Confidentiality Perceptions and Response: Preliminary Evidence from Business

Nick Greenia, Internal Revenue Service; J. Bradford Jensen, U.S. Census Bureau; and Julia Lane, American University, The Urban Institute, and LEHD program, U.S. Census Bureau

Concern about declining response rates on the decennial Census led to the development of a very interesting literature targeted at understanding the sources of the decline. One entire thread of this literature has examined the role of public perceptions of the degree to which Federal statistical agencies protect respondent data, and the link between this and respondent behavior (Singer, 2001). The results of the analysis have been extremely useful in informing statistical agency policy in the United States, and the model has served national statistical institutes abroad as well.

The burgeoning interest in perceptions research on individual respondents has not been matched by similar research on business respondents. This paper builds on one of the few studies of businesses (Greenia et al., 2001) and examines the links between business perceptions of statistical agency competence, and data protection and their willingness to provide data to those agencies.

The results of the study suggest that there is indeed a link between business perceptions and their willingness to provide data. Although firm conclusions cannot be drawn on the basis of such a preliminary study, if the results were to hold up under a more representative survey, they might well suggest that an initiative by statistical agencies to convey information to their survey populations about both how the data are protected and how they are used could improve response rates and data quality.

◆ Background and Literature Review

The lack of quantitative research does not mean that no attention has been paid to the confidentiality of business data provided for Federal statistical purposes [1]. In 1992, the U.S. Office of Management and Budget (OMB) established a working group that not only noted the differences between household and business perceptions of confidentiality, but also identified several factors that were likely to affect business trust in Government statistical agencies, and consequently their willingness to provide data [2]. One major factor was thought to be the existence and enforcement of penalties for unauthorized disclosure of information about an economic entity. Another was the use to which data were put--in other words, the benefit provided to policymakers as a result of the business responding to the survey. Finally, the OMB working group was concerned about the burden imposed on businesses by Federal statistical agencies.

Preliminary work by Willimack and coauthors supported the idea that businesses are primarily concerned about cost associated with data provision. In a series of interviews with some 30 large businesses, they found that business respondents would be willing to trade off privacy concerns and permit data-sharing across statistical agencies if it would mean that they had to fill out fewer surveys. However, their work did not quantitatively examine the interaction between businesses' willingness to provide data to Government and their views of the Government's ability to protect the data, nor did it examine whether small businesses had different views of statistical agencies than did large.

◆ Data

The data are derived from a recent mail survey of 5,000 businesses administered by the Urban Institute on behalf of the U.S. Census Bureau. The data design and

This study would not have been possible without the extraordinary efforts put into the development of the survey by Diane Willimack and Kristin Stettler, U.S. Census Bureau. The authors would also like to thank Dr. Frederick Knickerbocker for facilitating this work and Faye Schwartz for excellent research assistance. All views expressed represent those of the authors and not necessarily the institutions they represent.

collection are described in much more detail in Greenia et al. (2001)

Questionnaire Design

The survey itself was designed to answer a number of questions about business attitudes towards confidentiality. In this study, we analyse two issues in some depth: first, we provide some preliminary evidence of businesses' perceptions of the ability of statistical agencies to do their jobs, and second, we examine their perception of the ability of statistical agencies to protect their data. In each case, we are interested in examining the relationship between their perceptions and their willingness to provide the Federal Government with data.

There are three questions in the survey that address these questions. The first one, Question 5, which is reproduced in Box 1, examines the willingness of respondents to provide data to different types of recipients. In this paper, we are particularly interested in their responses to parts a and b of the question: the concern associated with providing data to Federal regulatory agencies (5a) and Federal statistical agencies (5b).

The second question of interest, Question 6 in the survey, attempts to quantify the respondent's sense of the ability of the Government to collect data (part a) and provide useful information to policymakers (part b). It follows with a question about the Federal Government's ability to protect data (part c), which is expanded on in the subsequent question.

