 Screenshot

**DETAILS OF SALES, SHIPMENTS, RECEIPTS, OR REVENUE**

Select ALL that apply.

Description	Select	Product Code
<b>Manufacturing of:</b>		
1. Sample books, swatches, and cards (color, carpet, upholstery, drapery, etc.), including sample mounting	<input type="checkbox"/>	2052525000
2. Hardwood rough lumber	<input type="checkbox"/>	2022225000
3. Contract manufacturing of women's, girls', and infants', team sport uniforms, costumes, and other apparel and apparel accessories, cut and sew	<input type="checkbox"/>	2052050000
<b>All other sales, shipments, receipts, or revenue from this establishment</b>		
4. Appropriations from general government revenues and intergovernmental transfers	<input type="checkbox"/>	8000225000
5. Deposit account service packages, except business	<input type="checkbox"/>	7005350000
6. Other, please specify: <input type="text" value="Describe..."/>	<input type="checkbox"/>	9783303993
7. Other, please specify: <input type="text" value="Describe..."/>	<input type="checkbox"/>	9783303994
8. Other, please specify: <input type="text" value="Describe..."/>	<input type="checkbox"/>	9783303995

[Back](#) [Save and Continue](#)

Regardless of which version of Screen 1 randomly assigned, all respondents saw the same second screen, the search results screen. This screen includes the search term from the first screen, as well as the 'best guess' of the machine learning algorithm running in the background.

Figure 3: Round 1 Screen 2 Screenshot

Add product/service not listed

You reported **restaurant**. Which of the following best describes this product/service?

Select from the results below.

If none of the results apply, please try a New Search below or select None of these and click Save and Continue to go to the next screen.

[New Search](#)

Select	Description	Product Code
<input type="checkbox"/>	Meals, snacks, other food items, and nonalcoholic beverages, prepared and served or dispensed, for immediate consumption	7000025000
<input type="checkbox"/>	Wholesale sales of restaurant, commercial food service and hotel equipment	4002560000
<input type="checkbox"/>	Construction services for new stores, restaurants, automobile service stations, parking garages, and other commercial buildings	3001250000
<input type="checkbox"/>	Construction services for maintenance and repair of stores, restaurants, automobile service stations, parking garages, and other commercial buildings	3002050000
<input type="checkbox"/>	Architectural services for apartment building and nonresidential building projects	7006775000
<input type="checkbox"/>	Store, restaurant, automobile service station, parking garage, and other commercial building constructions	3000775000
<input type="checkbox"/>	Construction services for additions, renovations, and alterations to stores, restaurants, automobile service stations, parking garages, and other commercial buildings	3001650000
<input type="checkbox"/>	Wholesale sales of commercial furniture	4002550000
<input type="checkbox"/>	Manufacturing of public bar, bowling center, cafeteria, and restaurant furniture	2018125000
<input type="checkbox"/>	Manufacturing of aluminum and other sheet metal work	2038750000
<input type="checkbox"/>	None of These	

[Add selected product\(s\)](#) [Close](#)

## Round 2 Overview

For Round 2, we completed a total of 36 interviews in November 2020. For this round, we interviewed respondents representing firms that had used the other/specify write-in capture box on the 2017 Economic Census. In targeting these firms, we were testing to see if the redesigned instrument curtailed the use of the write-in box. Note that for Round 2, we did not stratify our recruitment by firm size. For each industry except wholesale, we recruited firms to within five of the target for that industry; for wholesale, we underrecruited by about ten firms.

*Table 2: Distribution of Responding Firms for Rounds 1 and 2, and Target Distribution*

	Round 1	Round 2	Total	Target
Service	9	20	29	26
Wholesale	6	2	8	18
Retail	19	1	20	18
Manufacturing	11	7	18	18
Mining/Construction*	5	6	11	20
<b>Total</b>	<b>50</b>	<b>36</b>	<b>86</b>	<b>100</b>
*Industries collapsed to meet disclosure requirements.				

Based on the findings from Round 1 interviews, we tested one version of the tool in Round 2. Similar to Version 1, Screen 1 of Round 1, the Round 2 screen features a button to 'Add product/service not listed', but also includes additional instructions at the top of the screen as well as changed updated functionality of the buttons. The second screen – the search screen – is also similar to Round 1. In discussing the Round 2 findings, we will outline how this design came to be changed.

