

PUBLIC ATTITUDES TOWARD DATA SHARING: FINDINGS AND IMPLICATIONS

Eleanor Singer, *Survey Research Center, University of Michigan,*
Stanley Presser, *University of Maryland*

Eleanor Singer, Survey Research Center, P. O. Box 1248, University of Michigan, Ann Arbor, MI 48106

Key Words: Privacy, Confidentiality, Census, Administrative Records, Data Sharing

In a report by a Canadian parliamentary committee entitled "Privacy: Where Do We Draw the Line," the Canadian approach to privacy protection is characterized as "passive," as compared with that of European countries. The Committee surmises that this may be due to the influence of the United States, which has, they say, "long downplayed the importance of an independent and proactive data protection regime."

These differences between countries raise some interesting sociological questions. Does policy in this area follow public opinion, for example, or is opinion molded by policy? To answer such questions, we would have to compare the opinion of European, Canadian, and U.S. publics over time toward the privacy of information, and in particular toward the issue of data sharing among government agencies and between the government and the private sector. Unfortunately, no such comparative data are available, and in fact until recently we had no data even about U.S. public opinion in this area. Now, however, we have the results of two surveys, one carried out in 1995 by the Joint Program for Survey Methodology and the Maryland Survey Research Center, the other carried out in 1996 by Westat in behalf of the Census Bureau. Many of the same questions were asked on both surveys, in identical form; both were RDD surveys and achieved response rates around 65%. In this paper, we briefly describe some results from both surveys, and draw several implications for public policy from them.

Findings from the 1995 Survey

The 1995 JPSM survey was administered between late February and early July to a random digit dial sample of households in the continental United States. In each household, one respondent over 18 years of age was selected at random. The total number of completed interviews was 1443.¹

¹ The data were weighted to correct for unequal probabilities of selection (due to households containing different numbers of adults and phone lines) and

Because we suspected that questions about data sharing and about the confidentiality of census information would not be meaningful unless respondents had some idea of the kind of information involved, the interview began with the five demographic questions on the 1990 census short-form questionnaire, which asked for the respondent's name, age, race, ethnicity, and marital status. Subsequent questions about confidentiality and data sharing referred back to the content of these five questions. As a result, we hoped that responses to questions probing attitudes toward agencies' sharing of data with the Census Bureau and beliefs about the latter's safeguarding of identified data would be anchored in the specific content of the information involved.

Besides anchoring questions about data sharing in the specific information asked for on the short census form, we also acquainted respondents with the fact that the 1990 census had failed to count a significant number of people, and that the communities in which these people lived were, as a result, deprived of full political representation and economic benefits. Thus, the context for the questions on data sharing was the undercount and its consequences.

The questionnaire included several questions about the issue of data sharing. Respondents were asked whether they favored or opposed three specific agencies--the Social Security Administration (SSA), the Immigration and Naturalization Service (INS), and the Internal Revenue Service (IRS)--giving the Census Bureau the name, address, age, sex, marital status, and (in the case of SSA and INS) race of persons in their files *who are missed in the census*. These three agencies were selected because they are among the most likely candidates for supplementing lists of persons enumerated in the census, and also because two of the three might be perceived as threatening by respondents. Presentation of the three agencies was systematically rotated.

The questionnaire also included a large number of items we believed, on the basis of prior research (Blair,

poststratified to the March 1993 Current Population Survey distributions on region, gender, race, age, and education. The weights were scaled so that the weighted N equaled the actual number of interviews.

1994; Singer, Schaeffer, and Raghunathan, 1997), to be related to attitudes toward other statistical agencies' sharing data with the Census Bureau. We used exploratory factor analysis to reduce this large number of potential independent variables to seven indexes, each consisting of variables highly correlated with each other and all related in the same direction to the dependent variables: (1) *Importance* attached to the census; (2) *Knowledge about the census*; (3) *Belief* that the Census Bureau currently shares identifiable data with other agencies; (4) *Control*, which measures the respondent's perceived influence over government; (5) *Distress* at disclosure of personal information; (6) *Trust* in government; and (7) *Privacy invasion*. These indexes were then used to predict attitudes toward data sharing, controlling simultaneously for a number of demographic characteristics.

Sharing of data with the Census Bureau was widely approved by the 1995 sample. Between 70 percent (IRS) and 76 percent (INS, SSA), depending on the specific agency asked about, said they would favor the agency giving the Census Bureau the name, address, age, sex, and marital status of people who are missed in the census. These numbers are similar to those found in other surveys (e.g., Blair, 1994; Singer, Schaeffer, and Raghunathan, 1997). A considerably smaller number, but still a majority (54 percent), favor the Census Bureau's getting everyone's name, address, age, sex, race, and marital status from the records of other government agencies in order to eliminate the filling out of census forms altogether.

