H.R. 5215, CONFIDENTIAL INFORMATION PROTECTION AND STATISTICAL EFFICIENCY ACT OF 2002

HEARING

BEFORE THE

SUBCOMMITTEE ON GOVERNMENT EFFICIENCY, FINANCIAL MANAGEMENT AND INTERGOVERNMENTAL RELATIONS

OF THE

COMMITTEE ON GOVERNMENT REFORM HOUSE OF REPRESENTATIVES

ONE HUNDRED SEVENTH CONGRESS

SECOND SESSION

ON

H.R. 5215

TO PROTECT THE CONFIDENTIALITY OF INFORMATION ACQUIRED FROM THE PUBLIC FOR STATISTICAL PURPOSES, AND TO PERMIT THE EXCHANGE OF BUSINESS DATA AMONG DESIGNATED STATISTICAL AGENCIES FOR STATISTICAL PURPOSES ONLY

SEPTEMBER 17, 2002

Serial No. 107–227

Printed for the use of the Committee on Government Reform

Available via the World Wide Web: http://www.gpo.gov/congress/house
http://www.house.gov/reform

U.S. GOVERNMENT PRINTING OFFICE

88–326 PDF WASHINGTON : 2002

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

CONTENTS

H.R. 5215, CONFIDENTIAL INFORMATION PROTECTION AND STATISTICAL EFFICIENCY ACT OF 2002

TUESDAY, SEPTEMBER 17, 2002

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON GOVERNMENT EFFICIENCY, FINANCIAL MANAGEMENT AND INTERGOVERNMENTAL RELATIONS,
COMMITTEE ON GOVERNMENT REFORM,
Washington, DC.

The subcommittee met, pursuant to notice, at 1:32 p.m., in room 2154, Rayburn House Office Building, Hon. Stephen Horn (chairman of the subcommittee) presiding.

Present: Representative Horn.

Staff present: Bonnie Heald, staff director; Henry Wray, senior counsel; Dan Daly, counsel; Chris Barkley, clerk; David McMillen, minority professional staff member; Jean Gosa, minority clerk; and Earley Green, minority assistant clerk.

Mr. HORN. A quorum being present, the subcommittee on Government Efficiency, Financial Management, and Intergovernmental Relations will come to order.

Today, the subcommittee will consider a bill which I introduced on behalf of myself and Representatives Tom Sawyer of Ohio and Carolyn Maloney of New York. The bill is H.R. 5125, the Confidential Information Protection and Statistical Efficiency Act of 2002. H.R. 5215 has primary objectives. One objective is to enable the Federal Government's three principle statistical agencies, the Bureau of the Census, the Bureau of Labor Statistics, and the Bureau of Economic Analysis, to share the business data they collect. Such data sharing would substantially enhance the accuracy of economic statistics by resolving serious inconsistencies that now exist. It would also reduce reporting burdens on businesses that must now supply data separately to the individual agencies.

The bill's second and equally important objective is to ensure that the confidential data that citizens and businesses provide to Federal agencies for statistical purposes are subject to uniform and rigorous statutory protections against their unauthorized use. Currently, confidentiality protections vary among agencies and are often not based in law. This bill would raise confidentiality standards for all Federal statistical agencies to the highest standard that now exists.

The administration strongly supports H.R. 5215. This bill is similar to another bill I had introduced in the 106th Congress, H.R. 2885, the Statistical Efficiency Act of 1999. That bill received

strong bipartisan support, and passed the House under suspension of the rules. H.R. 5215 differs from its predecessor by narrowing the data sharing provisions and broadening the confidentiality protections.

H.R. 5215 is a bipartisan, common-sense bill that we should enact this year. Therefore, the subcommittee will hold a markup on the bill immediately following the hearing.

The Heritage Foundation has raised a concern that the confidentiality provisions in H.R. 5215 could be misconstrued to prevent the release of some data that is now available to the non-governmental researchers. This information is released in a form that does not directly or indirectly reveal the identity of the data provider. That is not the intent of H.R. 5215. With the support of the administration, I will offer an amendment at the markup to clarify the language of the bill in this regard. My amendment also will strengthen the bills oversight provisions.

[The prepared statement of Hon. Stephen Horn and the text of H.R. 5215 follow:]

Opening Statement
Chairman Stephen Horn,
Subcommittee on Government Efficiency, Financial Management
and Intergovernmental Relations
September 17, 2002

A quorum being present, the Subcommittee on Government Efficiency, Financial Management and Intergovernmental Relations will come to order.

Today the subcommittee will consider a bill that I introduced on behalf of myself and Representatives Tom Sawyer and Carolyn Maloney. The bill is H.R. 5125, the "Confidential Information Protection and Statistical Efficiency Act of 2002."

H.R. 5215 has two primary objectives. One objective is to enable the federal government's three principal statistical agencies - the Bureau of the Census, the Bureau of Labor Statistics and the Bureau of Economic Analysis - to share the business data they collect. Such data sharing would substantially enhance the accuracy of economic statistics by resolving serious inconsistencies that now exist. It would also reduce reporting burdens on the businesses that must now supply data separately to the individual agencies.

The bill's second and equally important objective is to ensure that the confidential data that citizens and businesses provide to federal agencies for statistical purposes are subject to uniform and rigorous statutory protections against unauthorized use. Currently, confidentiality protections vary among agencies and are often not based in law. This bill would raise confidentiality standards for all federal statistical agencies to the highest standard that now exists.

The Administration strongly supports H.R. 5215. This bill is similar to another bill I introduced during the 106th Congress, H.R. 2885, the "Statistical Efficiency Act of 1999." That bill received strong bipartisan support and passed the House under suspension of the rules. H.R. 5215 differs from its predecessor by narrowing the data-sharing provisions and broadening the confidentiality protections.

H.R. 5215 is a bipartisan, common-sense bill that we should enact this year. Therefore, the subcommittee will hold a mark-up on the bill immediately following the hearing.

The Heritage Foundation has raised a concern that the language of the confidentiality provisions in H.R. 5215 could be misconstrued to prevent the release of some data that is now available to non-government researchers. This information is released in a form that does not

directly or indirectly reveal the identity of the data provider. That is not the intent of H.R. 5215. With the support of the Administration, I will offer an amendment at the mark-up to clarify the language of the bill in this regard. My amendment also will strengthen the bill's oversight provisions.

I welcome all of our excellent witnesses and look forward to their testimony. I am particularly pleased that our colleague, Representative Tom Sawyer from Ohio, is here to start us off today. He has been an outstanding leader for many years in efforts to improve federal statistical activities.

I

107TH CONGRESS
2D SESSION

H. R. 5215

To protect the confidentiality of information acquired from the public for statistical purposes, and to permit the exchange of business data among designated statistical agencies for statistical purposes only.

IN THE HOUSE OF REPRESENTATIVES

JULY 25, 2002

Mr. HORN (for himself, Mr. SAWYER, and Mrs. MALONEY of New York) introduced the following bill; which was referred to the Committee on Government Reform

---

A BILL

To protect the confidentiality of information acquired from the public for statistical purposes, and to permit the exchange of business data among designated statistical agencies for statistical purposes only.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "Confidential Information Protection and Statistical Efficiency Act of 2002".

SEC. 2. DEFINITIONS.

As used in this Act:

(1) The term "agency" means any entity that falls within the definition of the term "executive agency" as defined in section 102 of title 31, United States Code, or "agency", as defined in section 3502 of title 44, United States Code.

(2) The term "agent"—

(A) means an employee of a private organization or a researcher affiliated with an institution of higher learning (including a person granted special sworn status by the Bureau of the Census under section 23(c) of title 13, United States Code) with whom a contract or other agreement is executed, on a temporary basis, by an executive agency to perform exclusively statistical activities under the control and supervision of an officer or employee of that agency; or

(B) means an individual who is working under the authority of a government entity with which a contract or other agreement is executed by an executive agency to perform exclusively statistical activities under the control of an officer or employee of that agency; or

(C) means an individual who is a self-employed researcher, a consultant, or a contractor,

or who is an employee of a contractor and with whom a contract or other agreement is executed by an executive agency to perform a statistical activity under the control of an officer or employee of that agency; or

(D) means an individual who is a contractor or who is an employee of a contractor engaged by the agency to design or maintain the systems for handling or storage of data received under this Act; and

(E) who agrees in writing to comply with all provisions of law that affect information acquired by that agency.

(3) The term "business data" means operating and financial data and information about businesses, tax-exempt organizations, and government entities.

(4) The term "identifiable form" means any representation of information that permits information concerning a specific respondent to be reasonably inferred by either direct or indirect means.

(5) The term "nonstatistical purpose"—

(A) means the use of data in identifiable form for any purpose that is not a statistical purpose, including any administrative, regulatory, law enforcement, adjudicatory, or other

purpose that affects the rights, privileges, or benefits of a particular identifiable respondent; and

(B) includes the disclosure under section 552 of title 5, United States Code (the Freedom of Information Act) of data that are acquired for exclusively statistical purposes under a pledge of confidentiality.

(6) The term "respondent" means a person who, or organization that, is requested or required to supply information to an agency, is the subject of information requested or required to be supplied to an agency, or provides that information to an agency.

(7) The term "statistical activities"—

(A) means the collection, compilation, processing, or analysis of data for the purpose of describing or making estimates concerning the whole, or relevant groups or components within, economy, society, or natural environment; and

(B) includes the development of methods or resources that support those activities, such as measurement methods, models, statistical classifications, or sampling frames.

(8) The term "statistical agency or unit" means an agency or organizational unit of the executive branch whose activities are predominantly the collection, compilation, processing, or analysis of information for statistical purposes.