5. In your opinion, how concerned is your company about providing business data (e.g., number of employe payroll, sales, profits, etc.) to each of the following? *Circle one response for each line.*

	Not at all Concerned	Somewhat Concerned	Very Concerned	Concerned
a. Federal regulatory authorities (e.g., Environmental Protection Agency, Securities and Exchange Commission, and Federal Trade Commission)	1	2	3	4
b. Federal statistical agencies (e.g., Census Bureau, Bureau of Labor Statistics and Bureau of Economic Analysis)	1	2	3	4
c. Not-for-profit researchers (e.g., universities, think tanks and research organizations)	1	2	3	4
d. For-profit researchers (e.g., market researchers and consulting firms)	1	2	3	4
e. Other businesses	1	2	3	4
f. The general public	1	2	3	4

6. Business data are collected for statistical purposes by both the Federal Government and the private sector (e.g., universities, think tanks, research organizations, market researchers, consulting firms, etc.). For each statement below, please indicate if you strongly agree, somewhat agree, somewhat disagree, or strongly disagree. *Circle one response for each line.*

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	
a. The Federal Government is better than the private sector at collecting business data for statistical purposes.	1	2	3	4	DK/NA
b. The Federal Government is better than the private sector at providing useful information to government policymakers.	1	2	3	4	DK/NA
c. The Federal Government is better than the private sector at protecting business data from being released to those without authority to have it.	1	2	3	4	DK/NA

Question 7 is directly focused on examining respondents' views and understanding of the Federal Government's ability to protect data. In this paper, we examine the links between the responses to parts a, b, and d and their responses to Question 5a and b.

The analytical section will examine the interactions across the responses to these different items.

Data Collection and Response Rates

The sample frame for the survey was obtained from Dunn and Bradstreet's commercial data base, which represents 11,300,000 businesses in the U.S. [3]. From this sample frame, we selected a stratified random sample of 5000 cases; 1,250 in each of four strata defined by the number of employees--1-49 employees; 50-249; 250-499; and 500 plus. Of the 5,000 total businesses, 2,530 were multiunit businesses with headquarters locations, and 2,470 were single unit companies with only one business [4]. While the geographic coverage was national, and the sample was representative of Dunn and Bradstreet's data base, we do not attempt to weight the sample to assign any kind of representativeness to the sample and simply view our results as preliminary evidence.

We followed the Dillman mail survey approach. The first survey was followed 2 weeks later by a reminder postcard. This was followed by a repeat survey mailout 2 weeks after that. The outcome was that 509 questionnaires were returned, resulting in a response rate of just over 10 percent [5]. While this response rate is quite reasonable for a private sector survey of businesses, the response rates for mandatory, Government surveys often exceed 80 percent. Since a more detailed review of the response rates is available in Greenia (2002), we simply summarize the results by noting that response rates declined with business size; and varied by industry type--manufacturing firms were less likely to respond than were service sector firms. Both the response rate and the sample size, however, mean that our results should be viewed as preliminary rather than definitive in nature.

7. For each statement below, please indicate if you strongly agree, somewhat agree, somewhat disagree or strongly disagree. *Circle one for each line.*

	Strongly Agree	Somewhat Agree	Somewhat Disagree	Strongly Disagree	
a. I believe that Federal statistical agencies (e.g., Census Bureau, Bureau of Labor Statistics, and Bureau of Economic Analysis) keep data provided by businesses confidential.	1	2	3	4	DK/NA
b. I believe that Federal statistical agencies do not release information by which a company or its data can be identified.	1	2	3	4	DK/NA
c. I believe that Federal statistical agencies do not share data provided by businesses with other Government agencies.	1	2	3	4	DK/NA
d. I believe that Federal statistical agencies do not release data provided by businesses to people outside the Government.	1	2	3	4	DK/NA
e. I believe that any Federal agency, such as the Internal Revenue Service, Small Business Administration, and Federal Trade Commission, can access business data my company has provided to other Federal agencies whenever it wants.	1	2	3	4	DK/NA

◆ Results

What are respondents' views in general? A brief examination of Table 1 reveals that, by and large, there is much higher trust in the Federal statistical system than other potential data collectors. However, one in three respondents is still quite concerned about providing data to these agencies, suggesting that in absolute terms, there is still a large group of businesses that mistrust statistical agencies.

There do appear to be differences in opinions by size of firm. The smallest businesses are substantially more reluctant to provide their data to Federal regula-

Table 1: Proportion of respondents who were very or extremely concerned about providing business data to the following entities (by employment size) (Question 5)

Type of recipient	1-49	50-249	250-499	500+	Total
Federal Regulatory Authority (EPA, SEC, FTC)	48.40%	43.20%	25.00%	32.00%	38.20%
Federal Statistical Agency (BLS, Census, BEA)	34.10%	33.10%	20.00%	27.80%	29.30%
Not for profit researchers (universities, think-tanks, research organisations)	42.20%	44.80%	41.00%	37.50%	41.60%
For-profit researchers	60.90%	62.40%	68.70%	56.70%	62.10%
Other businesses	73.60%	76.00%	77.80%	67.70%	73.90%
The general public	70.50%	73.60%	66.00%	53.60%	66.70%

Source: Greenia et al. 2001

tory agencies than are large ones, and slightly more reluctant to provide their data to Federal statistical agencies. All businesses, regardless of size, are most suspicious of for-profit researchers, however.