Figure 4: Round 2 Screen 1 Screenshot

**ITEM 22: DETAIL OF SALES, SHIPMENTS, RECEIPTS, OR REVENUE**

Of the total of Sales, Shipments, Receipts, or Revenue reported, what products and services were included?

**Select ALL that apply.**

**General** - Please do not combine product lines. If the information is not directly available from your records, reasonable estimates are acceptable.

The products and services listed below are generally made in your industry. If you make products or have revenue from sources not listed, click the "Add product/service not listed" button and search for an existing product or service.

Select	Description	Product Code
<input type="checkbox"/>	1. Manufacturing of other engineered structural wood members, excluding trusses and laminated veneer lumber (including glued laminated timber, beams, wood I-joists, I-beams)	2035500000
<input type="checkbox"/>	2. Manufacturing of laminated veneer lumber	2035475000
<input type="checkbox"/>	3. Wholesale sales of rough, dressed, and finished dimensional lumber, plywood and panels	4004725000
<input type="checkbox"/>	4. Manufacturing of wood poles, piles, and posts, treated	2035600000
<input type="checkbox"/>	5. Manufacturing of wood trusses (all types)	2035525000

[+ Add product/service not listed](#)

[Back](#) [Save And Continue](#)

## Findings and Recommendations

The remainder of this report lays out the findings and recommendations based on this testing. We present first the Round 1 findings and recommendations and then the Round 2 findings and recommendations. Because of the iterative nature of this work, the Round 2 findings will refer back to the Round 1 findings.

## Round 1 Findings and Recommendations

### Finding #1: Respondents were confused by the task.

In the Round 1 testing, we presented respondents with a randomly generated, sector agnostic list of products and services. The interviewers prompted respondents to the task at hand – adding products and services that are germane to the business but not currently on the list. Even with this prompting, respondents struggled to understand the task at hand. While some were confused about the listing, others were unsure what to do with the screen in general.

Some respondents immediately reacted to the prelisted products. For example, upon landing on the first screen, one respondent noticed the NAPCS numbers and asked “are these numbers related to

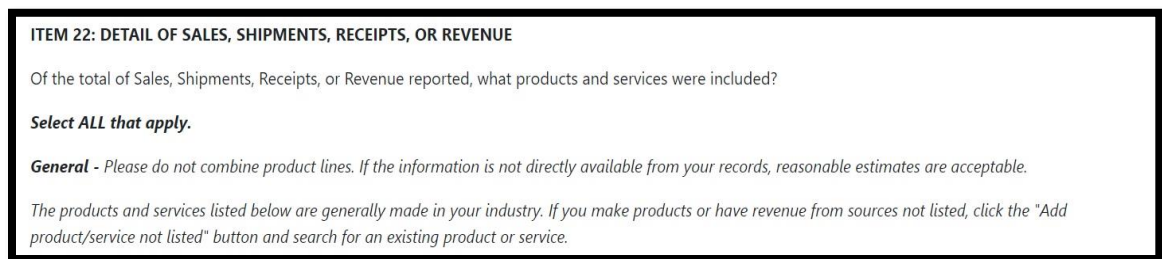
NAICS?” Another posited that the numbers are “OECD<sup>5</sup> universal business product codes”, while a third simply asked if the NAPCS values were “some kind of national coding system?”

A few other respondents were unclear on the point of the screen. Said one, this “wouldn’t apply to us” because “we don’t make the things on the list.” This particular respondent continued to struggle with the screen, eventually declaring “I am completely blank.”

**Response Error(s):** For testing, respondent confusion hindered the testing at hand, mainly, adding in missing products and services. Interviewers could guide respondents through the process, but in production, this kind of miscommunication could lead respondents to skip this question or to abandon the survey altogether, leading to higher item or unit nonresponse and less accurate estimates.

**Recommendation:** For the Round 2 testing, we included the standard instructions for Item 22 to the top of the screen.

*Figure 5: Screenshot of Round 2 Screen 1 Instructions*

A screenshot of a survey instruction box. The box has a black border and contains the following text: 'ITEM 22: DETAIL OF SALES, SHIPMENTS, RECEIPTS, OR REVENUE' in bold. Below this is the question 'Of the total of Sales, Shipments, Receipts, or Revenue reported, what products and services were included?'. Then, 'Select ALL that apply.' in bold. Next is a 'General' note in italics: 'Please do not combine product lines. If the information is not directly available from your records, reasonable estimates are acceptable.' Finally, another italicized note: 'The products and services listed below are generally made in your industry. If you make products or have revenue from sources not listed, click the "Add product/service not listed" button and search for an existing product or service.'

