Considering a Digital 2020 Census

Michael T. Thieme

U.S. Census Bureau
4600 Silver Hill Rd.
Washington, DC 20233 - 7200

Introduction

Every Decennial Census builds on the successes and works to address the challenges of the previous one. The scale of Decennial Census operations remains unparalleled by any other peacetime effort undertaken by the United States, but the goal remains the same: the Census Bureau must enable the location and enumeration of all US residents within a six month time period while meeting stringent demands for accuracy. As we look forward to the 2020 Census, societal issues such as immigration, the complexity of the modern family, and the perceived role of government will impact our ability to perform an accurate count. At the same time we are faced with an unprecedented opportunity to leverage the American society's increasing willingness to incorporate technology into their lives in ways we could not have anticipated only a few years ago. With these things in mind, a question arises – should we continue to build on the past or should we try a different approach?

Although the 2010 Census originally planned to capture data both on paper (through mail out/mail back operations) and digitally (through the Internet and handheld devices), in the end, we captured virtually all respondent data using paper. The initial plans for the 2020 Census propose a number of different combinations of technologies for data collection, including paper, Internet, and administrative records, with the intent to give respondents maximum flexibility and encourage self-response. While this is an admirable goal, the increase in available modes of response also comes with increased complexity, as each of the options will require a greater level of integration between groups and operations.

Given the observed evolution of technology and its impact on the Census, it seems more important than ever for us to analyze the true costs and benefits that greater complexity will bring to the next Census. Though a complex process is not inherently "good" or "bad", overly complex processes can unnecessarily increase costs, lower agility, cause unnecessary confusion among respondents, and present a greater level of risk. We are faced with an opportunity here at the beginning of 2020 planning to make a decisive move in the direction of simplicity. Though such a move will bring us out of our familiar comfort zone, it is the position of this paper that the Census Bureau should consider performing a virtually "All Digital" response Census and do away with the use of paper for data response in 2020.

Has the Decennial Census Gotten too Complex?

How We Arrived At the Current Census

How can something that sounds so simple: "count everyone once, only once, and in the right place," be so complex? The answer lies in the gradually increasing number and complexity of the specific operations we have designed over the years to address the individual needs of our many stakeholders. Since the initial execution of the Census in 1790, the degree of complexity incorporated to achieve the Census mission has continued to increase along with the expanding and changing population of the United States. Likewise, the 2020 Census is expected to continue a trend

of increasing cost in its attempt to deal with the complexities involved with "increased population diversity and decreased willingness to cooperate with self-response and non-response followup".[1]

In order to account for the new population challenges that 2020 holds, the Bureau is in the midst of evaluating alternative modes for data collection to reduce costs while maintaining quality and coverage. Though prior to the 2000 Census we suggested that the use of statistical sampling for non-response could significantly reduce the cost and operational complexity of that Census, the courts at that time ruled that statistical sampling is not a permissible means of obtaining Census data for the purposes of apportionment. In a similar attempt to reduce cost, initial 2020 Census design plans include research on the increased use of administrative records to supplement Census data collection. This is currently under review by the Bureau.

The True Complexity of a Paper-Based Response – Not as Simple as it Seems

As an organization, the Bureau is well accustomed to paper-based responses. Through the years, the limitations of paper have become so familiar as to seem normal and employees are very comfortable with its tangible nature.

But what is the true operational burden of processing a paper form? At great cost, hundreds of millions of paper forms have to be printed, stored, mailed out, re-mailed out, returned to processing centers, opened, optically converted to images, "read" by software, validated or keyed in by human beings, archived and eventually destroyed – all in a fashion that preserves privacy and security of sensitive information. The cost of the printing and scanning forms alone comes to approximately two billion dollars, showing us that the true cost of printing, sending and processing a paper response is somewhere between $12 and $14 (see Appendix 2).

Industry Comes to Realize the True Cost of Paper

Industry has realized that the true cost and implications of the use of paper is often buried in operational costs and inefficiencies that are not immediately apparent. As a result, organizations are going paperless and realizing significant benefits – from the financial services industry to hospitals storing life-critical patient information in digital files. Within the next decade, the trend away from paper will continue, with many businesses eliminating paper or making paper correspondence a paid option.

The use of paper also has significant effects on other Decennial operations. For instance, since paper can't be processed in real-time and has to be converted into a digital format before it is useful for processing activities, many workarounds have been developed. For example, the Bureau worked with the USPS to get real-time information from distribution centers on forms en route to the Bureau to overcome the time lag that the mail system generates. The USPS information was used to minimize the mailing of duplicate forms to households that had already returned a completed questionnaire.

A little bit of paper won't hurt…will it?

It has been proposed that both a paper and digital response option be made available in 2020 in order to give respondents a greater degree of flexibility while capitalizing on the benefits of digital response. Though this solution may appear to be a fair compromise to reduce complexity by decreasing paper responses, our analysis suggests that it is quite the opposite: offering both has a higher level of complexity than offering digital or paper alone.[2] The reality is that any introduction of paper, even with the intention of using it on a small scale, will require Census to plan and size contracts for the largest volume response possible for each option. As a result, offering both a paper and digital response will be both resource and cost intensive, and will create redundant processes and operations. On the other hand, a desirable and acceptable use of paper would be to make the initial contact through the mail and provide a unique code for respondents to use for self-response over the Internet, as Canada has done in their Census.

Defining Complexity

In the context of the Census, complexity can generally be defined as the degree to which people, processes, and systems must be coordinated, taking into account the number of touchpoints and handoffs required to successfully

[1] Daniel H. Weinberg, Plan for the Research and Testing Phase of the 2020 Census, (International Census Forum, 2010).