Who are the people who favor data sharing by one or more agencies, and how do they differ from others, who remain opposed to this practice? Estimating a logistic regression equation with controls for demographic characteristics, we found that two of the attitudinal variables described above significantly predicted attitudes toward data sharing (Table 1).

Those placing greater importance on the census were more likely to approve of this practice, and those more distressed at disclosure of personal information about them were more likely to be opposed. In addition, trust in government was marginally significant, greater trust being associated with greater favorability. Among the demographics, nonwhite race, and (marginally) income and gender were significant, with blacks and women more likely to be opposed to data sharing and those with higher income more likely to have favorable attitudes toward it. Those who did not disclose their income in the interview were significantly less likely to favor the sharing of administrative data across federal agencies, a finding also reported in Singer, Schaeffer, and Raghunathan (1997).

Trends in Public Attitudes toward Data Sharing

As already noted, many of these questions were replicated in a survey carried out by Westat in the fall of 1996, a little more than a year after the original survey. The Westat survey (Kerwin and Edwards, 1996) was also conducted with a sample of individuals 18 or older in U.S. households from June 11 to mid-September; the number of completed interviews was 1215.

The most significant finding emerging from a comparison of the two surveys was the *absence* of change with respect to attitudes relating to data sharing. In 1996 between 69.3% and 76.1%, depending on the agency, approved of other agencies sharing information from administrative records with the Census Bureau in order to improve the accuracy of the count, compared with 70.2% to 76.1% in 1995.² Responses to the INS, asked about in 1995, and the FSO, asked about in 1996, are comparable to those to the SSA. Responses are consistently least favorable toward the IRS. Westat documents only five significant changes ($p < .10$) among 22 questions asked about the Census Bureau on both surveys. These changes are mostly on the order of a few percentage points.

Trends in Attitudes toward Privacy

In contrast with attitudes toward data sharing and the Census Bureau, which showed virtually no change between 1995 and 1996, most questions about privacy and alienation from government showed significant change, all in the direction of more concern about privacy and more alienation from government. The relevant data are shown in Table 2.

There was a significant decrease in the percentage agreeing that "people's rights to privacy are well protected" and an insignificant increase in the percentage agreeing that "people have lost all control over how personal information about them is used." At the same time, there was a significant decline in the percentage *disagreeing* with the statement, "People like me don't have any say about what the government does," and a significant increase in the percentage agreeing that "I don't think public officials care much what people like me think" and in the percentage responding "almost never" to the question, "How much do you trust the government in Washington to do what is right?" The significant decline in trust and attachment to government manifested by these

² Text and tables use data weighted for number of residential phone numbers in the household and number of persons in the household, poststratified to census estimates of sex, race, age, education, and region.

questions is especially impressive given the absence of change in responses to the data sharing questions.

Willingness to Provide Social Security Number to Facilitate Data Sharing

One question of particular interest to the Census Bureau is the extent to which people would be willing to provide their Social Security Number (SSN) to the Census Bureau in order to permit more precise matching of administrative and census records. To clarify this issue, the Bureau asked Westat to include a question about SSN on the 1996 survey. The question (Q21) read as follows:

"The Census Bureau is considering ways to combine information from federal, state, and local agencies to reduce the costs of trying to count every person in this country. Access to Social Security numbers makes it easier to do this. If the census form asked for your Social Security number, would you be willing to provide it?"

About two thirds (65.9%) of the sample said they would be willing to provide the number; 30.5% said they would not; and 3.5% said don't know or did not answer the question.

The question about SSN was asked *after* the series of questions asking whether or not people approved of other administrative agencies sharing data with the Census Bureau. Therefore, it is reasonable to assume that responses to this question were influenced by opinions about data sharing, which the preceding questions had either brought to mind or helped to create. And, not surprisingly, there is a relationship between a large number--but not all--of the preceding questions and the question about providing one's SSN (see Table 3).

There are also significant relationships between political efficacy, feelings that rights to privacy are well protected, feelings that people have lost control over personal information, and trust in "the government in Washington to do what is right" and willingness to provide one's SSN (see Table 4).

These political attitude questions, it should be noted, were asked *after* the question about providing one's SSN, and so they could not have influenced the response to this question.