We also told respondents to enter their NAICS code on the first screen so that the pre-listed products and services matched the sector in which the firm is operating. This reduced confusion and made testing the search functionality more streamlined.

---

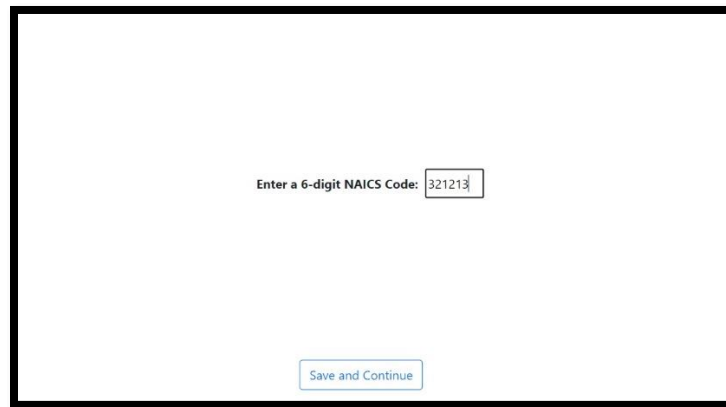
<sup>5</sup> Organisation for Economic Co-operation and Development

Findings and Recommendations from Usability Testing for the 2022 Economic Census, Item 22

02/08/2021

Page 12 of 24

Figure 6: Screenshot of Round 2 NAICS-specify Screen

A screenshot of a web form titled "Round 2 NAICS-specify Screen". The form has a white background with a black border. In the center, there is a text input field with the label "Enter a 6-digit NAICS Code:" to its left. The input field contains the text "321213". Below the input field, there is a blue button with the text "Save and Continue" in white.

**Finding #2: The search functionality is unclear to respondents.**

Before respondents interacted with the Item 22 test screen, interviewers asked them about how they anticipated the screen would function. In both version – Version 1 and Version 2 – respondents were not cued into the search functionality built into the screen.

For Version 2, which features three text capture boxes on screen 1, some respondents assumed that the box will simply record what is entered. Said one, “I think that what I enter [in the text capture box] will be entered into the system.” Another said similarly that the text capture box “would probably save as a default if I entered text here. I wouldn't think it would prepopulate anything.”

At the same time, for Version 1, where the capture box only appears once the respondent has clicked the “Add products/services not listed” button, respondents still saw the box as a verbatim text capture, not as a search box. Said one, “I expect that [clicking on the add button] will give me a blank space that I will write in what I want, what product or service I deal with” that are not already prelisted. Another specifically noted that she was looking for search functionality, saying “I would type something in [the text capture box], unless there is a way for me to search. There doesn't appear to be” a way to search from this page.

**Response Error(s):** We noticed that during the interviewing, the keywords that respondents entered into the search box were often long, complicated, and vague. The result was that the machine learning output was often not salient. Search terms that were not fruitful for the machine learning output include:

- Goods and services that help internet service providers with their content to businesses and homes;
- Equipment manufacturing;
- Food service;

- Pizza sales (dine in, carry out, delivery);
- Shelf-stable seafoods; and
- Transportation.

We expect that some of the mismatch in keywords and results has to do with respondents not knowing that they are entering keywords and not static text.

**Recommendation:** Respondents need more cues about the task at hand – searching for an appropriate good or service that may be missing from the prepopulated list. As such, we recommend better labeling and instructions to direct respondents to the search functionality. This recommendation was incorporated in Round 2 and will be discussed further in this report, as Finding 6.

### **Finding #3: Respondents are not always clear on what to do after entering a product.**

Once a respondent enters his unlisted product or service into the search capture box, we had expected he would click the ‘Save and Continue’ button to move forward to screen 2. However, we found that many respondents struggled on what the next step might be after entering an additional write-in response. A few respondents mentioned pressing the ‘Enter’ button on their keyboard, with no response from the test instrument.