[2] Sylvia Fisher, Jean Fox, William Mockovak and Christine Rho, Usability Issues Associated with Converting Establishment Surveys to Web-Based Data Collection, (Bureau of Labor Statistics, 2003) 1-2.

execute all operations. In the Census, as with other large operations, some process areas embody sub-functions that increase complexity; for instance, the need to integrate multiple technology systems. For the purposes of this paper we have identified three areas that will be discussed at greater length later in the document: management complexity, information systems complexity, and organizational complexity.

It is important to note that complexity does not increase in a linear fashion; any increase on a single dimension impacts the complexity of all dimensions.[3] For example, as the complexity of information systems security increases, there are secondary and tertiary impacts on the capture, systems integration, and responsiveness dimensions. Adding additional security measures to one system can make integrating data in the overall control system more complex, require changes to interfaces and ultimately impact the delay of receiving time-sensitive data.

The Hidden Cost of Complexity

Stephen A. Wilson, who has conducted extensive research on complexity costs, states "complexity costs are those costs in a business associated with having too many parts, products, services, systems, processes, business lines, plants, stores, vendors, customers, organization functions, relationships…"[4] Census must carefully consider if offering two fundamentally different technologies, paper and digital, merit the additional complexity costs that they bring. The decision to use paper, digital, or both to collect respondent data pervades many other processes that may not be immediately clear. For instance, the exclusive use of digital technology will likely decrease the complexity of NRFU, due to the enhanced ability to determine the non-response universe in real-time, saving unnecessary enumerator visits.

Research on the topic leads us to believe complexity costs can become significant very fast, and often times without notice.[5] While the first Census, in 1790, was relatively simple with enumeration required for each individual household, the mission has since ballooned in scope and scale requiring new operations in order to serve the variety of constituencies of Census data. For instance, Group Quarters (GQ) Enumeration is responsible for counting populations such as college dormitories, nursing homes, military barracks, prisons, juvenile institutions, migrant worker dormitories, convents, and group homes. There are 27 types of GQs, each requiring special processes in order to ensure complete coverage. For example, jails and prisons employ various enumeration processes where corrections staff become sworn enumerators. Further, the precise enumeration procedures vary site to site due to the unique demands of each facility. The need to create dedicated processes to count certain population groups significantly impacts complexity costs. It is reasonable to believe that additional operations will be needed to reach special population groups as the US becomes more diverse.

As the Bureau prepares for the 2020 Census, increased attention should be given to how the various population groups can be counted more efficiently though the use of different processes (i.e. using administrative records for prisons instead of counting individuals who are already closely monitored in IT systems) and new technologies.

What's on the Table for 2020

Modes are the fundamental tool used to facilitate data collection from the population and have far reaching implications. We have compiled the potential modes for the 2020 Census based on the operational design alternatives under consideration, as well as those modes that were used during the 2010 Census and those that could be used for a "Digital" Census.

It is almost without question that an Internet response option must be offered in the 2020 Census, and likely that mobile devices will be used by Census enumerators to record interview results. Though the use of an Internet response

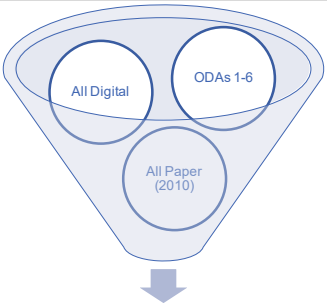

Figure 1: 2020 Options

[3] Stephen A Wilson and Andrei Perumal, Waging War on Complexity Costs, (New York: McGraw Hill, 2010).
[4] Wilson 70
[5] Wilson 36

option has already been established, the precise role of a digital response technology combined with the use of paper are still unknown and under consideration for 2020.

Considering the role of paper and digital in 2020 is important because they vary greatly in terms of how respondents interact with them and how the Bureau collects and integrates data from each source. The limited timeframe to complete the Census requires moving millions of people, surveys and systems across the country in a systematic and timely fashion. All moving pieces must be tightly secured and traceable. In the 2010 Census, millions of paper forms were mailed to housing units and then returned via mail for processing at various sites. In the case of the non-response operation, the LCO must get the right enumeration materials to the appropriate enumerators and then ship the completed enumerator questionnaires to Data Capture Integration (DCI) for capture.

Using Administrative Records…

The increased use of administrative records is under consideration for the 2020 Census to supplement data collected directly from respondents. Administrative records can be used in two ways: for characteristic imputation (supplement/verify data already collected) and for count imputation (to establish a count). The use of administrative records in either role will impact the scale of collection and/or validation operations, but does not alter the fundamental costs and resulting complexity of the response options selected for 2020.

Can We Make Do Without Paper?

Paper has played a significant role in recording Census data from respondents since the start of the 1790s. Indeed, the use of paper to make the initial contact with respondents and to provide a code that allows for response over the Internet will be a given in 2020. Though a Digital response survey appears to be more technologically advanced than an All Paper response survey, many fail to consider the immense amount of technology required to ultimately capture responses for inclusion in the Census. The technology needed to scan, capture, and digitize forms is quite advanced and complex, bringing significant costs and consuming valuable time needed to complete the process.

The Decennial program considered the use of an Internet response option in 2010. Though the Bureau determined that an Internet response Census was feasible, the Bureau ultimately chose not to proceed with it citing various challenges at the time. The advancements in technology and cultural changes that will take place by 2020 will likely ease the introduction of a digital response option.[6]

Studying Survey Modes

In order to gain a better understanding of how survey modes affect the Decennial program and influence operations, we decided to further examine the composition of what a mode is. Based on our research and analysis, we determined there are two basic dimensions that hold unique characteristics and are independent from one another: the response technology and the collection channel.

What is a Response Technology?

A response technology is the actual interface, whether paper or digital, used to capture data about the respondent. When analyzing response technologies we are primarily concerned with how easy it is to capture the data and then integrate it in to the greater Census universe, while the collection channel looks at the logistics, management, and operations behind getting the technology to the respondent.

[6] Digital Nation: Expanding Internet Usage, (U.S. Department of Commerce National Telecommunications and Information Administration, 2011) 1-4.