(9) The term "statistical purpose"—

(A) means the description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups; and

(B) includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support the purposes described in subparagraph (A).

SEC. 3. COORDINATION AND OVERSIGHT OF POLICIES.

(a) IN GENERAL.—The Director of the Office of Management and Budget shall coordinate and oversee the confidentiality and disclosure policies established by this Act.

(b) REVIEW AND APPROVAL OF RULES.—The Director shall review any rules proposed by an agency pursuant to this Act for consistency with the provisions of this Act and chapter 35 of title 44, United States Code, and such rules shall be subject to the approval of the Director.

SEC. 4. EFFECT ON OTHER LAWS.

(a) SECTION 3510 OF TITLE 44, UNITED STATES CODE.—This Act, including amendments made by this Act, does not diminish the authority under section 3510 of title 44, United States Code, of the Director of the Office of Management and Budget to direct, and of an agency to make, disclosures that are not inconsistent with any applicable law.

(b) SECTIONS 8, 16, 301, AND 401 OF TITLE 13 AND SECTION 2108 OF TITLE 44, UNITED STATES CODE.—This Act, including amendments made by this Act, does not diminish the authority of the Bureau of the Census to provide information in accordance with sections 8, 16, 301, and 401 of title 13 and section 2108 of title 44, United States Code.

(c) SECTION 9 OF TITLE 13, UNITED STATES CODE.—This Act, including amendments made by this Act, shall not be construed as authorizing the disclosure for nonstatistical purposes of demographic data or information collected by the Census Bureau pursuant to section 9 of title 13, United States Code.

(d) SECTION 12 OF THE FEDERAL ENERGY ADMINISTRATION ACT OF 1974.—In accordance with the provisions of this Act, data acquired for exclusively statistical purposes under a pledge of confidentiality are exempt from mandatory disclosure in identifiable form for nonstatis-

tical purposes under section 12 of the Federal Energy Administration Act of 1974 (15 United States Code 771).

(e) PREEMPTION OF STATE LAW.—Nothing in this Act shall preempt applicable State law regarding the confidentiality of data collected by the States.

(f) STATUTES REGARDING FALSE STATEMENTS.—Notwithstanding section 102, information collected by an agency for exclusively statistical purposes under a pledge of confidentiality may be provided by the collecting agency to a law enforcement agency for the prosecution of submissions to the collecting agency of false statistical information under statutes that authorize criminal penalties (such as section 221 of title 13, United States Code) or civil penalties for the provision of false statistical information, unless such disclosure or use would otherwise be prohibited under Federal law.

(g) CONSTRUCTION.—Nothing in this Act shall be construed as restricting or diminishing any confidentiality protections or penalties for unauthorized disclosure that otherwise apply to data or information collected for statistical purposes or nonstatistical purposes, including, but not limited to, section 6103 of title 26, United States Code.

TITLE I—CONFIDENTIAL INFORMATION PROTECTION

SEC. 101. FINDINGS AND PURPOSES.

(a) FINDINGS.—Congress finds the following:

(1) Individuals, businesses, and other organizations have varying degrees of legal protection when providing information to the Federal Government for strictly statistical purposes.

(2) Pledges of confidentiality by the Federal Government provide assurances to the public that information about individuals or organizations or provided by individuals or organizations for exclusively statistical purposes will be held in confidence and will not be used against such individuals or organizations in any Federal Government action.

(3) Protecting the confidentiality interests of individuals or organizations who provide information for Federal statistical programs serves both the interests of the public and the needs of society.

(4) Declining trust of the public in the protection of information provided to the Federal Government adversely affects both the accuracy and completeness of statistical analyses.

(5) Ensuring that information provided for statistical purposes receives protection is essential in

continuing public cooperation in statistical programs.

(b) PURPOSES.—The purposes of this title are the following:

(1) To ensure that information supplied by individuals or organizations to an agency for statistical purposes under a pledge of confidentiality is used exclusively for statistical purposes.

(2) To ensure that individuals or organizations who supply information to the Federal Government for statistical purposes will neither have that information disclosed in identifiable form to anyone not authorized by this Act nor have that information used for any purpose other than a statistical purpose.

(3) To safeguard the confidentiality of individually identifiable information acquired under a pledge of confidentiality for statistical purposes by controlling access to, and uses made of, such information.

SEC. 102. LIMITATIONS ON USE AND DISCLOSURE OF DATA AND INFORMATION.

(a) USE OF STATISTICAL DATA OR INFORMATION.—Data or information acquired by an agency under a pledge of confidentiality and for exclusively statistical purposes

shall be used by officers, employees, or agents of the agency exclusively for statistical purposes.

(b) DISCLOSURE OF STATISTICAL DATA OR INFORMATION.—

(1) Data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes shall not be disclosed by an agency in identifiable form, for any use other than an exclusively statistical purpose, except with the informed consent of the respondent.

(2) A disclosure pursuant to subparagraph (b)(1) above is authorized only when the head of the agency approves such disclosure and the disclosure is not prohibited by any other law.

(3) This section does not restrict or diminish any confidentiality protections in law that otherwise apply to data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes.

(c) RULE FOR USE OF DATA OR INFORMATION FOR NONSTATISTICAL PURPOSES.—A statistical agency or unit shall clearly distinguish any data or information it collects for nonstatistical purposes (as authorized by law) by a rule that provides that the respondent supplying the data or information is fully informed, before the data or

information is collected, that the data or information could be used for nonstatistical purposes.

(d) DESIGNATION OF AGENTS.—A statistical agency or unit may designate agents, by contract or by entering into a special agreement containing the provisions required by section 2, who may perform exclusively statistical activities, subject to the limitations and penalties described in this Act.

SEC. 103. FINES AND PENALTIES.

Whoever, being an officer, employee, or agent of an agency acquiring information for exclusively statistical purposes, having taken and subscribed the oath of office, or having sworn to observe the limitations imposed by section 102 of this title, comes into possession of such information by reason of his being an officer, employee, or agent and, knowing that the disclosure of the specific information is prohibited under the provisions of this Act, willfully discloses the information in any manner to a person or agency not entitled to receive it, shall be guilty of a class E felony and imprisoned for not more than 5 years, or fined not more than $250,000, or both.

TITLE II—STATISTICAL EFFICIENCY

SEC. 201. FINDINGS AND PURPOSES.

(a) FINDINGS.—Congress finds the following:

(1) Federal statistics are an important source of information for public and private decision-makers such as policymakers, consumers, businesses, investors, and workers.

(2) Federal statistical agencies should continuously seek to improve their efficiency. Statutory constraints limit the ability of these agencies to share data and thus to achieve higher efficiency for Federal statistical programs.

(3) The quality of Federal statistics depends on the willingness of businesses to respond to statistical surveys. Reducing reporting burdens will increase response rates, and therefore lead to more accurate characterizations of the economy.

(4) Enhanced sharing of business data among the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics for exclusively statistical purposes will improve their ability to track more accurately the large and rapidly changing nature of U.S. business. In particular, the statistical agencies will be able to better ensure that businesses are consistently classified in appropriate industries, resolve data anomalies, produce statistical samples that are consistently adjusted for the entry and exit of new businesses in a timely manner,

and correct faulty reporting errors quickly and efficiently.

(5) Congress passed the International Investment and Trade in Services Act of 1990 that allowed the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics to share data on foreign-owned companies. The Act not only expanded detailed industry coverage from 135 industries to over 800 industries with no increase in the data collected from respondents but also demonstrated how data sharing can result in the creation of valuable data products.

(6) With title I of this Act, the sharing of business data among the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics continues to ensure the highest level of confidentiality for respondents to statistical surveys.

(b) PURPOSES.—The purposes of this title are the following:

(1) To authorize the sharing of business data among the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics for exclusively statistical purposes.

(2) To reduce the paperwork burdens imposed on businesses that provide requested information to the Federal Government.

(3) To improve the comparability and accuracy of Federal economic statistics by allowing the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics to update sample frames, develop consistent classifications of establishments and companies into industries, improve coverage, and reconcile significant differences in data produced by the three agencies.

(4) To increase understanding of the United States economy, especially for key industry and regional statistics, to develop more accurate measures of the impact of technology on productivity growth, and to enhance the reliability of the Nation's most important economic indicators, such as the National Income and Product Accounts.

SEC. 202. DESIGNATION OF STATISTICAL AGENCIES.

For purposes of this title, the following shall hereafter be referred to as Designated Statistical Agencies:

(1) The Bureau of the Census in the Department of Commerce,

(2) The Bureau of Economic Analysis in the Department of Commerce, and

(3) The Bureau of Labor Statistics in the Department of Labor.

SEC. 203. RESPONSIBILITIES OF DESIGNATED STATISTICAL AGENCIES.

Designated Statistical Agencies shall:

(1) identify opportunities to eliminate duplication and otherwise reduce reporting burden and cost imposed on the public in providing information for statistical purposes;

(2) enter into joint statistical projects to improve the quality and reduce the cost of statistical programs; and

(3) protect the confidentiality of individually identifiable information acquired for statistical purposes by adhering to safeguard principles, including—

(A) emphasizing to their officers, employees, and agents the importance of protecting the confidentiality of individually identifiable information,

(B) training their officers, employees, and agents in their legal obligations to protect the confidentiality of individually identifiable information and in the procedures that must be followed to provide access to such information,

(C) implementing appropriate measures to assure the physical and electronic security of confidential data,

(D) establishing a system of records that identifies individuals accessing confidential data and the project for which the data were required, and

(E) being prepared to document their compliance with safeguard principles to other agencies authorized by law to monitor such compliance.