However, many respondents simply did not know how to proceed. Said one after typing in an additional product, “it isn’t clear if you anticipated for additional options to pop up here” after typing in a product. Said another, “after typing, it only offers me to save and continue, or to go back,” suggesting that neither of these options was what she was looking for on screen. One respondent was forthright with his confusion, saying “I’m not sure what to do. I don’t know the product code; if I select ‘other’ does this number mean ‘other’?”, referring to the code that auto-populates once an additional product or service has been written in.

**Response Error(s):** In this case, respondents are unable to move forward in the survey without understanding how to execute the search on their additional products and services. This could lead to underreporting of products and services, introducing bias.

**Recommendation:** Throughout the survey, we have trained respondents to use the ‘Back’ and ‘Save and Continue’ buttons as navigational; these two buttons help the respondent move forward and backward through the survey. By using the ‘Save and Continue’ button to execute the search function, we have introduced an opportunity for confusion. We recommend retaining these two buttons solely for navigation, and either introducing a new button for search functionality or executing the search automatically without respondent input. This change was implemented in Round 2 testing, and we discuss it further later in this report in Finding 7.

**Finding #4: Some search results are not salient to respondents.**

Moving on to screen 2 – the search results screen – first we note that some respondents took issue with the search results. This took the form of dichotomous stances: results are too specific or results are too broad.

Some respondents note that the results are more specific than they were expecting. Said one, “the results...there’s quite a few. I’m surprised on how specific [these products] are.” Said another, the results “are more specific than what I searched on,” noting that her results list was “leading toward a specific type of manufacturing” and that she “would say none of them are overtly applicable to what we do.” Said another, “generally on these surveys, I can find a ‘big bucket’ to fit [our products and services].” He went on to note that the specificity of the results was an issue, noting that he “used the keywords that make sense, but the descriptions that are generated are somewhat limiting.” This became particularly important when looking forward in the survey – respondents did not know at the time that their answers to Item 22 would determine additional questions in the survey, but nevertheless, one respondent noted that if she selected all of the various products her firm produces, “I’m going to have to respond to different questions” about each and that, “depending on the question, I don’t know if we would have that level of detailed information, especially the production [information] as opposed to the sales [information].”

At the same time, other respondents were discouraged by the broadness of the returned products and services. The results screen prompted one respondent to say that the list doesn’t “necessarily have [a product] that’s very close” to her firm’s products and services. Echoing this sentiment, one respondent mentioned that “there’s nothing on the list that fits exactly what we do.” Another noted that the search results “aren’t related at all to [what the firm does] but they do have [one search term used] in them.” One respondent called her search results “close to being in the general ballpark” of what her firm does, going on to say that “depending on what you’re looking for, [the results] could be too broad.”

**Response Error(s):** When respondents get frustrated or are unsatisfied with their results, we see them generally take three actions:

1. They may use some other scheme to organize themselves: we had some respondents who got to the search page (screen 2) and wondered if they should enter their NAICS code or some other coding scheme to try to elicit more accurate search results.
2. They may not recognize appropriate selections: some respondents get so frustrated that the first few search results are out of scope for them that they fail to notice that the product or service that they are trying to add to their list is in fact within the search results; these respondents may select ‘none of these,’ which does not reduce the number of write-in responses.
3. They may abandon the survey altogether: again, when a respondent is overly cognitively challenged on a survey, he may simply terminate his survey session. Respondents who terminate a session – ‘partial completers’ – can be unlikely to log back in and complete the survey. In an analysis of paradata from the 2012 Economic Census, Hoeffel et al.

noted that more than three in four respondents (78.3 percent) logged into their survey only once, “implying that if a user submitted forms, the forms were completed in one sitting”<sup>6</sup>. We do not want to incur break-offs at this point because our machine learning algorithm is insufficient.

None of these three strategies will result in reduced write-in responses and increased accuracy of the resultant data.

**Recommendation:** We recommend that the machine learning team continue to refine the training data sets and resultant algorithms to ensure as salient results as possible.

**Finding #5: Respondents are confused by the looping search results for multiple entries.**

Version 2 of Item 22 allows for up to three write-in responses for respondents to specify additional products and services. Respondents are then ‘looped through’ each of the write-in responses they had provided: each search term produces a separate results list. Part of the confusion stems from the way that the instrument presents results – we use a ‘one to many’ match, wherein the algorithm produces best guess matching products and allows respondents to select as many as apply. However, regardless of what a respondent chooses on the first screen of results, she will still be next directed to search results for the subsequent search terms. This led respondents to be confused about what was populated on their screen.