What is a Collection Channel?

Though many survey modes use the same underlying response technology, there may be variance in the delivery mechanism, or collection channel, involved. For instance, the use of paper to capture responses can be delivered via mail for self response, or be completed by a staff enumerator on behalf of respondents. The collection channel addresses the means in which respondent data is communicated to the Census. Though paper may be the response technology used to record the response, USPS is the collection channel to actually return the data to Census. In the case of digital, the Internet would be the collection channel and an Internet website is the response technology.

The table below depicts the breakdown of modes into their two comprising parts. Each of these dimensions share common features in terms of characteristics and complexity.

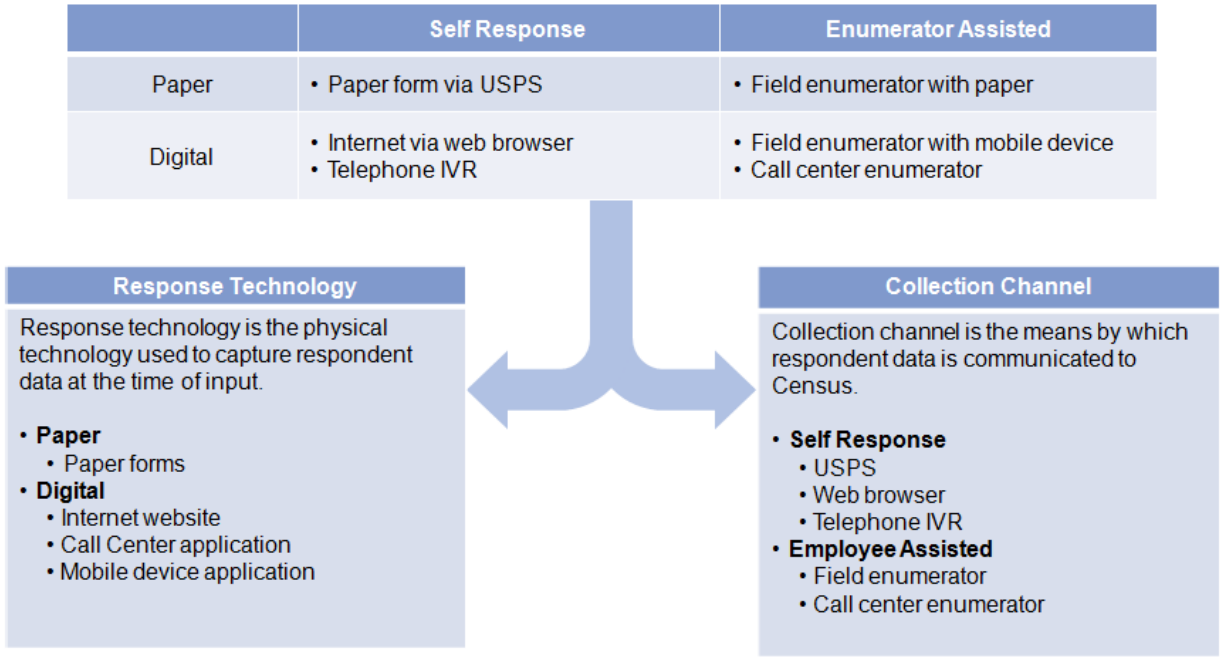

Figure 2: Survey Modes Combine a Collection Channel and a Response Technology

While we defined that a survey mode consists of a collection channel and a response technology, the response technology is our principle focus for this analysis.

A Closer Look at Analyzing Modes

In order to gain a better understanding of the implications of employing multiple survey modes to collect data, we must look closer at the aspects that make modes complex. Modal complexity can be assessed by evaluating the response technology and collection channel independently in order to ascertain, and delineate, components of complexity and how they impact the greater process.

The collection channel and response technology have very different complexity components. While collection channel complexity focuses on logistics and management, response technology is more concerned with how the data is captured and integrated. In order to better focus our discussion around a Digital Census, we are going to concentrate on the complexity of different response technologies. Though channel complexity is an important consideration, it becomes more prevalent when evaluating actual modes rather than generalized response options.

In order to perform an analysis of the response technology alternatives for 2020, we identified critical sub-components specific to the Decennial program that are likely to have a significant impact on an option's overall complexity. The following table outlines principle areas of complexity that we are going to explore in order to analyze variances in options for 2020.

Response Technology Areas of Complexity	
Management Complexity	
Data Availability	Complexity associated with the delays in managing operations if data is not available. Quicker data availability enables more timely decision making and better resources allocation while preventing waste (for instance, preventing an enumerator from visiting a household that submitted a questionnaire late).
Data Management	Complexity associated with management of the data for the Decennial and effort required to manage data privacy and security, to prevent data loss, to improve data accuracy, to reduce fraud and to improve data reliability.
Respondent Support	Complexity associated with the need to provide assistance to respondents to support them in successfully complete a questionnaire accurately, including the need to provide access to a proper questionnaire (such as a form in a different language or a duplicate form).
Information Systems Complexity	
Response Capture	Complexity associated with capturing responses from questionnaires for integration in to the Census universe, whether provided via self-response, an enumerator, or other means.
Systems Integration	Complexity associated with the number and types of information systems that must be integrated in order to successfully capture, store and manage respondent data.
Organizational Complexity	
Organizational Scale	Complexity associated with managing a number of disparate parties (internal or external) required to support the response technology.
Organizational Experience	Complexity associated with managing new, updated or changed response technology(ies) and supporting systems at an organizational level.

Table 1: Response Technology Evaluation Criteria

Evaluating Options for the 2020 Decennial

Three High Level Response Alternatives for 2020

In order to simplify the analysis we have created three high level response technology alternatives: Paper Only, Mixed Paper and Digital, and Digital Only. The Mixed Paper and Digital alternative represents the current alternative being considered for the 2020 Census design. It is important to note that this discussion focuses on how Census responses are returned to the Census Bureau. Paper will likely be used in all scenarios in order to inform respondents of the upcoming Census and provide response instructions. Table 2 illustrates the breakdown of response technologies used under the three high level design alternatives. We are also able to see the technology used for self response and enumeration.