From the foregoing, it appears that there are two reasons underlying reluctance to provide one's SSN. *First, there are reasons associated with beliefs about the census:* People who are less aware of the census, who consider it less important, and who are less favorable toward the idea of data sharing are significantly less willing to provide their SSN. Low levels of education are also associated with these characteristics. *Second, however, is a set of beliefs and attitudes concerning privacy, confidentiality, and trust:* People who are more

concerned about privacy, who have less trust in the Bureau's maintenance of confidentiality, and who are less trusting of government in general are much less likely to say they would provide their SSN to the Census Bureau. The relationships described above suggest that the same clusters of beliefs predicting attitudes toward data sharing are relevant for willingness to provide one's Social Security Number, as well.

Knowledge, Trust, and Attitudes toward Data Sharing

It is an article of faith at the Census Bureau that respondents' willingness to provide data depends on their knowledge that the Bureau guarantees the confidentiality of their answers. In an effort to understand the meaning of confidentiality to respondents, and to assess its importance for their willingness to have agencies share data with the Bureau, we asked a split-ballot question near the end of the 1996 survey. One asked whether the Census Bureau was required by law to keep census information confidential; the other, whether the Bureau was forbidden by law from giving identified census information to other agencies.

The responses to the two versions of this question (Q22) are shown in Table 5.

Majorities of those who have an opinion give the correct answer to both questions; but the proportion answering DK is larger, and the proportion giving the correct answer smaller, when the question asks about giving other agencies identified information than when it asks about maintaining confidentiality.

As a follow-up to both questions, we asked those who said the Bureau is required to protect the information or forbidden from disclosing it, *whether or not they trusted the Bureau to uphold the law--that is, to keep the information confidential, or to refrain from disclosing it to other agencies.* Regardless of which version of Q22 they got, only two thirds of those who answered Yes to the factual question about legal requirements said they trusted the Bureau to comply with the law. Both knowledge and trust are independently (and positively) related to attitudes toward data sharing (data not shown).

Conclusions and Implications

There is some concern, on the part of OMB and others, that nonresponse to the two data sharing surveys may have introduced major error into the conclusions. The available evidence does not support that concern. There are few significant differences between the responses of those who initially refused and those who never refused. The differences that exist indicate that initial refusers are somewhat less knowledgeable about the census and somewhat more concerned about privacy,

suggesting that if the response rate had been higher, a larger percentage might have opposed data sharing, since both of these characteristics are predictive of opposition to sharing. But the differences are very small, the two characteristics predict only a small part of the variance in attitudes toward data sharing, and refusers make up no more than half of the total nonresponse. Moreover, there is little reason to suppose that nonresponse due to noncontact is related to attitudes toward data sharing. Thus, assuming final refusers were not dramatically different from converted refusers with respect to privacy attitudes, it seems unlikely that the implications for policy would change very much if nonresponse error were considerably reduced. That assumption does not seem an unreasonable one to make.

What are the implications for policy? Some of them can be readily inferred from the data. On the one hand, almost three quarters of the public currently favors data sharing if this would reduce the census undercount. This proportion has remained stable over at least a year, and in 1996 two thirds said they would be willing to provide their SSN to the Bureau to facilitate such sharing. On the other hand, data sharing to reduce or eliminate the burden of the long form garners a much smaller majority in favor, suggesting that the routine sharing of data among agencies might evoke greater opposition. Furthermore, opposition to data sharing, and to making the SSN available, is significantly related to privacy concerns, and such concerns show a small but significant increase between 1995 and 1996. Thus, it is possible that if privacy concerns continue to increase, they may erode the support for data sharing that currently exists. The same implication can be drawn from our findings concerning belief in the Census Bureau's assurance of confidentiality. Information about the law is apparently not enough; trust is also required. And the latter is a much more difficult message to communicate effectively.

The fact that a large majority of the public regards the sharing of identifiable data as an acceptable practice for reducing the census undercount does not imply that individuals should not be notified of the possibility that their data will be shared, or given an opportunity to opt out of this practice. The burden should be on the statistical agencies to convince their clients, the public, that in any particular case this practice is defensible and will result in a worthwhile outcome that could not be

achieved otherwise, or achieved only less well or at greater expense. In the case of voluntary surveys, individuals who are not persuaded should have the right not to have their data shared. Otherwise, such data collections become mandatory rather than voluntary. The decennial census, of course, is mandatory, and thus some would argue that it has no need to obtain permission. Yet even in this instance it is important for policy makers to weigh the benefits of data sharing against the infringement of the right to consent and, by extension, the right to refuse. Not only is the right to refuse a fundamental social value, but abrogating it may further reduce the trust on which all data collection activities, including the census, depend.