SEC. 204. SHARING OF BUSINESS DATA AMONG DESIGNATED STATISTICAL AGENCIES.

(a) A Designated Statistical Agency may provide business data in an identifiable form to another Designated Statistical Agency under the terms of a written agreement among the agencies sharing the business data that specifies—

(1) the business data to be shared;

(2) the statistical purposes for which the business data are to be used;

(3) the officers, employees, and agents authorized to examine the business data to be shared; and

(4) appropriate security procedures to safeguard the confidentiality of the business data.

(b) The provision of business data by an agency to a Designated Statistical Agency under this title shall in no way alter the responsibility of the agency providing the data under other statutes (including the Freedom of Information Act and the Privacy Act) with respect to the provision or withholding of such information by the agency providing the data.

(c) Examination of business data in identifiable form shall be limited to the officers, employees, and agents authorized to examine the individual reports in accordance with written agreements pursuant to this section. Officers, employees, and agents of a Designated Statistical Agency who receive data pursuant to this title shall be subject to all provisions of law, including penalties, that relate:

(1) to the unlawful provision of the business data that would apply to the officers, employees, and agents of the agency that originally obtained the information; and

(2) to the unlawful disclosure of the business data that would apply to officers, employees, and agents of the agency that originally obtained the information.

(d) NOTICE.—Whenever a written agreement concerns data that respondents were required by law to report and the respondents were not informed that the data could

be shared among the Designated Statistical Agencies, for exclusively statistical purposes, the terms of such agreement shall be described in a public notice issued by the agency that intends to provide the data. Such notice shall allow a minimum of 60 days for public comment.

SEC. 205. LIMITATIONS ON USE OF BUSINESS DATA PROVIDED BY DESIGNATED STATISTICAL AGENCIES.

(a) Business data provided by a Designated Statistical Agency pursuant to this title shall be used exclusively for statistical purposes as defined by this Act.

(b) Publication of business data acquired by a Designated Statistical Agency shall occur in a manner whereby the data furnished by any particular respondent are not in identifiable form.

SEC. 206. CONFORMING CHANGES IN LAW.

(a) DEPARTMENT OF COMMERCE.—Section 1 of the Act of January 27, 1938 (15 United States Code 176a) is amended by—

(1) striking "The" and inserting "Except as provided in the Confidential Information Protection and Statistical Efficiency Act of 2002, the";

(2) chapter 10 of title 13, United States Code, is amended by adding after section 401 the following:

"SEC. 402. PROVIDING BUSINESS DATA TO DESIGNATED STATISTICAL AGENCIES.

"The Bureau of the Census may provide business data to the Bureau of Economic Analysis and the Bureau of Labor Statistics ('Designated Statistical Agencies') if such information is required for an authorized statistical purpose and the provision is the subject of a written agreement with that Designated Statistical Agency, or their successors, as defined in the Confidential Information Protection and Statistical Efficiency Act of 2002."; and

(3) the table of sections for chapter 10 of title 13, United States Code, is amended by adding after the item relating to section 401 the following:

"402. Providing business data from the Bureau of the Census to Designated Statistical Agencies.".

○

Mr. HORN. I welcome all of our excellent witnesses, and look forward to their testimony. I am particularly pleased that our colleague, Representative Tom Sawyer from Ohio, he is a very distinguished legislator and we are sorry to see him leave after this Congress. He has been an outstanding leader for many years in many efforts to improve in this issue, to improve the Federal statistical activities. And if Mr. Sawyer would like to come forward and make his statement, and if you wish to, after your statement, come to the dais. And you are welcome to answer questions and all the rest of it. So, Tom, we are looking forward to it.

STATEMENT OF HON. TOM SAWYER, A REPRESENTATIVE IN CONGRESS FROM THE STATE OF OHIO

Mr. SAWYER. Thank you very much, Mr. Chairman. And thank you for those very kind words. It has been a pleasure to work with you over all the years that we have worked here together, and on no topic more than the one that brings us here together today.

You are right, I have spent a lot of my time in Congress in working on ways to improve Federal statistical systems so that the policies that we struggle over and the formulas that we agonize over have some meaningful life when they go forward in the real world and distribute dollars and lead to changed policies.

I am really pleased that you are moving forward on this legislation today. As you point out, this measure has been several years in the making. It builds on your approach to provide limited data sharing among agencies for the efficiency of the work of those agencies and my bill to strengthen the confidentiality of Government statistics. Together, these two approaches, I believe, will reduce statistical errors in many important arenas of Federal endeavor. Both parts of this bill are worthwhile, but I will focus my remarks on confidentiality today.

In that regard, the bill provides a clear and consistent standard for the use of confidential statistical information, and prohibits its use for any non-statistical purpose. It ensures that data gathered under a pledge of confidentiality are used only for statistical purposes, and imposes penalties for the willful disclosure of confidential information. It would replace the current patchwork of rules, and extend those protections to all individually identifiable data collected for statistical purpose, and in that way, encourage greater public cooperation with Government surveys and improve the quality of Federal statistics.

This measure, Mr. Chairman, I think is both timely and necessary. Confidence in Government data-gathering is fragile at best. In the 2000 Census, more than a quarter of the occupied housing units in this country did not respond to the mailed questionnaire. And while that was an improvement over the previous Census, it is an enormously difficult problem to overcome. There are many reasons for this. One is the growing unease about confidentiality that has grown with the unprecedented surge in the immigrant population. While there are no known cases of Federal agencies misusing such personal information in exactly this way, the risk is real. The temptation to diminish civil liberties in the name of national security leaves ordinary people and businesses feeling vulnerable to disclosure.

We saw that risk in the post-Pearl Harbor anti-Japanese tide. At that time, the Census Bureau worked closely with the War Department to help locate Japanese Americans. They didn't turn over actual records or violate the less stringent standards of the time, but the appearance of complicity acknowledged only 2 years ago shook the confidence of many immigrant communities, especially Asian American communities. Now, after September 11, the firewalls between individually identifiable information and aggregate statistics appear to be at risk.

I am pleased that the House Homeland Security bill made clear that it would not alter existing law on information collected for statistical purposes. However, in too many cases, existing law does not ensure that such personal information will remain confidential. More than 70 Federal agencies or statistical units collect such data, but only 12, as you suggest, Mr. Chairman, are covered by regulations to protect personally identifiable information from disclosure, and only a handful of those have the stronger protection of law.

Some of these uncovered units collect information on highly sensitive topics ranging from health care and substance abuse and mental health. It involves millions of dollars of sensitive data, and deserves the most stringent of protections from disclosure.

While agency policy may have once been enough, real public trust requires that information be shielded by the force of law. Statutory protection under H.R. 5215 would prevent any regulatory or law enforcement misuse of these data. This recommendation was first made under the Privacy Act of 1944. However, that act has several loopholes that allow for the disclosure of personally identifiable information without the informed consent of those who supplied the information.

There are 12 categories of such exemptions, and the act fails to distinguish between data collected for research purposes and that collected for administrative purposes, and so offers minimal protection from improper disclosure.

The commission at that time that arose from the Privacy Act recommended that no record or information collected for statistical purpose be used in identifiable form to make any decision or take any action directly affecting the person to whom the record pertains. H.R. 5215 embodies the commission's recommendation in that regard.

In summary, these improvements, Mr. Chairman, are long overdue. They are needed to protect the public and ensure continued public participation in essential governmental research. Informed public policy relies on it. Thank you, Mr. Chairman, for the chance to be here.

Mr. HORN. Thank you. And if you wish to come up here, why, without objection we are delighted to have you with us.

Mr. SAWYER. Thank you, Mr. Chairman.

[The prepared statement of Hon. Tom Sawyer follows:]

Congressman Tom Sawyer
9/17/02
Testimony: H.R. 5215

Thank you Chairman Horn and Ranking Member Schakowsky for holding this hearing today. I appreciate the opportunity to testify on this important legislation. As a former Chairman of the Subcommittee on Census, Statistics and Postal Personnel, I have spent much of my Congressional career working on ways to improve the government's statistical capabilities, and I am pleased that this committee is moving forward on this legislation. I am the lead Democratic sponsor of this bill and would like thank Chairman Horn for the opportunity to work with him to introduce this bill and for his leadership on this issue.

H.R. 5215, the Confidential Information Protection and Statistical Efficiency Act of 2002 has been years in the making. This bill builds upon legislation[1] introduced by Chairman Horn last Congress to provide limited data sharing among agencies as well as legislation[2] I introduced last year to strengthen the confidentiality of government statistics. My remarks today will focus on the

[1] H.R. 2885, the Statistical Efficiency Act of 1999.

[2]H.R. 2136, the Confidential Information Protection Act.

confidentiality provisions of the current legislation. I am confident that witnesses later today will testify on how H.R. 5215 will reduce statistical errors in estimating the Gross Domestic Product and in compiling industrial output, productivity and employment statistics.

The bill:

- Provides a clear and consistent standard for the use of confidential statistical information, and prohibits the federal government from using such information for any non-statistical purpose;
- Provides safeguards to ensure that data provided under a pledge of confidentiality are used only for statistical purposes; and
- Imposes criminal penalties on Federal employees who willfully disclose confidential information.

In sum, H.R. 5215 would create a uniform set of protections for statistical information collected by the government under a pledge of confidentiality. It would replace the current patchwork of rules and extend these protections to all individually identifiable data collected for statistical purposes. These enhanced

confidentiality protections will encourage greater participation and cooperation among the public with government surveys, which will in turn improve the quality of federal statistics.