	Self Response	Enumerator Assisted
Paper Only	Paper	Paper
Mixed Paper and Digital	Paper/Digital	Digital
Digital Only	Digital	Digital

Table 2: Technology & Response Channel Matrix

Analyzing the Alternatives

In order to evaluate the relative complexity of our three groups, we will apply the criteria to an All Paper, a Mixed Paper and Digital Census, and a Digital Census. The following radar charts illustrate the relative complexity of these three options. The scale being used is low, medium and high complexity. The further the points and lines extend outward in each table, the greater the complexity of that given response technology. The rating rationales can be found in Appendix 1.

Evaluation of an All Paper Response Survey

At a first look, "All Paper" seems to be low-tech and simple: a form is mailed out, the respondent completes it, and then mails it back. The complexity of paper is largely a result of the data availability, response capture, and data management complexity resulting from the delay incurred when the respondent mails back the Census form, until it is processed and captured. Paper requires a high level of support due to the systems that are needed to capture and store data that is communicated on the Census form. Figure 2 depicts the complexity of an All Paper Census. We can clearly see that the All Paper Census has a low degree of organizational experience complexity as a result of the Bureau's extensive experience using paper and capture technologies.

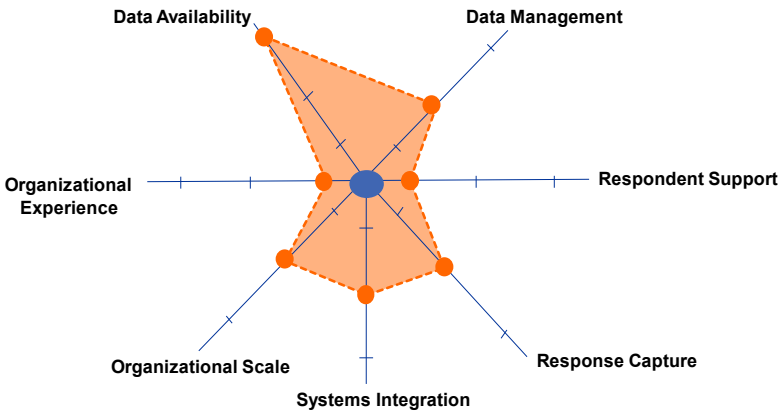

Figure 2: All Paper

One of the key challenges of paper is the lag of data due to the time required to digitize the paper forms. As a result, management does not always have the most accurate, up-to-date data to make timely decisions. At times, this can cause inefficiencies, especially when non-response follow ups are deployed while forms were in route by mail. Though paper provides the greatest accessibility for self-response due to its portability, the back-end complexities and processing challenges may in-fact pose greater risks and threats than a Digital response option.

Evaluation of a Mixed Paper & Digital Response Survey

A Census that employs paper and digital response technologies will naturally be more complex than a Census that only uses one response technology. The level of complexity increases exponentially with the addition of a second response technology due to the need to manage a separate set of systems, but also due to the higher level of coordination and integration required between capture systems.

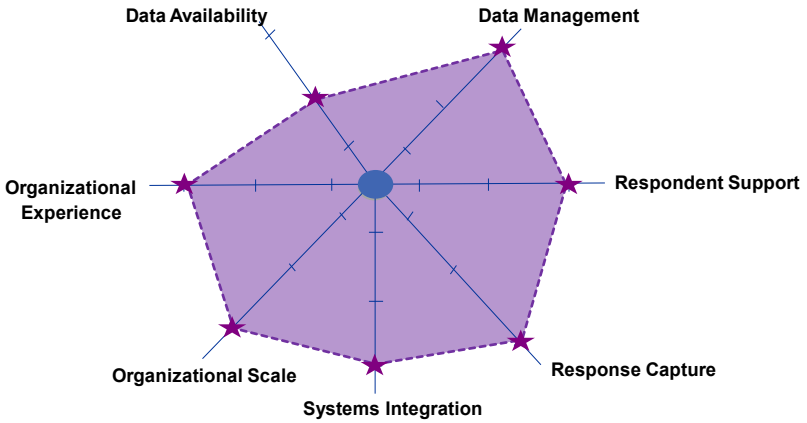

Figure 3: Mixed Paper and Digital

Figure 3 illustrates the complexity of the Mixed Paper and Digital option. All but one area of complexity is scored as "high," primarily because the limitations of paper hold true in addition to the increased complexity of having to manage and integrate two key disparate systems.

The perception is that more response technology options will lead to a higher self-response rate and better quality data by making it easier to "do business" with the Census Bureau. The reality is that it can likely cause confusion among respondents and studies have demonstrated that offering multiple response technologies does not necessarily increase response rate.[7] Research also indicates that operating a multimode survey presents data consistency challenges due to "mode change," where the respondent's answer may vary depending on the response channel used, even when questions are worded exactly the same.[8] Similarly, a study conducted by the ACS in 2001 found that "…offering multiple modes of response in a mailing led to a drop in overall response."[9]

Evaluation of a Digital Response Survey

Based on our evaluation criteria, a Digital response Census bears the least complexity. This is primarily due to the elimination of the paper capture operations and the better data integration that results from collecting data in a digital format from all sources.

Advancements in technology have brought digital media to the forefront of every life. Americans are becoming increasingly comfortable with the Internet and it is reasonable to expect for the Internet to become even more prevalent than it is today.