References

- Blair, Johnny. 1994. "Ancillary Uses of Government Administrative Data on Individuals: Public Perceptions and Attitudes." College Park: University of Maryland Survey Research Center.
- House of Commons Committee on Human Rights and the Status of Persons with Disabilities. 1997. "Privacy: Where Do We Draw the Line?" Report of the Committee. Ottawa, Canada.
- Hyman, Herbert H., Charles Wright, and John Reed. 1975. *The Enduring Effects of Education*. Chicago: University of Chicago Press.
- Kerwin, Jeffrey and Sherman Edwards. "The 1996 Survey on Privacy and Administrative Records Use." Rockville, MD., Westat.
- Singer, Eleanor and Stanley Presser. 1996. "Public Attitudes toward Data Sharing by Federal Agencies." Paper presented at the Annual Research Conference, Bureau of the Census.
- Singer, Eleanor, Nora Cate Schaeffer, and Trivelloro Raghunathan. 1997. "Public Attitudes toward Data Sharing." *Journal of International Public Opinion Research*, forthcoming.

Table 1
Predictors of Favoring Data Sharing by at Least
One Agency to Improve Count

	B	(Std. Error)	Sig	B	(Std. Error)	Sig
Intercept	.7059	(.7601)	.35	-.0321	(.8684)	.97
Importance	1.0328	(.1265)	.00	.9537	(.1321)	.00
Knowledge	-.0667	(.2450)	.79	-.1211	(.2692)	.65
Belief	-.3134	(.1343)	.02	-.2201	(.1423)	.12
Control	-.1175	(.1475)	.43	-.2142	(.1551)	.17
Distress	-.4767	(.1070)	.00	-.4901	(.1085)	.00
Trust	.2158	(.1331)	.11	.2546	(.1370)	.06
Paranoia	-.2251	(.2419)	.35	-.2227	(.2500)	.37
Age				.0058	(.0049)	.24
Education				.0563	(.0362)	.12
Income				.1062	(.0560)	.06
Income Miss				-.6243	(.1866)	.00
Gender				-.3169	(.1684)	.06
Nonwhite				-.4490	(.2237)	.04
Hispanic				.2500	(.3629)	.49
	N = 1401			N = 1391		

Table 2
Concerns about Privacy and Alienation from Government, by Year

	<i>% Agree Strongly or Somewhat</i>			
	<i>1995</i>		<i>1996</i>	
People's rights to privacy are well protected	41.4	(1413)	37.0	(1198)
People have lost all control over how personal information about them is used	79.5	(1398)	80.4	(1193)
		<i>(ns)</i>		
People like me don't have any say about what the government does	59.2	(1413)	62.9	(1200)
I don't think public officials care much what people like me think	65.4	(1414)	71.1	(1202)
How much do you trust the government in Washington to do what is right? (Almost never)	19.2	(1430)	25.0	(1204)

Table 3
Willingness to Provide SSN and Attitudes to Census Bureau

<i>Attitude/Opinion</i>	<i>Would Not Provide SSN</i>	<i>Would Provide SSN</i>
	<i>%</i>	<i>%</i>
Believes counting population is "extremely" or "very" important	63.8	79.7
Is aware of census uses	43.1	54.8
Would favor SSA giving Census Bureau short-form information	56.3	85.0
Would favor IRS giving Census Bureau long-form information	30.4	61.2
Would favor "records-only" census	45.6	60.0
Trusts Bureau to not give out/keep con-fidential census responses	45.0	76.7
Would be bothered "a lot" if other agency got census responses	54.1	29.9
Believes benefits of record sharing outweigh privacy loss	36.0	51.1
Believes the five items on short form are invasion of privacy	31.3	13.4

Table 4
Willingness to Provide SSN, by Concerns about Privacy
and Alienation from Government

<i>Concern/Alienation</i>	<i>Would Provide SSN</i>	<i>Would Not Provide SSN</i>
	<i>%</i>	<i>%</i>
Disagrees strongly that rights to privacy are well protected	24.2	45.6
Agrees strongly people have lost control over personal information	37.9	54.2
Agrees strongly “people like me” have no say about what government does	27.7	43.7
Agrees strongly public officials don’t care much about “what people like me think”	31.2	45.4
Almost never trusts government in Washington to do what’s right	19.5	37.8
Privacy loss outweighs economic benefit of data sharing	47.1	56.0
Economic benefit of data sharing outweighs privacy loss	47.9	30.4

Table 5
The Effect of Question Wording on Knowledge of
Laws Regarding Sharing of Census Information

	Is the Census Bureau forbidden by law from giving other government agencies census information identified by name or address?	Is the Census Bureau required by law to keep census information confidential?	Total
	<i>%</i>	<i>%</i>	<i>%</i>
Yes	28.3	51.1	40.2
No	17.1	11.6	14.2
Dont’t Know	54.6	37.3	45.5
<i>N</i> (unweighted)	591	624	1215