This legislation is both timely and necessary. Confidence in government data gathering is fragile at best. In the 2000 census, more than a quarter of occupied housing units did not respond to the mailed questionnaire. There are many reasons people do not answer the census and other federal surveys, or do so only reluctantly. One is the growing unease about the confidentiality of personal information collected for statistical purposes, a concern that has grown with the unprecedented surge in the immigrant population. To my knowledge, there are no cases of federal agencies disclosing or misusing personal information collected exclusively for statistical uses. But the risk of misuse is real. The temptation to diminish civil liberties in the name of national security can be high. It leaves just the sort of personal and sensitive information ordinary people and businesses provide to the nation's statistical entities vulnerable to disclosure.

We saw that risk 60 years ago, in the wake of the post-Pearl Harbor anti-Japanese tide. The Census Bureau, in a misguided attempt to contribute to the war

effort, worked closely with the War Department over a significant period of time to identify the location of Japanese Americans. The bureau did not turn over actual records collected in the 1940 census, nor did it violate what were then far less stringent legal standards for data confidentiality. Nevertheless, the *appearance* of complicity – acknowledged publicly by the Census Bureau only two years ago in light of new research into the incident – shook confidence in the entire census process, particularly in Asian American and other immigrant communities.

Now, in the aftermath of the September 11th attacks, the resilience of the firewalls between individually identifiable information and aggregate statistics could be tested once again. I was pleased that the House legislation creating the Homeland Security Department included an amendment by Representative Chambliss making it clear that the creation of the new Department would not alter existing law regarding the confidentiality of information collected by the Federal Government solely for statistical purposes.

However, in many instances, existing law does not ensure that personal information collected by the government for statistical purposes will remain confidential. More than 70 federal agencies or statistical units have a role in

collecting data from individuals and businesses to support statistical activities. Yet, only 12 are covered by government regulations that seek to protect personal, identifiable information from disclosure. Only a handful of those have the stronger protection of the law.

Some of the agencies or statistical units that are not covered by regulation or a specific law often collect information on personal or sensitive topics, including the Agency for Health Care Research and Quality, the National Institutes of Health, and the Substance Abuse and Mental Health Services Administration. These agencies conduct hundreds of millions of dollars worth of statistical work each year. Such sensitive data deserves the most stringent of protections from disclosure. While pledges of confidentiality or administrative policy may have been sufficient in the past, the public should know that information they provide is shielded by the force of law. Only through such measures will we gain the public's utmost trust.

By extending statutory protection to all information collected for statistical purposes under a pledge of confidentiality, H.R. 5215 would prevent any misuse of these data through a regulatory or law enforcement action. This was a

recommendation made back in the late 1970s by the Privacy Protection Study Commission, which was created by the Privacy Act of 1974. The 1974 Act was the first attempt by Congress to provide comprehensive protection of personally identifiable information collected by government agencies.

However, the Privacy Act has several loopholes that allow for disclosure of personally identifiable information without the informed consent of the business or individual who supplied the information. There are twelve categories of exemptions that allow an agency to disclose personally identifiable information without informed consent. Furthermore, the Act fails to distinguish between data collected for research purposes and data collected for administrative purposes. Consequently, the law offers minimal protection from improper disclosure of statistical information for non-research purposes.

The commission recommended that no record or information collected for statistical purpose be used in identifiable form to make any decision or take any action directly affecting the person to whom the record pertains. H.R. 5215 embodies the commission's recommendation.

Improvements that the bill would make in our nation's statistical programs are long overdue. The measures are needed not only to protect the public but also to ensure the public's continued cooperation and participation in essential government research. Informed public policy relies on it.

Mr. HORN. We will now go to the panel two, and that is the Honorable Randall S. Kroszner, member, Council of Economic Advisers, Executive Office of the President; the Honorable Kathleen B. Cooper, Under Secretary for Economic Affairs, Department of Commerce; the Honorable Kathleen P. Utgoff, Commissioner of the Bureau of Labor Statistics, Department of Labor. And since it is a big table, and three more—we will throw in panel three, if you would like to come up to the table and just grab one of the chairs if we are missing them. And Maurine Haver, doctorate, Chair, statistics committee, representing the National Association for Business Economics; William D. Nordhaus, Sterling professor of economics, Department of Economics at Yale; Dr. Ralph Rector is the research fellow and project manager, Center for Data Analysis, the Heritage Foundation.

And we have to swear witnesses. If you will stand and raise your right hand. And if you have anybody supporting you there, get them behind you, too, so we don't have to do it halfway through the panel.

[Witnesses sworn.]

Mr. HORN. And the clerk will note that all six affirmed. And we will start then with what it is on the agenda, and that is the Honorable Randall Kroszner, member, Council of Economic Advisers. Those were people that were authorized by law under President Truman; and the old humor that the President had, he said he was tired of somebody saying on this bit and that bit and so forth, I want some people that can give me some good economic data and not just tell it on the right, he is saying, and then the left hand. And this is the Council of Economic Advisers.

STATEMENTS OF RANDALL S. KROSZNER, MEMBER, COUNCIL OF ECONOMIC ADVISERS, EXECUTIVE OFFICE OF THE PRESIDENT; KATHLEEN B. COOPER, UNDER SECRETARY FOR ECONOMIC AFFAIRS, DEPARTMENT OF COMMERCE; AND KATHLEEN P. UTGOFF, COMMISSIONER, BUREAU OF LABOR STATISTICS, DEPARTMENT OF LABOR

Mr. KROSZNER. Thank you very much, Mr. Chairman, and members of the subcommittee. And I am very, very pleased to be speaking before you today on what I consider an extremely important initiative that has very big implications, but one that has no budget implications. And so what I am going to argue is that this is a very, very valuable piece of legislation that does not take any additional costs on the—for the public sector, and actually can reduce burdens on the private sector. And of course, since I am on leave from the University of Chicago, our main theme there is, "there ain't no such thing as a free lunch."

So, why hasn't this happened before? Well, members of the staffs of all of the agencies that have been involved as well as the leaders of those can tell you that this was not a free lunch. It required a lot of work to make sure that we could get the language correct, to make sure that we had the appropriate ideas in there, appropriate scope. And we got just fabulous support and cooperation among the different agencies and with Capitol Hill.

So I am extremely pleased to be able to speak to you about this very important issue.

As we well know, Federal statistics in the United States are among the best, if not the best, in the world. But that should not make us complacent. We can still improve them. The U.S. economy is an extremely dynamic one, and one in which we must respond to changes, constant changes. And that requires constant improvement in our statistics. As we well know, Government statistics play an important role not only for Government decisions, whether they be Social Security decisions, budgetary decisions, monetary policy decisions, small errors can have very large effects when we do our budget projections. And this is something that is extremely important to the Council of Economic Advisers in providing information to the President to be able to provide accurate forecasts.

The private sector, of course, relies very heavily on statistics for their own budget and planning purposes, and, of course, academics around the world rely on U.S. statistics to be able to do an appropriate analysis of both the effects of policy as well as more fundamental research. And so what we need to do is to improve those statistics in a way that will help to reduce some of the data problems that we have seen recently.

We have had very large revisions of GDP. We have different estimates of productivity, one of the most important aspects of our economy going forward. Depending on which measure you use from the income and project accounts versus other accounts, the difference can be 35 basis points. Now, that is about a third of a percent. That doesn't seem like very much, but that is a third of a percent difference in growth every year going out into the future. That makes an enormous difference for our well-being over time, it makes an enormous difference to our budget projections. There is an approximate effect of about $200 billion for every point 1 percent difference in GDP growth over a 10-year budget horizon. So we are talking real money here.

Also, when we think about the classification of firms, a very limited study was done a few years ago looking at the differences between the Census and the Bureau of Labor statistics on how they classified firms. They found about a 30 percent difference in which industries firms are classified into. This has led to very big differences in the estimates for the size of particular sectors of the economy, the chemical sector, high-tech sector. And in particular, in the new sectors, in emerging sectors it becomes very difficult to try to classify these firms.

And that is how these data anomalies and these areas creep in. Even though I think all the agencies do a superb job, just trying to look over something as large as the U.S. economy is a very difficult thing, given the very limited budgets they have. By being able to talk with each other and say, well, we have classified this firm this way, but you've classified it that way, let's try to understand what would be the best classification, that is a very simple, straightforward thing that is basically impossible to do now. And if we can allow for that, we can improve our Government statistics and improve the numbers that we get out of that process and be much more accurate.

What this act would allow us to do is do that through two means, through formal data sharing through a variety of memoranda of understanding among the different agencies, but also by having

consistent high confidentiality protections amongst all of the agencies. And this is very valuable for allowing the data sharing to occur, because there has to be the same level of protection, and we want to make sure that it is a high level of protection if the data is going to move from one agency to another. And so that becomes an extremely important part of the legislation itself.

And what this will also do is reduce the burdens on the private sector. If the agencies can coordinate with each other, eventually they may be able to reduce duplicative surveys, they can, in general, just reduce the burdens on the private sector.

And so I consider this something that is a triple win. First, by improving these statistics directly through improved business lists, we get better data from what is provided to the statistical agencies. But by boosting confidence and by lowering the burdens on the private sector, we are likely to get much more accurate data from the private sector so we also have more efficient Government, we have lower burdens on the private sector, we have no budget cost. I consider this a triple win.

I think this is something that should have bipartisan support going forward for everyone's agenda because it is something that I think clearly improves the Government, improves the private sector, has no budget costs, and I see very little—actually, absolutely no downside. Thank you very much, Mr. Chairman.

Mr. HORN. Thank you. We appreciate your fine succinct presentation, and we will wait and go with your colleagues and then we will start the question period.