One of the key advantages to a digital response is that a digital questionnaire can enable alternate presentations of the form (such as larger fonts for Americans with Disabilities Act (ADA) compliance and alternate language support) that may ease the completion of the form and ultimately lead to a higher level of

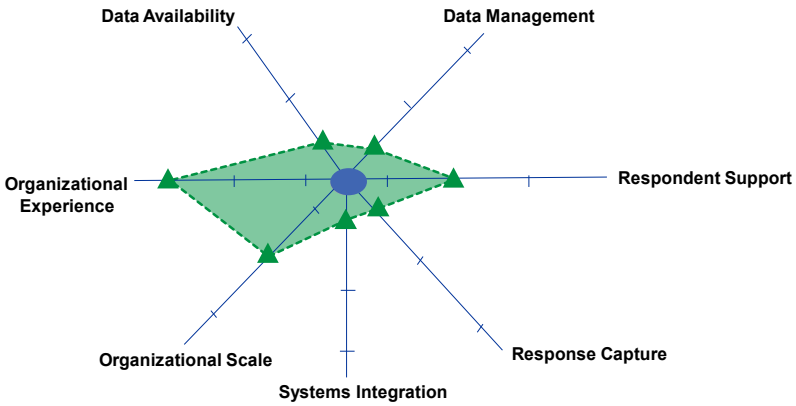

Figure 4: Digital

[7] Edith D. De Leeuw, To Mix or Not to Mix Data Collection Modes in Surveys, (Journal of Official Statistics, 2005) 240-41.
[8] D.A. Dillman and L.M. Christian, Survey Mode as a Source of Instability in Responses across Surveys, (2003) 12.
[9] Deborah H. Griffin, Donald P. Fischer, and Michael T. Morgan, Testing an Internet Response Option for the American Community Survey, (Washington: U.S. Bureau of the Census, 2001) 16.

respondent-provided data accuracy.[10] For instance, a "wizard" format may be

used where the user is guided through a series of screens depending on prior responses. "Screen tips" can enable the respondent to receive help on a specific form field right on the form.

Figure 4 represents the Digital option. The overall complexity is relatively lower than All Paper in all areas except for systems security, risk management and customer support. The Digital option is highly complex in customer support because many demographics will require significant training and customer support.

Other Key Factors for the Bureau

While complexity is a key factor that we have applied to 2020 Decennial design options, other key factors include cost, quality and coverage. Each of these factors are broad in scope and are topics for future analysis. The technology or combination of technologies that is used will have sweeping impacts on the Decennial. Some examples of the questions that need to be explored are: 1) How will use of the Internet impact coverage?; 2) Can Internet self-response and enumerator with mobile device reach the same, or greater coverage, than paper based collection?; 3) Will upfront investments that are required for building out digital infrastructure drive down long-term costs of executing the 2020 Census and future Decennials?; and 4) Does digital or paper lead to a higher quality of data?

Summary of Findings

Figure 5 illustrates the relative complexity of each high level option. As described above, All Paper and Mix of Paper and Digital are relatively more complex than the Digital alternative. Our analysis of the Mixed Paper and Digial illustrates that combining technologies causes a significant increase in complexity due to the replication of systems and processes nessesary to drive the response option. Simply put, the scale of a given response option has little effect on the resulting complexity.

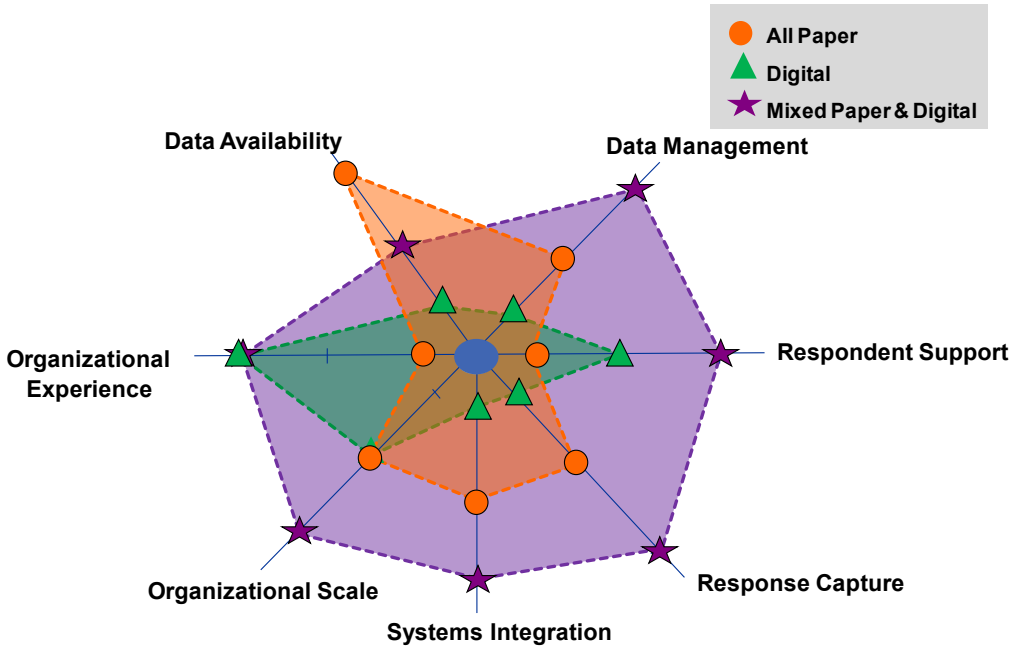

Figure 5: Summary of Findings

[10] Daniel Castro, e-Census Unplugged: Why Americans Should Be Able to Complete the Census Online, (Washington: The Information Technology & Innovation Foundation, 2008) 5.

There are certain pros and cons inherent to the All Paper, Mixed Paper and Digital, and Digital alternatives that we touched on above. The following table analyzes select pros and cons in areas of the likely response rate, data capture and data integration. Those pros and cons are highlighted in Table 3.