The perception of the statistical system's competence at doing its job seems, unfortunately, not particularly good. As Table 2 indicates, only one in three respondents believe that the "Federal Government" collects data better than the private sector, or that it provides more useful information than the private sector. The differences across size class are not marked, however.

An examination of the results presented in Table 3 suggests that, by and large, more businesses believe that their data are protected than do not. Here, however, we do see a marked difference by size of firm--the smallest firms have the least confidence in Federal statistical agencies. This is consistent regardless of the way in which the question is phrased: the responses to the questions in the first two rows are consistent. The apparent contradiction between the answers to the first two questions and the second two (which is that the Government releases data to other agencies and to the public) may possibly be due to a confusion between the release of microdata and aggregate level data.

In any event, possibly the most interesting result in this table is the response to the last question--businesses overwhelmingly believe that particular Federal agencies can access business data at any time they want, despite the existence of very strong legal constraints on such access.

Table 2: Proportion of respondent who somewhat or strongly agreed with the following statements (Question 6)

	1-49	50-249	250-499	500+	Total
The Federal Government is better than the private sector at collecting business data for statistical purposes.	32.60%	23.80%	38.00%	33.30%	31.50%
The Federal Government is better than the private sector at providing useful information.	38.80%	26.20%	31.30%	29.20%	31.60%
The Federal Government is better than the private sector at protecting business data from being released to those without authority to have it.	35.90%	39.20%	42.00%	24.00%	35.60%

Source: Greenia et al., 2001

Table 3: Proportion of businesses somewhat or strongly agreeing with the following statements (Question 7)

	1-49	50-249	250-499	500+	Total
I believe that Federal statistical agencies keep data provided by businesses confidential.	48.10%	69.80%	69.70%	60.80%	61.60%
I believe that Federal statistical agencies do not release information by which a company or its data can be identified.	41.10%	66.40%	65.70%	57.70%	57.10%
I believe that Federal statistical agencies do not share data provided by businesses with other Government agencies.	21.10%	21.60%	27.60%	21.60%	22.80%
I believe that Federal statistical agencies do not release data provided by business to people outside the Government.	28.30%	41.70%	43.30%	38.50%	37.60%
I believe that any Federal agency, such as the Internal Revenue Service, Small Business Administration, and Federal Trade Commission can access business data my company has provided to other Federal agencies whenever it wants.	71.10%	68.50%	67.30%	73.20%	70.00%

Source: Greenia et al., 2001

How do these different beliefs held by businesses correlate with their concerns about providing data to Federal statistical agencies? We examine this by estimating simple linear probability models which regress the response to questions 5a and b (coded 1 if the respondent was extremely or very concerned about providing data to Federal regulatory or statistical agencies respectively) against their responses to question 6. These results are reported in Table 4, and the results of a similar regression against the responses to question 7 are reported in Table 5.

The most interesting result from Table 4 is that the data support the a priori expectations that the more the business believes the Government does a good job in collecting and producing information, the less concerned it is about providing data to the Government, as evidenced by the negative and significant coefficients reported in the first three rows of the table. Indeed, the size of the coefficients is remarkably similar regardless of whether the data collection entity is a Federal statistical or a Federal regulatory agency. However, the constant term is markedly higher when the respondents are asked about Federal regulatory agencies--suggesting that the level of concern is higher for regulatory agencies.

One other interesting result is apparent from an examination of the table--namely, that the response varies significantly by industry. In particular, firms in the Financial, Insurance, and real estate industry as well as in manufacturing are significantly less likely to be concerned about Federal regulatory agencies than are other firms--but they are no different in their concerns about statistical agencies than are firms in other industries.

Table 4: The Relationship between response and business perceptions of Government competence

Dependent Variable: Respondent said he/she was very or extremely concerned about providing business data to the following entities