[The prepared statement of Mr. Kroszner follows:]

Testimony
of
Randall Kroszner
Member, Council of Economic Advisers

before the
Subcommittee on Government Efficiency, Financial Management and Intergovernmental Relations
Committee on Government Reform
U.S. House Of Representatives

Tuesday, September 17, 2002
1:30 P.M.

Chairman Horn and members of the Subcommittee, it is a pleasure to appear before you to discuss the important issues of data sharing and statistical confidentiality. Enhanced data sharing will boost our understanding of the U.S. economy, especially in fast-growing regions and industries. This will lead to more accurate measures of the impact of technology on productivity growth, and improve the accuracy of the Nation's most important economic indicators, such as the Gross Domestic Product (GDP) and other key aggregates that are the cornerstones of budget and monetary policy. Enhanced confidentiality protection will strengthen the confidence that respondents place in government statistical organizations and their willingness to participate in government surveys. All of this would happen at no additional cost to the taxpayer.

The Administration strongly supports this legislation, and is grateful for your leadership on this issue. As you know, the Administration has consulted with the Congress extensively on the development of legislation to address these important goals. These collaborative efforts have produced a bill that is good, common sense government. Thank you for your efforts in developing this important piece of legislation.

The data-sharing initiative as embodied in H.R. 5215, the Confidential Information Protection and Statistical Efficiency Act, would authorize the sharing of business data among the Bureau of the Census (Census Bureau), the Bureau of Economic Analysis (BEA), and the Bureau of Labor Statistics (BLS). The ability to share data would improve the accuracy and reliability of economic statistics, and reduce the duplicative paperwork burdens imposed on businesses. H.R. 5215 also would establish a uniform set of statutory protections to ensure the confidentiality of all information acquired from the public under a pledge of confidentiality for exclusively statistical purposes; these protections would include tough criminal and civil penalties for inappropriate disclosure.

Let me discuss each aspect of the proposal in turn. The sharing of business data among the Census Bureau, BEA, and BLS will facilitate statistical projects that improve the quality and reduce the burden and cost of statistical programs. Only business data will be shared, for exclusively statistical purposes. Household and demographic data are not part of the data sharing proposal. Written agreements will specify the business data to be shared, the statistical purposes for which the business data are to be used, and the appropriate security procedures to safeguard the confidentiality of the business data.

What benefits accrue from allowing greater sharing of data? Enhanced data sharing will improve the ability of the Census Bureau, BEA, and BLS to track rapidly changing trends in the U.S. economy. The most important result of the proposal would be allowing these agencies to improve employment, price, and shipments data by better classifying establishments in appropriate industries. A limited research study compared the Census Bureau's and BLS's business lists for 1994 and found that 30 percent of the same single-establishment firms had been assigned different industry codes at the 4-digit Standard Industrial Classification (SIC) level. Industry analyses that rely on employment or price survey data from BLS and shipments survey data from the Census Bureau may well provide unreliable characterizations of changes in real output and productivity for particular industries. Moreover, it is expected that this discrepancy will worsen as the economy changes.

The initiative would permit these agencies to increase accuracy and resolve data anomalies. At the end of 2001, the statistical discrepancy in the National Income and Product Accounts (the difference between income and output) was historically high, approaching two percent of GDP. The discrepancy distorts our readings of such critical indicators as output and productivity growth. And these are critical for budget policy: an error of just 0.1 percentage point in long-term real GDP growth can result in an error of approximately 200 billion dollars in a ten year budget forecast.

Sharing of business data would help provide for more accurate measures of industry output and compensation trends that could help reduce the statistical discrepancy. It would permit the statistical agencies to keep abreast of our dynamic economy by producing statistical samples that are consistently adjusted for the entry and exit of new businesses in a timely manner, and allowing the agencies to correct errors quickly and efficiently. This is especially important for fast-growing industries such as information technology. A comparison of BLS-based payroll data in the information technology sector was 13 percent higher than the payroll as reported by the Census Bureau in the last economic census year of 1997, and similar differences exist in other industries.

At present, Federal statistical agencies sometimes conduct separate collections of information from the public on similar subjects. This occurs when multiple agencies have a critical need for such information but are prohibited by law from sharing it. Data sharing can provide new opportunities for the agencies to coordinate their efforts by integrating or eliminating duplicative collections of information.

The second important aspect of the legislation will strengthen safeguards to protect confidential statistical information provided by the public. Individuals, businesses, and other organizations currently have varying degrees of statutory protection for confidential statistical information that they provide to the Federal Government. This legislation would apply clear and uniform statutory restrictions on the use of confidential statistical information. In particular, information about individuals or organizations acquired for exclusively statistical purposes and under a pledge of confidentiality could only be used for statistical purposes.

In addition to uniform restrictions, the legislation would provide consistent, tough penalties for unauthorized disclosure of confidential statistical information. In doing so, it would not diminish or restrict the applicability of harsher penalties that already exist in the law for unauthorized disclosure.

Finally, the two aspects of the proposal work together. Reducing reporting burdens and providing greater assurances of confidentiality will raise the likelihood that businesses will respond to surveys, and therefore lead to more accurate descriptions of the economy.

The Administration looks forward to working with the Congress on enactment of H.R. 5215 so that the American public can start benefiting from higher quality economic statistics for public and private decision-making, greater Federal Government efficiency, and increased protection of confidential statistical information. Thank you for the opportunity to be here today. I would be happy to answer your questions.

Mr. HORN. We now have the Honorable Kathleen Cooper, Under Secretary for Economic Affairs at the Department of Commerce.

Ms. COOPER. Thank you very much, Chairman Horn, Congressman Sawyer. I am very pleased to be here with you today to discuss H.R. 5215. I need to get the—better volume. Is that better? All right. I hope it is not too loud. But I also want to thank you, Congressman Horn—I mean, Chairman Horn and Congressman Sawyer, but certainly Chairman Horn, for your leadership on this issue for a good number of years. As Under Secretary for Economic Affairs of the Department of Commerce, I have the privilege of overseeing the fine work of two of the jewels of the Federal statistical system, the Bureau of Economic Analysis and the Census Bureau.

And our economic statistics from these Bureaus are calculated by experts and professionals who produce the most accurate numbers possible with the tools available to them. The President has made enhancing our economic data a priority and wants to give the Bureaus the tools they need to measure the twenty-first century economy. He appreciates that better information is fundamental to better public and private decisionmaking. With the President's budget request for the Census Bureau and BEA, you will see a range of economic data sooner. The release of international trade data, for example, will be available 20 days earlier. We plan to implement an annual measure of investment in information technology and quarterly measures of the services sector, the Census Bureau's first new economic indicator in 50 years.

Unfortunately, while these agencies are striving to improve economic statistics, the fiscal year 2003 funding level approved by Senate appropriators is significantly below the President's request, and we simply will not be able to undertake these important improvements with the Senate's flat funding level.

But today, I would like to discuss one way to improve our Federal statistics at next to no cost. If enacted, this legislation will help us improve the measurement of inventories, one of the most volatile components of GDP. We will develop more efficient samples, reduce the reporting burden, improve regional and State data, and reduce revisions. And I would like to share with you some real-life examples of gaps in our Federal statistical system the data-sharing bill would close.

Most of BEA's data comes from elsewhere, the Census Bureau and the Bureau of Labor Statistics being the main building blocks for BEA products. The Census Bureau, BLS, and BEA already work hand in hand. BEA is the Census Bureau's most important customer, and the two agencies are in contact every day. The staffs at BLS and the Census Bureau meet routinely with their counterparts at BEA. Throughout the year, managers collaborate and ensure that our statistical infrastructure is efficient and productive. However, H.R. 5215 would allow BEA, the Census Bureau, and BLS to work even more efficiently together, to share knowledge, and to borrow strengths. The most important result would be, as my colleague mentioned, consistent classification of businesses by the Census Bureau and BLS. BEA would be the first to stand and cheer such an accomplishment. The Census Bureau and BLS place

one third of the businesses in different boxes, and the BEA has to sort out the resulting data.

In determining real output, for example, BEA looks at shipments from Census and prices for BLS, and must untangle the classification confusion. With the data-sharing bill, the statistical agencies could cross-validate their company lists and determine the most appropriate classification. By comparing corporate financial reports with BLS and Census surveys, BEA can improve estimates of profits and of wages and salaries. This would help to reconcile the near $100 billion statistical discrepancy between gross domestic income and gross domestic product.

At the end of July, we saw large revisions to profits that indicated that corporate profits had peaked in 1997 rather than in 2000, and that other incomes were weaker than earlier estimated. If BEA were able to access the firm level data from the Census Bureau, if you compare that information with corporate return data from the IRS and publicly released financial reports and in this way BEA could better capture the impact of corporate activity where there is a difference between tax and financial accounting methods.

I would like to see an improvement in the accuracy of State personal income as well. For the year 2000, estimates of growth and payrolls for Delaware vary from a BLS estimate of 6 percent to Census Bureau's 14 percent estimate. Likewise, for Virginia, there are sizable differences. Even for New York where the differences in growth are smaller, the difference in dollars in over $7 billion in wages or $380 million in State and local taxes for the State.

Amid the dynamic economy, how can our statistical agencies keep track of businesses that come and go? The Census Bureau and BLS have different sources of information that provide insight into companies births and deaths. Combining the two measures should give us better information than from either source alone.

Other legislation under congressional consideration would allow the agencies limited access to IRS data. Today, Census and BEA have access to difference set office tax data. Enhanced access to IRS material would allow BEA to make its measure of corporate profits and other business income all the more accurate.

As you will hear and have heard in part already from my colleagues, my fellow panelists, this bill also builds on the agency's unmatched record of confidentiality. It provides equally stringent protection for all data, and avoids any perception of inappropriate use.