	Pros	Cons
Paper Only Response	• Tangible • User friendly (at least to subsets of the population) • Broader reach • Perceived as more secure than digital • Less change management required within the Bureau – "tried and tested" approach	• Logistics involved with printing, mailing and scanning paper forms can be time consuming, expensive and resource intensive • Data visibility is not in real-time • Need to print and support multiple languages • Cost of converting results to digital format
Mixed Paper and Digital Response	• Provides greater flexibility to respondents • Built-in redundancy	• Multiple response channels has the potential to cause confusion among respondents and increase integration complexity • Coordinating across multiple response channels can dramatically increase complexity • More susceptible to duplicate responses • Paper mode capacity, cost and complexity needs to be scaled for maximum possible paper response rate • Paper to digital conversion would still be required
Digital Only Response	• Logistics involved with printing, mailing, managing, archiving and scanning paper forms get eliminated • No need to convert from paper to digital • Integration is seamless and near real-time, providing improved ability to respond to changes • Reduces human error through standardization • Provides respondents with instant access to alternate language and customized help (screen tips)	• Not physically tangible like paper • Less reach than paper – potential to reduce self-response rate as older populations and low income households may either not have access or could be not digitally savvy • Perceived risk of fraud is higher • Increased customer service/helpdesk capabilities required at the Bureau • Requires new process for authentication

Table 3: Paper and Digital Pros & Cons

Envisioning a Digital Census

While the specific details of how a Digital Census would function have yet to be determined, the following 15 points provide a basic, high level approach as a starting point for discussion:

1. The Communications Directorate creates a Communication Plan to socialize/advertise the Digital Census design to the American public starting early in the decade.
2. Geography Division's GSS initiative facilitates a highly accurate and complete address frame by the end of the decade.
3. Perform a mail-out of letters/postcards to distribute the unique codes required for use with Internet or phone Census response. The code links the respondent's address and Census data with a Master Address File ID (MAFID).
4. After the letter/postcard mail-out, the public begins self-response using the Internet, or using the phone (both options result in digital data collection).
5. After a predetermined time period, or after Internet and phone response rates decline to a predetermined threshold, begin Nonresponse Followup in the field using secure smartphones equipped with a Census-designed data collection application. Internet and phone response should stay open during NRFU in order to continue removal of completed responses from the NRFU workload.
6. Use of secure smartphones for data collection should take advantage of wireless communication capability, but not be critically dependent on it to complete operations.
7. The smartphone data collection application should be streamlined enough to work for both Housing Units and Group Quarters, and should include payroll capability.
8. Field physical infrastructure should be minimized to incorporate a higher number of "micro-offices" equipped with minimal accommodations and staff. Enumerators should work out of their homes.
9. Make use of administrative records available in a digital format as appropriate to reduce NRFU workload, but do not make the overall Census design dependent on administrative records.
10. Research strategies to minimize the number of potential responses per MAFID.
11. Permit addition of new addresses during NRFU, using the smartphone technology equipped with GPS
12. Advertise for those "missed" in the Census to respond using the phone. Verify these added addresses and responses during NRFU.
13. Through QA design and application coding for the Internet and the smartphone, correct data issues as close to the source (the HU/GQ) as possible, not on the back-end.
14. Minimize the number of HQ database systems needed to conduct the Census.
15. Discourage data *collection* of any kind using paper.

Conclusion

Complexity costs are not only related to what we might call "bad" complexity, but also include the costs across the portfolio and throughout the enterprise that result from all the complexity in the business. Therefore, attacking complexity costs requires a combination of removing bad complexity as well as removing non value-added costs across all complexity dimensions.[11]

Based on the positive momentum driving us to transform the way that the Bureau executes the Census, and the basic assumption that at least some digital technology will be used in 2020, we must critically evaluate the options on the table. The deliberation and decision making process must consider how the introduction of digital technology will impact all stakeholders.

Finalizing the initial 2020 plan that calls for using both paper and digital modes invariably increases complexity especially in the areas of data integration, capture, responsiveness, and customer support. This conclusion is clearly illustrated in Figure 5 above. Moreover, the hidden costs for paper (both time and materials) must be considered. Most vital is the fact that the introduction of any quantity of paper for data collection requires the Bureau to build a capacity to handle the maximum scale of response.

The decision that Bureau leadership must consider is whether it is in its best interest to pursue a mixed paper and digital Census or whether a digital option is a better approach. While this paper provides only a general outline of how the Bureau could conduct a Digital 2020 Census, the hope is that it demonstrates, at a minimum, that a Digital Census is not only a feasible option, but a desirable option for the Bureau and the general public. The transition to a Digital Census design will require major systems development and integration work, change management, and training. However, a Digital Census design arrived at early in the decade will provide a clear and defined set of goals and challenges that we can approach with a sense of stability and a spirit of enthusiasm. The Digital approach will also likely be well received by organizations charged with Census oversight. Finally, moving decisively to a Digital Census will signal the Census Bureau's willingness to take part in the inevitable and continuing expansion of the role of digital technology in American society, allowing us to produce a more responsive, efficient, and manageable program for 2020 and beyond.

[11] Wilson 72

Appendix 1

Complexity of All Paper Response

Dimension	Criteria	Analysis	Rating (High/Med/Low)
Management Complexity	Data Availability	An All Paper response produces the greatest lag time between the questionnaires being completed and the data being integrated in to the response universe. As a result, operations such as NRFU are increasingly difficult to plan and execute due to the unavailability of timely response data. Workarounds such as working with the USPS to try to determine what questionnaires are en-route leads to inefficiencies that are inherent to paper.	High
	Data Management	An All Paper response lacks certain data management capabilities that are available in other options. The risk of respondent data loss is highest compared to a digital response due to: a. Responses lost due to mail or sorting problems b. Handwriting that cannot be interpreted 100% accurately c. Stray markings that otherwise result in a misread, preventing complete data capture. Similarly, there is no redundancy of the data until it is captured by DRIS. Finally, paper brings the least control of access to the data. While digital response technologies can employ the use of authentication, encryption, and logging, paper is dependent on physical security.	Medium
	Respondent Support	An All Paper response requires that the Census send forms in an alternate language and/or assist respondents with completing the questionnaire.	Low
Information Systems Complexity	Response Capture	An All Paper response requires a complex process in order to prepare the paper forms for capture, and then to digitize the data. Forms have to be opened, optically converted to images, "read" by software, validated or keyed in by human beings, archived and eventually destroyed. Forms requiring translation or that otherwise cannot be scanned must be manually transcribed and re-processed.	Medium
	Systems Integration	An All Paper response requires a relatively moderate level of systems integration, compared to the other two options, due to the need to integrate the DRIS data capture systems with the greater response universe.	Medium