	Federal Regulatory Authority (EPA, SEC, FTC)			Federal Statistical Agency (BLS, Census, BEA)		
The Federal Government is better than the private sector at collecting business data for statistical purposes.	-0.123 (2.59)**			-0.116 (2.44)*		
The Federal Government is better than the private sector at providing useful information.		-0.125 (2.42)*			-0.123 (2.59)**	
The Federal Government is better than the private sector at protecting business data from being released to those without authority to have it.			-0.132 (2.64)**			-0.149 (3.24)**
Manufacturing	-0.234 (2.35)*	-0.211 (2.10)*	-0.209 (2.09)*	-0.115 (1.24)	-0.103 (1.11)	-0.102 (1.11)
Wholesale	-0.168 (1.35)	-0.138 (1.11)	-0.140 (1.13)	-0.048 (0.41)	-0.035 (0.30)	-0.039 (0.34)
Retail	-0.185 (1.71)	-0.167 (1.54)	-0.151 (1.39)	0.001 (0.01)	0.013 (0.13)	0.031 (0.31)
FIRE	-0.318 (2.88)**	-0.299 (2.69)**	-0.290 (2.61)**	-0.211 (2.04)*	-0.207 (2.01)*	-0.200 (1.95)
Services	-0.266 (2.96)**	-0.232 (2.58)*	-0.219 (2.45)*	-0.174 (2.09)*	-0.158 (1.91)	-0.146 (1.78)
Constant	0.641 (7.45)**	0.595 (6.99)**	0.593 (7.03)**	0.440 (5.53)**	0.429 (5.50)**	0.435 (5.62)**
Observations	404	404	404	411	411	411
R-squared	0.05	0.03	0.04	0.04	0.04	0.05
Mean of Dependent Variable	0.36			.28		

Absolute value of t statistics in parentheses; * significant at 5%; ** significant at 1%; omitted industry is agriculture, mining, contract construction

A very similar pattern is uncovered by an examination of the relationship between trust and data provision (Questions 5 and 7), the results of which are reported in Table 5. Just as in the previous table, the concern a business exhibits at providing data to Federal statistical and regulatory agencies is highly correlated by its view of the agencies' competence--in this case, by the trust in the ability of the agency to protect its data.

As before, there are interesting industry differences in businesses' views of regulatory and statistical agencies. In this case, however, businesses in retail trade as well as in finance, insurance, real estate, and manufacturing are significantly less likely to distrust Federal regulatory agencies than are firms in other industries--but their views of Federal statistical agencies are not different.

It is easy to speculate about the reasons for these industry differences--but clearly only empirical analysis will provide the answer. An obvious possibility is that Federal regulatory agencies have more interaction with these industries than others. Hence, firms in these industries have actually had contact with the regulatory agencies, and have found them generally competent. An alternative possibility is that the reverse is the case!

◆ Summary

The preliminary research described in this study suggests that there is a link between businesses' willingness to provide confidential data to statistical agencies and their trust in both the competence of the agency in using the data and the ability of the agency to protect it. The research also suggests that a worryingly high number of respondents are quite distrustful of Federal statistical agencies, and this was substantiated by many of the handwritten comments on the returned surveys.

It is clear, however, that more representative data should be collected in order to support these preliminary results. Our rather disappointing response rates, combined with the evidence that businesses trust statistical agencies more than think tanks or outside researchers, suggest that the appropriate entities to conduct such research are Federal statistical agencies themselves. In addition to the imprimatur and data collection infrastructure provided by the statistical system, the agencies stand to gain the most from such research--both in terms of developing a relationship with the respondent community and in terms of supporting policy decisions. As statistical agencies well know, good decisions are driven by good data.

Table 5: The relationship between response and businesses' perceptions of data protection

Dependent Variable: Respondent said he/she was very or extremely concerned about providing business data to the following entities						
	Federal Regulatory Authority (EPA, SEC, FTC)			Federal Statistical Agency (BLS, Census, BEA)		
I believe that Federal statistical agencies keep data provided by businesses confidential.	-0.238 (4.98)**			-0.199 (4.45)**		
I believe that Federal statistical agencies do not release information by which a company or its data can be identified.		-0.192 (4.05)**			-0.150 (3.38)**	
I believe that Federal statistical agencies do not release data provided by business to people outside the Government.			-0.166 (3.45)**			-0.119 (2.64)**
Manufacturing	-0.213 (2.17)*	-0.187 (1.89)	-0.197 (1.98)*	-0.106 (1.17)	-0.082 (0.89)	-0.091 (0.99)
Wholesale	-0.167 (1.37)	-0.130 (1.06)	-0.129 (1.04)	-0.057 (0.50)	-0.022 (0.19)	-0.024 (0.20)
Retail	-0.214 (2.01)*	-0.150 (1.40)	-0.172 (1.59)	-0.034 (0.34)	0.026 (0.26)	0.007 (0.07)
FIRE	-0.311 (2.86)**	-0.280 (2.56)*	-0.279 (2.53)*	-0.217 (2.14)*	-0.187 (1.83)	-0.187 (1.82)
Services	-0.243 (2.78)**	-0.225 (2.55)*	-0.230 (2.59)*	-0.169 (2.08)*	-0.150 (1.83)	-0.154 (1.87)
Constant	0.719 (8.18)**	0.656 (7.63)**	0.616 (7.29)**	0.529 (6.47)**	0.466 (5.86)**	0.431 (5.52)**
Observations	404	404	404	411	411	411
R-squared	0.08	0.06	0.05	0.07	0.05	0.04