Finally, Mr. Chairman, I would like to alert you to a concern that I have about the most basic step in our data gathering, voluntary participation in our surveys. In April, reporting by large semi-conductor companies in the Census Bureau's monthly survey—in the Census Bureau's monthly survey of manufacturing activities dropped to the point where the Bureau had to discontinue publishing data on semi-conductors. As a result, the Census Bureau could no longer produce bellwether sales and inventory data for this very important industry.

The Census Bureau and the semi-conductor industry have agreed to a test to determine whether the industry can provide the desired data. The results will not be known for 5 months, and in the mean-

time, gaps in our data persist. We will continue to work with businesses to find efficient means for them to report. Ultimately, Congress and the administration must encourage participation that yields information vital to informed decisionmaking.

I thank you, Mr. Chairman. I appreciate your efforts to improve the quality and the efficiency of our Nation's statistics while protecting its confidentiality. Thank you.

Mr. HORN. Thank you. We are delighted to have you with us.

And we now will have the last part on panel one, and that is the Honorable Kathleen Utgoff, Commissioner of the Bureau of Labor Statistics.

During the Eisenhower administration, I was assistant to the Secretary of Labor, and the first thing he said to me was: Nobody around here fools with the commissioner of Bureau of Labor Statistics. And all of us were told he is untouchable. Now, you are untouchable. So.

Ms. UTGOFF. Thank you for that piece of history, Mr. Chairman. Good afternoon and—good afternoon, Mr. Chairman, Congressman Sawyer. I appreciate the opportunity to testify in support of H.R. 5215. Having come on board as BLS commissioner just last month, I am particularly pleased that my first chance to appear before Congress is an opportunity to support something that is so important to the BLS. There are three main reasons we believe the proposal currently before the subcommittee is a good one, and you will hear the other panelists give similar reasons.

First, all data users, researchers, analysts, policymakers, private citizens, Government agencies, corporate decisionmakers will benefit from a higher quality economic statistics. Second, the major Federal statistical agencies will be able to operate more efficiently. Third, the protections of confidential statistical information will be enhanced.

This carefully crafted bill is designed to meet all these aims. It has come about through an extraordinary level of cooperation among representatives from the Council of Economic Advisers, the Bureau of the Census, the Bureau of Economic Analysis, the BLS, the Office of Management and Budget, and the Internal Revenue Service, coupled with outstanding congressional collaboration. The bill is a testament to hard work and a shared belief in better Government.

Because BLS has an extensive network of cooperative arrangements with the State's statistical agencies to produce State and subcommittee—sublabor market estimates, I would like to note that the intended benefits from data sharing will also extend to the States. In addition to being large producers of their own economic data, State governments are large consumers of federally produced data, and therefore have an interest in Federal program improvement.

The enhanced data sharing that would be permitted under Title 2 will improve the ability of BLS, BEA, Census, and the States to track rapidly changing trends in the U.S. economy. It will facilitate joint projects among the agencies to improve data quality and to reduce the reporting burden and costs of programs. In particular, both the Federal and States' statistical agencies will be better able to classify business establishments in appropriate industries, they

will be able to resolve data anomalies and correct reporting errors more quickly and more efficiently.

Reconciling discrepancies between the BLS and the Census Bureau business registers is one critical example of why enhanced data sharing matters. The tangible benefits will include improved employment, unemployment, and income measures, better survey sampling frames, improved payroll data for forecasting State government revenues, and a better foundation for economic development plans. More accurate business classification will ultimately allow for the production of more accurate industry statistics, a vital part of our national and State economic intelligence picture.

Another important area of potential improvement to BLS data series is the measurement of multifactor productivity. These data series track the contributions of capital, technology, and labor to output. Productivity is widely regarded by analysts as a key ingredient of economic performance and the standards of living. By linking BLS work force and occupational data to Census Bureau production inputs and outputs data, better measures of productivity can be developed. These better measures allow more complete research into understanding the factors that cause productivity change particularly at the firm level.

Having access to data the Census Bureau collects on firm revenues by specific product lines would allow BLS to improve sampling and reduce respondent burden for the producer price index, or the PPI. The PPI is one of the Nation's most watched economic indicators. It measures price pressures at the wholesale level of the economy.

Turning now from the data-sharing provisions of the bill to the confidentiality provisions, I would like to emphasize the importance of Title I to Bureau of Labor Statistics. The BLS has long needed the explicit statutory confidentiality protection that this legislation offers for the data it collects. The fact that BLS lacks the specific comprehensive protections already in place with the Census Bureau, to cite one example, is a historical omission that needs correcting. BLS has been successful in protecting the confidentiality of the data it collects through an amalgam of statutes, precedents, rules, and practices. A clear example—a clear, explicit, and comprehensive statutory assurance of confidentiality is essential to maintain and improve our response rates by increasing respondent confidence in improving our ability to protect their data.

In fact, it is not an overstatement to say that confidentiality is the lifeblood for Government statistical agencies like BLS that depend upon the voluntary participation of survey respondents. We rely on individual citizens in their private households and on companies of all sizes in all industries and across all States to entrust us with their vital economic data. The voluntary transmittal of citizens to the Government of private information on, for example, job status or earnings, or by a company of its core business information such as employment wages and revenues is a remarkable example of public/private partnership. Without the trust of our survey respondents, the BLS surveys and the important statistics they gen-

erate would be in jeopardy. H.R. 5215 will strengthen that trust.

Thank you, Mr. Chairman. I would be happy to answer any questions that you have.

Mr. HORN. Thank you.

[The prepared statement of Ms. Utgoff follows:]

TESTIMONY OF KATHLEEN P. UTGOFF
COMMISSIONER OF LABOR STATISTICS

Before the

SUBCOMMITTEE ON GOVERNMENT EFFICIENCY, FINANCIAL MANAGEMENT AND INTERGOVERNMENTAL RELATIONS
HOUSE COMMITTEE ON GOVERNMENT REFORM

SEPTEMBER 17, 2002

Good afternoon, Mr. Chairman and Members of the Subcommittee. I appreciate the opportunity to testify in support of H.R. 5215, the Confidential Information Protection and Statistical Efficiency Act of 2002. Having come on board as BLS Commissioner just last month, I am particularly pleased that my first chance to appear before Congress is an opportunity to support something that is so important to the Bureau.

Title I of the proposed legislation would establish uniform statutory confidentiality protections for information collected by the Federal Government for exclusively statistical purposes under a pledge of confidentiality. Title II would facilitate the exchange of data collected from businesses by the Federal Government's principal economic statistics agencies.

There are three main reasons we believe the proposal currently before the Subcommittee is a good one, and you no doubt will hear these themes emphasized by the other panelists. First, all data users – researchers, analysts, policy-makers, private citizens, government agencies, corporate decision-makers – will benefit from higher quality economic statistics. Second, the major Federal statistical agencies will be able to operate more efficiently. Third, the protections of confidential statistical information will be enhanced.

This carefully crafted bill is designed to meet those aims. It has come about through an extraordinary level of cooperation among representatives from the Council of Economic Advisers, the Bureau of the Census, the Bureau of Economic Analysis (BEA), the BLS, the Office of Management and Budget, and the Internal Revenue Service, coupled with outstanding congressional collaboration. It is a testament to hard work and a shared belief in better government.

Because BLS has an extensive network of cooperative arrangements with state statistical agencies to produce State and sub-State labor market estimates, I would like to note that the intended benefits from data sharing also will extend to the States. In addition to being large producers of their own economic data, State governments are large consumers of federally-produced data, and therefore have an interest in Federal program improvement.

The enhanced data sharing that would be permitted under Title II will improve the ability of BLS, BEA, Census, and the States to track rapidly changing trends in the U.S. economy. It will facilitate joint projects among the agencies to improve data quality and reduce the reporting burden and cost of programs. In particular, both the Federal and State statistical agencies will be better able to classify business establishments consistently in appropriate industries, resolve data anomalies, and correct reporting errors more quickly and efficiently.

Reconciling discrepancies between the BLS and Census Bureau business registers is one critical example of why enhanced data sharing matters. The tangible benefits will include improved employment, unemployment and income measures, better survey sampling frames, improved payroll data for forecasting State government revenues, and a better foundation for economic development plans. More accurate business classification will ultimately allow for the production of more accurate industry statistics, a vital part of our national and state economic intelligence picture.

Another important area of potential improvement to BLS data series is in the measurement of multifactor productivity. These data series track the contributions of capital, technology, and labor to output. Productivity is widely regarded by analysts as a key ingredient of economic performance and standards of living. By linking BLS workforce and occupational data to Census Bureau production inputs and outputs data, better measures of productivity can be developed. These better measures allow more complete research into understanding the factors that cause productivity change, particularly at the firm level.

Having access to data the Census Bureau collects on firm revenues by specific product lines would allow BLS to improve sampling and reduce respondent burden for the Producer Price Index (PPI). The PPI is one of the Nation's most watched economic indicators. It measures price pressures at the wholesale level of the economy.

Turning now from the data-sharing provisions of the bill to the confidentiality provisions, I would like to emphasize the importance of Title I to the Bureau of Labor Statistics. The

BLS has long needed the explicit statutory confidentiality protection that this legislation offers for the data it collects. The fact that BLS lacks the specific comprehensive protections already in place for the Census Bureau, to cite one example, is an historical omission that needs correcting. BLS has been successful in protecting the confidentiality of the data it collects through an amalgam of statutes, precedents, rules, and practices. But, BLS gathers most of its data voluntarily and under a pledge of confidentiality. A clear, explicit, and comprehensive statutory assurance of confidentiality is essential to maintain and improve our response rates by increasing respondent confidence in our ability to protect their data.