Organizational Complexity	Organizational Scale	An All Paper response requires cross-functional support from internal departments and external contractors. For instance, the party responsible for developing the questionnaire must work closely with DRIS in order to ensure that the form design is optimized for scanning (in addition to respondent usability).	Medium
	Organizational Experience	An All Paper response is highly familiar to the organization and the Bureau already has the experience and processes developed to execute an All Paper response.	Low

Table 4: All Paper

Complexity of a Digital Response

Dimension	Criteria	Analysis	Rating (High/Med/Low)
Management Complexity	Data Availability	A Digital response enables questionnaires completed by all actors to be immediately integrated in to the response universe. As a result, respondent data may be made available in real-time for better planning and decision making.	Low
	Data Management	A Digital response provides a high level of support for data management, including safe-guards to protect against data loss. Because the content of the questionnaire is immediately submitted electronically, it can be replicated and backed up with minimal risk of data loss. Further, encryption and access control techniques provide the Bureau with greater control and flexibility over who has access to what, and when. An electronic audit trail enables management to identify potential problems in the process and monitor data collection status.	Low
	Respondent Support	A Digital response will require the Bureau to provide a basic level of technical support for respondents responding via the Internet. Conversely, a digital questionnaire can enable alternate presentations of the questionnaire that may ease the completion of the form and ultimately lead to a higher level of respondent-provided data accuracy. For instance, a "wizard" format may be used to guide the respondent through a series of screens depending on prior responses. "Screen tips" can enable the respondent to receive help on a specific form field right on the questionnaire.	Medium
Information Systems Complexity	Response Capture	A Digital response will enable the Bureau to capture respondent data digitally, eliminating the need to perform a time-consuming and complex digitization. Digital questionnaires enable real-time validation and ensure standardization of data, permitting the instantaneous integration of respondent data in to the response universe.	Low
	Systems Integration	A Digital response eases systems integration as a result of the data already being digital and the ability for respondent data to be directly entered in to the response universe.	Low
Organizational	Organizational Scale	A Digital response would likely require a relatively moderate level of	Medium

Complexity		organizational coordination in order to implement supporting systems such as firewalls, security, etc.	
	Organizational Experience	A Digital response will require the Bureau to develop new processes and procedures to support the use of this new technology.	High

Table 5: Digital Response

Mixed Paper and Digital Response

Dimension	Criteria	Analysis	Rating (High/Med/Low)
Management Complexity	Data Availability	A Mixed Paper and Digital response will enable management to make more timely decisions based on responses submitted via digital means. The use of paper will limit management's ability to fully capitalize on the benefits of a digital response since paper will still be used.	Medium
	Data Management	A Mixed Paper and Digital response will likely increase the complexity involved with managing data. The use of two response technologies will result in the need to develop tracking, storage, and security procedures for both sources of data.	High
	Respondent Support	A Mixed Paper and Digital response increases the resources needed to provide support for respondents using paper and digital. The degree of complexity will be dependent on the role of paper, though respondent support will inevitably be more complex than the use of paper or digital alone.	High
Information Systems Complexity	Response Capture	A Mixed Paper and Digital response increases the effort and resources required to support a dual-technology response.	High
	Systems Integration	A Mixed Paper and Digital response requires a high degree of systems integration due to the number and variety of systems needed to support a digital and paper response. A high degree of coordination is required to ensure the proper integration and flow of data.	High
Organizational Complexity	Organizational Scale	A Mixed Paper and Digital response requires cross-functional support from internal departments and external contractors to a higher degree than either paper or digital alone.	High
	Organizational Experience	A Mixed Paper and Digital response is new to the Bureau. Though the Bureau has significant experience with handling paper responses, the introduction of a digital response will require new procedures to be introduced to handle the interaction of the two systems.	High

Table 6: Mixed Paper and Digital Response

Appendix 2

The following table of lifecycle costs was derived from a study conducted by the National Academy of Sciences which breaks down major costs from the 2010 Decennial Lifecycle (2002-2013). We have added the Estimated Percentage Cost from Paper and Estimate Paper Lifecycle Cost columns to attempt to capture a conservative estimate of the paper-based costs in Decennial operations. Items that are in bold in the table indicate that they add a significant paper-based cost to the operation.

Category	Key Activity	Lifecycle Cost ($ M)	Estimated Percentage Cost from Paper (%)	Estimated Paper Lifecycle Cost ($ M)	Rationale
Major Contracts	**Decennial Response Integration System (DRIS)**	**$981**	**100%**	**$981M**	DRIS processed paper-based responses for the 2010 Decennial. Scanning was required to process paper.
	Field Data Collection Automation (FDCA)	$801	0	0*	
	Communications	$308	0	0*	
	Data Access and Dissemination System (DADS)	$176	0	0*	
	Printing	**$179**	**100%**	**$179M**	A large contract was required for printing millions of paper questionnaires.
	Mail Out/Mail Back Postage	**$257**	**100%**	**$257M**	The cost of Mail Out/Mail Back postage was due to the paper operations.
Office Space & Staff	Regional Census Centers (RCC)	$828	0	0*	
	Local Census Offices (LCO)	$1,301	0	0*	
	Address Canvassing	$386	0	0*	
	Group Quarters Advance Visit	$17	0	0*	
	Group Quarters Enumeration	$80	0	0*	
	Group Quarters Validation	$71	0	0*	
	Coverage Measurement	$83	0	0*	
	Puerto Rico	$62	0	0*	
	Island Areas MOAs	$37	0	0*	
	Field Verification	$39	0	0*	
	Non Response Follow-Up	$2,744	0	0*	
	Vacant/Delete	$341	0	0*	
	Military	$5	0	0*	
	Remote Alaska	$4	0	0*	