**Response Error(s):** When a survey functions in a way that confuses respondents, they are less likely to complete the survey. See Finding 4 for the response errors embedded in ‘break-offs’.

**Recommendation:** In this case, we recommend having respondents search for each additional product and service separately. In this way, respondents can keep track of which term they are looking for, which terms they have already captured, and when they have a complete picture of the products and services offered by their firm. We recommend once a respondent has added additional products, looping them back to Screen 1 as a final ‘check’ at the full list before moving on to other survey questions. This is further discussed in Finding 7 from Round 2.

---

<sup>6</sup> Hoeffel, Elizabeth M., Joanna Fane Lineback, and Gianna Dusch. 2013. “What Can Web Paradata Tell Us About 2012 Economic Census Business Reporting?” Research Report: Federal Committee on Statistical Methodology: 9. Available: [https://nces.ed.gov/FCSM/pdf/D3\\_Hoeffel\\_2013FCSM.pdf](https://nces.ed.gov/FCSM/pdf/D3_Hoeffel_2013FCSM.pdf)



## Round 2 Findings and Recommendations

### Finding #6: Search functionality cues help orient respondents.

Based on the findings from Round 1, the Round 2 test instrument included additional cues about the search functionality. We changed two features to the Item 22 screen to direct respondent attention to this functionality, including:

- Removing the capture box from the default question block: Based on interactions with the looping search functionality as well as the drive to reduce the number of write-ins, we ultimately abandoned the text capture box from the screen altogether. Instead, respondents are presented with the 'Add products/services not listed' button, which upon clicking, activates the search screen.
- Additional instruction in the header that explicitly tells respondents that clicking on the button will let them search for an existing product or service.

Interviewers noted that fewer respondents needed guidance that Item 22 had search functionality built in while conducting Round 2 interviews in comparison to Round 1 interviews.

**Response Error(s):** Cueing respondents into the task at hand – searching for products and services – may lead to better search results produced by the machine learning algorithm, which may, in turn, lead to fewer write-in responses overall. These search cues seem to reduce response error, when response error is defined as supplying a written-in response.

**Recommendation:** Retain the search functionality cues included on the Round 2 instrument.

### Finding #7: Separating navigation buttons from a search button helps orient respondents.

In response to Finding 5 from Round 1, Round 2 tested search functionality that activated automatically once a respondent clicked on the 'Add products/services' button, and retained the 'Save and continue' and 'Back' buttons for navigation through the survey only. We found that respondents identified the purpose of the button more readily, and still understood the navigation buttons to move to and from questions. Asked how he may include unlisted products, one respondent said bluntly, "I would click 'Add product/service not listed'." Another expanded, saying "I would have to add [a service to the list] – I offer dining, hospitality, entertainment. I would click on the button and search."

**Response Error(s):** The recognition that the add button activates a search may lead to better search terms, and increased salience in the search results. This could decrease response error and write-in responses.

**Recommendation:** Retain navigation buttons for navigation, and use the add button to call up the search window.

### Finding #8: Respondents struggled with the 'Add selected products/services' button.

Once a respondent clicks on the 'Add products/services not listed' button, he is immediately directed to a search page. This search page features instructions, a search bar, a search button, an add button, and a close button.

Figure 7: Screenshot of Round 2 Search Screen

**Add product/service not listed** [Close]

Enter search terms for additional products or services in the search bar below. Then select the result(s) that best match(es).

If none of the results apply, please try a new Search, or select "None of these" and click the "Add Selected Products/Services" button.

If you have come to this screen in error, please click the "Close" button.