In fact, it is not an overstatement to say that confidentiality is the lifeblood for government statistical agencies like BLS that depend upon the voluntary participation of survey respondents. We rely on individual citizens in their private households and on companies of all sizes, in all industries, and across all States, to entrust us with their vital economic data. The voluntary transmittal by citizens to the government of private information on, for example, their job status or earnings, or by a company of its core business information, such as employment, wages, and revenues, is a remarkable example of public/private partnership. Without the trust of our survey respondents, the BLS surveys and the important statistics they generate would be in jeopardy. H.R. 5215 will strengthen that trust.

Thank you, Mr Chairman. I would be happy to respond to any questions you may have.

Mr. HORN. We will ask a few questions before we get to the next panel, and then if you could all stay around, why, we would appreciate it. And that way we would have a better look at it.

Could you elaborate on how the bill will reduce data disparities. For example, how will it solve the problem of inconsistent standard industrial classification codes? Let's just go right down the way.

Mr. KROSZNER. Thank you very much, Mr. Chairman. As I had mentioned in my earlier remarks, a limited survey or limited study that had been done in 1994 trying to compare the classifications of firms by BLS versus the Census found about a 30 percent difference in the classifications. And so that's a very, very concrete illustration of how, by working together, they can try to work out some of the inconsistencies in the classification of firms.

Now, part of this is because firms change in a dynamic economy and it is very difficult to keep up as rapidly with them. So the Census may have done their classification in 1995. That firm today may still have the same name, but it may be doing something completely different that may be in a new area. Also, it may be doing something that wasn't even part of the classification codes from 5 to 10 years ago. And that is why it is very important to do this. And also, I want to emphasize that I think the agencies do a superb job of this, but it is just hard to keep track of such a vibrant entrepreneurial economy that we have.

And so I think it is extremely important to allow for this, and we can reduce the anomalies and inconsistencies by allowing the data to be shared, the different agencies to talk to each other to keep up with our very dynamic economy.

Mr. HORN. Can you give us an idea of how much duplicate data collection now occurs among the three agencies?

Ms. COOPER. How much? I'm sorry?

Mr. HORN. Just give us an idea of how much duplicate data collection now occurs among the three prime agencies here.

Ms. COOPER. Well, I think it's hard to give a general statement. My fellow panelist, Mr. Kroszner, mentioned with regard to classification of business firms. I think we also have to think about what will happen over time among the three agencies. I think it is just a very difficult question to answer; but we think that over time, we will reduce the duplication, and that clearly is a plus for the reporters, for businesses out there in this economy. It will lower the burden on them. And with the comparisons that we could make of cross agencies, that has to be a very real plus. But it is one of those numbers that is very difficult to figure out.

Mr. HORN. Well, this next question is about the same as the last one. But can you estimate the time and resources that this bill will give the agencies by reducing duplicate information collection?

Ms. UTGOFF. Again, I don't think we could give you a numerical answer to that question. But there are many instances. For instance, the International Price Program. If we are able to use the Census data, we will be able to reduce the number of people who have to respond to the survey, and we will be able to reduce the information that we get from each respondent.

Ms. COOPER. I might just add there, from the Census Bureau's point of view, that if we were able to reduce one survey across our broad set of businesses that we tend to survey, that, in and of

itself, would save us some $2 million. So I think we can begin to see that there are some real savings out there as we go forward, but it is just very difficult to come up with a precise number.

Mr. HORN. Now I am going to yield 10 minutes, to start with, with the gentleman from Ohio on any questions you want to ask.

Mr. SAWYER. Thank you very much, Mr. Chairman.

Ms. Cooper, it's my understanding that the Census Bureau is currently conducting research on matching data from household surveys with data from business surveys and censuses. The idea is to use them to model changes in labor force composition, future pension demands, and a variety of other important economic conditions. It's my understanding that under the terms of this legislation, that those data could not be shared with researchers at BLS or BEA, despite the fact of the substantial expertise in those agencies and the work that they might be able to do to understand those data. Can you explain why the legislation has been changed to exclude those data? Or am I incorrect in my understanding of it?

Ms. COOPER. That this legislation has been changed to exclude it?

Mr. SAWYER. I believe so.

Ms. COOPER. I honestly don't know the answer to that one.

Ms. UTGOFF. There were questions—this bill deals only with the information for firms. It does not deal with individual household response. In the previous efforts to have this legislation passed, the household response became very controversial, so it was not included in this round of the bill.

Mr. SAWYER. I would simply hope that as we gain experience with this kind of information sharing, that we could find a way to enable that kind of sharing. It is important. I couldn't agree with you more about the sensitivity of personal individual household information. But where it can be used to shape policy in important ways, I think it would valuable to do.

It leads me to my second question: Does the administration have a plan of action with regard to how, in future, this kind of legislation might be expanded to include sharing household data in order to improve our understanding of the changing nature of poverty, access to early childhood education, pension coverage, and a myriad of other kinds of social statistics that shape an awful lot of the debate that characterize our work here?

Mr. KROSZNER. There is no specific plan right now. I think what we want to do is gain experience with the data sharing with the business data to avoid the concerns and controversies that had been raised with the individual data. And I think as we have the experience with that, both the researchers outside of the Government and internally, we will be able to see how best to shape the memoranda of understanding to make sure that no data—no confidentiality agreements are violated.

And so I think it is an important foundation and first step. Much like we had some experience with sharing some international data back in 1990, I think that provides the basis for how well the agencies can work with that data to show that this will be functioning very well. And then perhaps in the future we can take it another

step, but there is no particular plan right now for that next step to be taken.

Mr. SAWYER. Go ahead.

Ms. COOPER. And could I just add, Congressman Sawyer, going back again to your question before and tying it to this one. I agree that the work that is being done that ties the labor market and other information together, the Census Bureau, is innovative and can be very helpful longer term, but as long as we—at this stage we can only do that in the confines of the Census Bureau. But as my colleague has indicated, once we can demonstrate that there will not be concerns longer term—I am not sure how long that takes, but we are all hopeful that it will be sooner rather than later—then perhaps this could be considered later, much later.

Mr. SAWYER. I appreciate the concerns that you have for confidentiality, which is what led me to my half of this legislation. I don't want those concerns, however, to stand in the way of serious innovation that can come about as a result of better sharing.

Let me ask each of you to respond to the fact that the administration sought and was granted a provision under the PATRIOT Act that provides the Attorney General access to individually identifiable survey records held by the National Center for Education Statistics. If the administration sought access to similar records held by the Census Bureau or BEA or even the Council of Economic Advisers, would you support or oppose access to survey records for law enforcement purposes? Each of you, please.

Mr. KROSZNER. Well, we have no records at CEA. So, for our data, we are happy to share what we have with anyone. Certainly, that's one issue that has arisen about maintaining confidentiality versus collecting data for law enforcement purposes, and there can sometimes be a tension there.

Mr. SAWYER. They are very much in tension, and I appreciate that. I shouldn't interrupt. Go ahead.

Mr. KROSZNER. No, no. And so we have tried to ensure that nothing that we have done in this legislation would, in any way, inhibit the ability of the Department of Justice to mete out justice to wrongdoers. But I don't think we at CEA have a particular view on that broader question.

Mr. SAWYER. Others?

Ms. COOPER. I don't think I have more to say on that, either. I think that is something that has to be settled elsewhere, and that is a very real tension and concern.

Ms. UTGOFF. I agree with the other panelists on this point.

Mr. SAWYER. Let me touch on something that the chairman touched on. This is kind of off the track. Nearly 20 years ago when I was a mayor in Akron, Ohio, we had gone from an—about 80 years where our signature industry was in the tire and rubber industry. It was 1984, and we hadn't made a passenger car tire in Akron since 1979. And the truth is, we were trying to chart a new future for ourselves as a community with some sense of realism.

We had done a great deal of work in expanding the product applications of synthetic materials and other kinds of polymeric applications in a wide range of different kind of product fields. And, but before we committed ourselves to that kind of work and to bringing together resources, as a city we decided we wanted to get a sense

of the current state of play in what we were loosely referring to as the polymer industry, and discovered to our frustration that it was very difficult to track that because standard industrial codes simply did not reflect the way in which the industry had shifted its field, not just individual companies—B F Goodrich, by the middle of the 1980's, was no longer a tire company, they were a chemical and aerospace company. But that's different from when an entire industrial field shifts its ground.

Can you talk to me just briefly about the how the ability to share data will help you track not the changing character of companies, but the changing character of large-scale enterprise in the United States?

Mr. KROSZNER. By being able to—actually, in some sense, by being forced to resolve anomalies between the different agencies when they've classify one firm one way and another firm another way, that forces the agencies to address exactly that issue much more head-on than they otherwise would have to, because suddenly they now have the same firm classified in two different ways, and neither of those classifications is appropriate for that firm anymore. And so by talking to each other, they'll say, well, maybe our classifications aren't appropriate. We have to try to build on our standard industrial classifications to take into account this dynamism.

And so I think that is one of the ways in which could help to have the agencies speak with each other, because they can then use their expertise together to say there is something wrong here, we've both misclassified this firm, it should be something new, and we have to innovate to come up with a new classification for it.

Mr. SAWYER. Thank you, Mr. Chairman.

Mr. HORN. Thank you. About 5 years ago, probably at maybe the same relevant question of this, I wanted, when I had all those brains looking at me, and I knew you had the answer. And that is, how does the OMB have one set of assumptions for their accuracy of Federal statistics, and the CBO on Capitol Hill use different assumptions? And is there any way we can get both of those wonderful, powerful operations that they can agree on a base? And how can we handle that? I realize that might have nothing to do with what you are saying. But I just want that now that I have got a few bright economists.