	Service-Based Enumeration	$41	0	0*	
	Transient Night	$11	0	0*	
	Update Enumerate	$108	0	0*	
	Update Leave	$116	0	0*	
	Urban Update Leave	$2	0	0*	
	Fingerprinting	$148	0	0*	
Other	**National Processing Center Census Operations**	**$364**	**15%**	**~ $55M**	NPC supports various other operations and functions in addition to paper. Approximately 15% of its costs stem from paper (i.e. Mail Out/Mail Back, NRFU, etc).
	HQ Staff and All Other (2002-2013 lifecycle)	**$2,986**	**10%**	**~ $300M**	HQ staff and others are responsible for a wide variety of activities that span IT, survey methodology, contract management, communications, etc. Approximately 10% of Census staff must support paper, in one form or another, from printing, and scanning.
Total Cost ($B)		**$12.5B**		**$1.8B**	
Cost of paper per household				**$12 - $14**	138 million households were counted and paper forms were used to support data collection
Source: Envisioning the 2020 Census, National Academy of Sciences, http://www.nap.edu/catalog/12865.html (p.36)					
*We do not have sufficient data to identify the cost of paper for these operations. The focus of this analysis is on the primary cost contributors.					

The pie chart below provides a high level illustration of the cost of paper across the five major contributing areas: DRIS, the printing contract, Mail Out/Mail Back Postage, National Processing Center Census Operations and HQ Staff and all other. The chart illustrates that DRIS is the single highest cost contributor over the lifecycle comprising approximately 55% of the total paper-based cost.

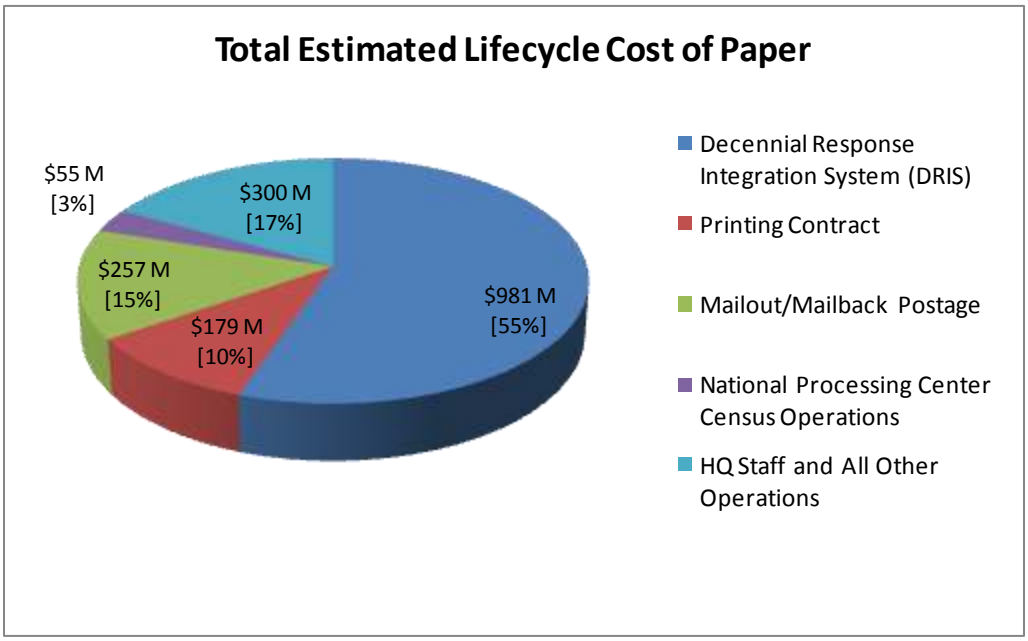

Figure 6: Total Estimated Lifecycle Cost of Paper

Works Cited

Brown, Lawrence D., Michael L. Cohen, Daniel L. Cork, and Constance F. Citro. “Envisioning the 2020 Census.” Washington: National Academies, 2010.

Castro, Daniel. “E-Census Unplugged: Why Americans Should Be Able to Complete the Census Online.” Rep. Washington: Information Technology and Innovation Foundation, 2008.

De Leeuw, Edith D. "To Mix or Not to Mix Data Collection Modes in Surveys." Journal of Official Statistics, 2005.

“Digital Nation: Expanding Internet Usage.” National Telecommunications and Information Administration, 2011.

Dillman, D. A. and Christian, L. M. “Survey Mode as a Source of Instability in Responses across Surveys. Revised version of a paper presented at the Workshop on Stability of Methods for Collecting, Analyzing and Managing Panel Data.” Cambridge: American Academy of Arts and Sciences, 2003.

Fox, Jean, William Mockovak, Sylvia Fisher, and Christine Rho. “Usability Issues Associated with Converting Establishment Surveys to Web-Based Data Collection.” Bureau of Labor Statistics, 2003.

Griffin, Deborah H., Donald P. Fischer, and Michael T. Morgan. “Testing an Internet Response Option for the American Community Survey.” U.S. Bureau of the Census, 2001.

Weinberg, Daniel H. "Plan for the Research and Testing Phase of the 2020 Census." International Census Forum, 2010. Presentation.

Wilson, Stephen A and Andrei Perumal. “Waging War on Complexity Costs.” New York: McGraw Hill, 2010.