Absolute value of t statistics in parentheses * significant at 5%; ** significant at 1% ; omitted industry is agriculture, mining, contract construction

◆ References

Biemer, P.B. and Fecso, R.S. (1995), "Evaluating and Controlling Measurement Error in Business Surveys," Chapter 15 in B. Cox, et al. (editors), *Business Survey Methods*, Wiley, NY.

Christianson, A. and Tortora, R. (1995), "Issues in Surveying Businesses: An International Survey," Chapter 14 in B. Cox, et al. (editors), *Business Survey Methods*, Wiley, NY.

Cox, B.G. and Chinnappa, B. N. (1995), "Unique Features of Business Surveys," Chapter 1 in B. Cox, et al. (editors), *Business Survey Methods*, Wiley, NY.

Dillman, D. A. (1978), *Mail and telephone surveys: The total design method*, Wiley, NY.

Edwards, W.S. and Cantor, D. (1991), "Toward a Response Model in Establishment Surveys," Chapter 12 in Biemer, P.P.; Groves, R.M.; Lyberg, L.E.; Mathiowetz, N.A.; and Sudman, S. (editors), *Measurement Errors in Surveys*, Wiley, NY.

Greenia, N.; Jensen, J. B.; and Lane, J. (2001), "Business Perceptions of Confidentiality," in Doyle, Pat; Lane, Julia; Zayatz, Laura; and Theeuwes, Jules (editors), *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, North Holland.

Mackie, C and Bradburns, N. (editors), *2000 Improving Access to and Confidentiality of Research Data: Report of a Workshop*, National Research Council.

Nichols, E.; Willimack, D. K.; and Sudman, S. (1999), "Balancing Confidentiality and Burden Concerns in Censuses and Surveys of Large Businesses," paper presented to the Washington Statistical Society, U.S. Bureau of the Census, Washington, DC, September.

Singer, E. (2001), "Public Perceptions of Confidentiality and Attitudes Towards Data Sharing by Federal Agencies" in Doyle, Pat; Lane, Julia; Zayatz, Laura; and Theeuwes, Jules (editors), *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, North Holland.

Singer, Eleanor; Presser, S.; and Vanhoewyk, J. (1999), "Public Attitudes Toward Data Sharing by Federal Agencies," in *Record Linkage Techniques--1997 Proceedings of an International Workshop and Exposition*, National Academy Press.

Wallman, K. and Coffey, J. (1999), "Sharing Statistical Information for Statistical Purposes" in *Record Linkage Techniques--1997 Proceedings of an International Workshop and Exposition*, National Academy Press.

Willimack, Diane K.; Nichols, Elizabeth; and Sudman, Seymour (forthcoming), "Understanding Unit and Item Nonresponse in Business Surveys," in Groves, Robert M.; Dillman, Don A.; Eltinge, John L.; and Little, Roderick J.A. (editors), *Survey Nonresponse*, Wiley, NY.

Willimack, Diane K; Seymour Sudman; Nichols, Elizabeth; and Mesenbourg, Thomas (1999), Cognitive Research on Large Company Reporting Practices: Preliminary Findings and implication for Data Collectors and Users, unpublished paper presented to the Census Advisory Committee of Professional Associations, U.S. Bureau of the Census, Washington, DC, March 2000.

◆ Footnotes

[1] The protection of confidentiality in this paper is defined as the restriction of access to information about the individual party/entity once it has been provided--for statistical or administrative purposes--to a second party charged with the collection responsibility. The confidentiality protection responsibility is traditionally viewed as residing with the collecting party, even when the law permits third and fourth parties to access the data. Indeed, the consequences of any breach of confidentiality would almost always be borne by the collecting

party in the form of reduced response rates and less precise responses as the cost exacted for such violations.

[2] These are similar to the differences between collecting household and business data generally (see, e.g., Box and Chiannapa, 1995).

[3] This total is derived from the Dunn and Bradstreet report and differs from the size of the Census Bureau's Business Register, most likely due to differences in definition of a "business."

[4] We did not ask Dunn and Bradstreet to strip out subsidiaries from the file.

[5] We did not separate out subsidiaries or non-headquarters responses. ■