**Specify additional product or service:**

lv [X] [Search]

[Add Selected Products/Services] [Close]

	laminated timber, beams, wood I-joists, I-beams)	
<input type="checkbox"/>	2. Manufacturing of laminated veneer lumber	2035475000
<input type="checkbox"/>	3. Wholesale sales of rough, dressed, and finished dimensional lumber, plywood and panels	4004725000
<input type="checkbox"/>	4. Manufacturing of wood poles, piles, and posts, treated	2035600000
<input type="checkbox"/>	5. Manufacturing of wood trusses (all types)	2035525000

[+ Add product/service not listed]

[Back] [Save And Continue]

While the automatic execution of the search functionality ultimately helped respondents to navigate the survey, we found that providing the 'Add selected products/services' button immediately caused some respondent confusion. Several respondents entered their search term in the search bar, but instead of clicking the 'Search' button, they instead clicked the 'Add' button. Because no search was executed, and so, no product was selected, clicking on 'Add' here simply took the respondent back to the first page without recording what they had entered. For example, one respondent narrated his interaction with the screen: "I would click add product or service. Then, I am typing into the search bar [a description of our business]. Then, I would click the add product or service button." As designed, the instrument took him back to the first screen, and he then noted "I don't see what I wrote on the list."

At the same time, a number of respondents were able to execute the search using the 'Search' button, scroll through the results and make selection, but then instead of clicking on the 'Add' button, they clicked on the 'Close' button. Since no products were 'added' from the selection, when the respondent is redirected to the first screen, the selected products are not populated in the list.

**Response Error(s):** The 'Add' button is being used inappropriately by respondents, leading to missing data.

**Recommendation:** The first recommendation to increase the usability of the buttons on the search screen is to make them available only as they are needed. Respondents can only add an additional product once they have selected one from the search results. Therefore, the 'Add' button does not need to be available to respondents until that step has been completed. This button should be grayed out or withheld from screen altogether until a respondent has checked a product or service.

The second recommendation in response to this finding is to test the location of the 'Add' button relative to the 'Close' button. That is, one future research question might involve moving these two buttons closer together or further apart on the screen to induce appropriate interactions. We may also consider an error message or warning to alert respondents that they have not selected any products or services before closing.

**Finding #9: The layout of the search results screen limits respondents' answers.**

Keeping with the search results screen, we also noticed that some respondents did not see that the list of results had a scroll bar. That is, respondents saw the top results, and did not realize that there were additional results available to them if they just scrolled down. For example, one respondent narrating her interaction with the search results screen said: "It's not any of these [from the results list] - none apply because [my business's service] simply is not here... wait... [I just] scrolled down and saw more than 4 [results]. I didn't notice at first that there were more on the list to scroll down to."

**Response Error(s):** Because respondents are not seeing the breadth of the search results, they may be selecting suboptimal choices, using the 'None of these' selection, or leaving products and services from their listing altogether.

**Recommendation:** We recommend additional research into the placement of the buttons and the instructions. For example, if we move the next action buttons – the 'Add' and the 'Close' buttons – to the bottom of the screen, only accessible once a respondent has scrolled through the results, he may be more likely to review the list of results. At the same time, we recommend 'floating' the instructions on the search screen, such that as a respondent scrolls, he always sees the instructions at the top of his screen.

## About the Data Collection Methodology and Research (DCMR) Branch

The Data Collection Methodology and Research (DCMR) Branch in the Economic Statistical Methods Division assists economic survey program areas and other governmental agencies with research associated with the behavioral aspects of survey response and data collection. The mission of DCMR is to improve data quality in surveys while reducing survey nonresponse and respondent burden. This mission is achieved by:

- Conducting expert reviews, cognitive pretesting, site visits and usability testing, along with post-collection evaluation methods, to assess the effectiveness and efficiency of the data collection instruments and associated materials;
- Conducting early stage scoping interviews to assist with the development of survey content (concepts, specifications, question wording and instructions, etc.) by getting early feedback on it from respondents;
- Assisting program areas with the development and use of nonresponse reduction methods and contact strategies;
- And conducting empirical research to help better understand behavioral aspects of survey response, with the aim of identifying areas for further improvement as well as evaluating the effectiveness of qualitative research.

For more information on how DCMR can assist your economic survey program area or agency, please visit the [DCMR intranet site](#) or contact the branch chief, [Amy Anderson Riemer](#).

## Appendix A: Interview Protocol

### NAPCS Debriefings Protocol

#### General Probes (Use these whenever necessary)

- In your own words, what would you say is the purpose of this survey question?
- In your opinion, is the information/ instructions on this screen helpful or not helpful?
- Can you tell me what you are thinking about?
- Can you tell me more about that?
- I want to make sure I understand; can you explain that again?
- Reflect back on R's answer ("I want to make sure I have it right. I think you said, "...?")

#### 1. Participant Background

- How long have you been with the company/in your current position?
- What is your title/role?
- What are your major responsibilities?
- What are your government reporting responsibilities?
- What other government surveys or filings, if any, do you also handle?

#### 2. NAPCS Background

- Can you tell me about your process for filling out Economic Census?
  - *Additional probes (if necessary):*
    - Was this your first time filling out this survey?
    - What was your role completing the Economic Census? Are others involved in helping gather the data? If so, Did you complete Item 22 which asks about the *Details of Sales, Shipments, Receipts or Revenue* in the 2017 Economic Census survey? –may reference content that was emailed if needed
    - How do you typically gather the data for this survey item?
    - Did someone else have to review the data to this survey item before you submitted it?

*Now I would like to ask you specific questions about this survey item. Please open the attachment/ click on the link that I sent you via e-mail (if haven't already done so).*

#### 3. Item 22 (Part A)

- In your own words, what was the purpose of this survey item?
- Describe to me how you reported to this item?

## Cleared for Public Release

- Via the spreadsheet or form (screen by screen) view?
  - Tell me about your past experience reporting to this survey item? Any recommendations for improvements?
- Did you use any records or other documentation?
  - If so, describe the nature and content of these records.
- Did you obtain assistance from others?
- How easy or difficult was it to report to this survey item? Describe.
  - Any thoughts on the pre listed options that were presented to you?
    - Number of items
    - Ordering
    - Display
  - How easy or difficult was it to find the services/products that you needed to report?
    - Please tell me about how you went about identifying your products/services among the items that were listed.
- Were the product/services listed relevant/applicable to your company? Did they accurately reflect your products/services?
- Were you able to report all of your products/services from the listing? Were there any missing?
  - If so, how did you account for the products and services that were not pre listed?
- To what extent did the descriptions used for your pre listed products/services match the way your business describes them in your records?
  - How easy or difficult was it for you to identify the products/services from your records on the listing
- **Machine learning**
  - Option A: What do you think is the purpose of line item #X?
  - Option B: What do you think is the purpose of the *Add additional products/services* button at the bottom of the screen? What do you expect to happen upon clicking on this button?
  - For which option would you prefer to use to write in additional products/services that you did not find on the listing.
  - Any thoughts on having the ability to type in products/services not listed?
    - Describe this process for me.
  - What type of words do you think someone could enter?
  - Phase 1: *Describe the search functionality for write in responses*: Lets say someone entered X into the field because they did not see it on the pre listed products/services. After typing in those keywords and clicking on the *Save and Continue* button to proceed, the survey presents them with the following screen. Is it clear what happened here? Describe.

- How should one proceed from this screen. If the results were not applicable, how should one proceed?
- Any thoughts on the display?
  - Number of results displayed
  - Accessing additional results
- Phase 2: *If respondent had an Other write in response in prior reporting:* Describe to me what you may enter based on your establishment.
- Lets enter those words from this screen. After typing in those keywords and clicking on the *Save and Continue* button to proceed, the survey presents you with the following information. Is it clear what happened here? Describe.
- How should one proceed from this screen. If the results were not applicable, how should one proceed?
  - What if you entered the wrong words resulting in results that were not applicable?
- Any thoughts on the display?
  - Number of results displayed
  - Accessing additional results
  -

#### 4. Item 22 (Part B)

- Now I would like for you to navigate to *Item 22 Details of Sales, Shipments, Receipts or Revenue*
  - In your own words, what was the purpose of this screen?
    - Anything missing (e.g., expectations)
  - Any thoughts on this screen? What is it showing you/asking you to do here?
  - How do you feel about being shown only products/services that you selected on that previous screen?
- Please describe how you arrived at your answers to these items? That is, how did you go about determining the revenue/receipts/sales values requested here?
  - Did you use any records?
    - If so, describe the nature and contents of these records.
    - How are these records maintained?
  - How easy or difficult was it to report to this survey item? Describe.
  - Are you able to report revenue/receipts/sales values for all of your products/services selected from the previous listing?
    - *If no*, tell me more about that.

#### Wrap-up

**Cleared for Public Release**

- Do you have any final recommendations or comments on any of the screens we went over today?

***This concludes our meeting. Thank you for your time and valuable feedback.***