Mr. KROSZNER. Well, I think it is a very important issue, because it is sharing of a different type of sense of data and different types of assumptions across different groups. I think this is certainly an area of—in which reasonable people can differ about looking out, let's see, 10 years hence about what economic growth will be.

Actually, the long-term economic forecasts which we developed through the process of CEA and OMB and Treasury, much of the economics is actually quite similar to what we find in the CBO forecasts for—especially for the long-range growth assumptions? There is some differences on how the business cycle will move over the short-term and then differences on views on how you turn GDP into revenues. And so I think increasing the dialog between our groups would be very—would certainly be very valuable, but I don't unfortunately have sort of the magic bullet that can make us all agree. And if you look at private sector forecasters, they often have

very different assumptions that they make, and in some sense that's appropriate that people have different views, and those should be taken into account. I am not sure that we want to have just one single view. I think we are relatively close; we're not too far apart, but I think further dialog would certainly be helpful.

Mr. HORN. Ms. Cooper.

Ms. COOPER. I might just add that indeed they certainly have been fairly close in the last year, 2 years. They operate—the two organizations operate under different rules in terms of what the CBO's goal is and what the OMB's goal is. And the timing of their forecasts clearly is slightly different, not largely different. So I am impressed that they are as close to one another as they are, especially after they have a little bit of time to readjust.

But I would agree with my colleague that it's good to have not just one view going forward. It would be surprising if we had terribly different points of view on potential GDP growth and we actually don't. It's trying to wrestle at this point more than anything else with what level of revenues come with each dollar of GDP.

Mr. HORN. Ms. Commissioner.

Ms. UTGOFF. I don't have anything to add to that. I agree with what the panelists have said.

Mr. HORN. Well, maybe we will get something out of the next panel. Because up here—and, you know, various presidents have said, well, gee, I've got OMB; at that point it is what the President wants to have done. And then up here, we'd like to having something because, we are trying on spending and we are trying on not spending. And so everything we do, going through the appropriations, we have got to show them that this is going to be in either 5 years or whatever. And that tightens up things quite a bit around here. And so I still think there ought to be some way that they get together and they say, look, this is what is here and this is what we ought to know, and deal with it.

So, then in other administrations, why, they say, gee, we like what CBO did and so forth. And it's all back and forth. It's like a tennis game and little people return running every day over the, with getting the ball halfway over the net. So, anyhow, that's one little gripe I have. And I don't think it will ever be solved, but it would be nice if we had sort of a treaty of Versailles here for economics and, you know, maybe the mirrors in the palace of Versailles would do it with all the things they can get it through the sun, through the windows and everything.

Mr. SAWYER. That was as remarkable of a mixed metaphor as I think I've heard all day. I just thought maybe we could think of these guys as line judges in your game of tennis.

Mr. HORN. My colleague here and I have both been dealing with the European parliament since we got here, and he's very good with that group. And so I will leave it right there.

OK. Is there anything that we should have asked you that we didn't ask you? Well, that's to make a good conscience somewhere.

But we will get the next panel up. And then if we could get all—if you can stay, we can get the questions with all of you there.

We have Maurine Haver, William D. Nordhaus and Ralph Rector. They were previously sworn. Let's take Dr. Haver, start with

her. She is the chair of statistics committee of the National Association for Business Economics.

STATEMENTS OF MAURINE HAVER, PH.D., CHAIR, STATISTICS COMMITTEE, NATIONAL ASSOCIATION FOR BUSINESS ECONOMICS; WILLIAM D. NORDHAUS, PH.D., STERLING PROFESSOR OF ECONOMICS, DEPARTMENT OF ECONOMICS, YALE UNIVERSITY; RALPH RECTOR, PH.D., RESEARCH FELLOW AND PROJECT MANAGER, CENTER FOR DATA ANALYSIS, THE HERITAGE FOUNDATION

Ms. HAVER. I am pleased to testify today in support of H.R. 5215. I am speaking in my capacity of chair of the statistics committee of the National Association for Business Economics, on behalf of our 3,000 members. The members of our association have a keen interest in the quality of economic statistics produced by the BEA, Census and the BLS. We use these statistics daily in our work to help our companies and clients make informed business decisions that have real dollar consequences.

We believe that passage of this legislation will protect the confidentiality of our companies' proprietary information supplied to the Government. It will minimize the burden imposed on our companies by the statistical agencies because duplicate surveys can be eliminated. And, finally, it will improve the quality of our national economic information, especially statistics at the industry level, because of more complete and consistent source data.

The business community and financial markets derive significant benefits from the collection and dissemination of economic data. Complaints of respondent burden are often misinterpreted. Most companies recognize the value of Government statistics and actively use statistics produced by these agencies for the basis of many operation and planning decisions. Companies need industry and national statistics to have highest quality and are willing to do their part as long as confidentiality of their proprietary information can be assured and data collection is done efficiently.

This bill is important because it extends confidentiality protection for respondents to all Federal agencies that collect data for statistical purposes under a pledge of confidentiality, and it prohibits the use of those data for any other purpose. It also specifically prohibits disclosure of information under the Freedom of Information Act. This uniform set of confidentiality protections will do a great deal toward reducing concern about reporting to national Government agencies.

I believe the prior witnesses have devoted sufficient time to talking about the improvements that data sharing will make in our statistics. Let me just say that current statutory barriers to the sharing of business data do result in duplicate surveys that not only increase respondent burden, but also introduce classification errors that reduce accuracy. This bill will address both of these obstacles.

In summary, we believe this bill will encourage business participation in Government surveys and will improve the quality of the statistics available to business and policymakers.

We strongly urge passage of H.R. 5215. I would like to thank you, Mr. Chairman, for your invitation to participate in this hear-

ing today, and also for your efforts ever since 1995, when I was President of the NABE. You were there to help us work toward a more efficient Federal statistical system, and we greatly appreciate that. I would be happy to answer your questions.

[The prepared statement of Ms. Haver follows:]

Testimony of

Maurine A. Haver

Chair, Statistics Committee
National Association for Business Economics
and
President, Haver Analytics

Before the House Committee on Government Reform's Subcommittee on Government Efficiency, Financial Management, and Intergovernmental Relations

September 17, 2002

Chairman Horn, members of the Subcommittee, I am pleased to testify today in support of H.R. 5215, the "Confidential Information Protection and Statistical Efficiency Act of 2002". I am speaking in my capacity as chair of the Statistics Committee of the National Association for Business Economics (NABE) on behalf of our members. The members of our association have a keen interest in the quality of economic statistics produced by the Bureau of Economic Analysis (BEA), the Bureau of the Census and the Bureau of Labor Statistics (BLS). We use these statistics daily in our work to help our companies and clients make informed business decisions that have real-dollar consequences.

We believe passage of this legislation will

- Protect the confidentiality of our company's proprietary information supplied to the government.
- Minimize the burden imposed on our companies by the statistical agencies because duplicate surveys could be eliminated.
- Improve the quality of our national economic information, especially statistics at the industry level, because of more complete and consistent source data.

The business community and financial markets derive significant benefits from the collection and dissemination of economic data by these agencies. Complaints of respondent burden are often misinterpreted. Most companies recognize the value of

government statistics and actively use statistics produced by our federal statistical agencies as the basis of many operation and planning decisions. Companies need industry and national statistics of the highest quality and are willing to do their part through participation in government surveys as long as confidentiality of their proprietary information can be assured and data collection is done efficiently.

This bill is important because it extends confidentiality protections for respondents to all Federal agencies that collect data for statistical purposes under a pledge of confidentiality, and it prohibits the use of these data for any other purpose. It also specifically prohibits disclosure of information under the Freedom of Information Act. This uniform set of confidentiality protections will do a great deal toward reducing concern about reporting to Federal government agencies.

Current statutory barriers to the sharing of business data among agencies result in duplicate surveys that not only increase respondent burden but also introduce classification errors when different respondents in the same firm classify an establishment's primary business differently. This can happen because an employee responsible for payroll information may respond differently than one who is knowledgeable about revenues by business line. For those of us who work in businesses that were not recognized under the old SIC classification system, it is also not surprising that different classification choices are made by the same business establishment on different surveys. This legislation will not only reduce the number of opportunities for misclassification but will also allow the statistical agencies to flag inconsistencies and investigate data anomalies. The bottom line for all of us is greater efficiency and higher quality statistics.

Since NABE launched its Campaign for Quality Economic Statistics in 1995, it has conducted several surveys of its members. In our first survey (1995), 71 percent indicated that decisionmaking in their firm had been adversely impacted by lack of quality data. This percentage has climbed over time to over 80 percent. When asked the "most important step" that could be taken to improve the quality of government data other than

increased spending, 98 percent cited data sharing (52.8%) or a consolidation of statistical agencies (45.2%) that would permit data sharing. Our organization has always supported balanced budgets and has recognized that to meet the challenge of measuring this increasingly complex economy, the federal statistical agencies must be as efficient as possible. Data sharing is essential to increased efficiency.

Problems of data quality or lack of economic information pose heavy costs on our society. We must begin a renewed effort to improve our statistical system so it can provide us with the information we need to make appropriate decisions for the twenty-first century. We believe that this bill will encourage businesses to do their part through streamlined reporting and greater assurance of data confidentiality. We urge passage of this bill.

Thank you, Mr. Chairman, for your invitation to participate in this hearing and for your efforts over several years to improve our federal statistical system. I would be happy to answer questions.

Mr. HORN. Thank you, and we will pursue the questions with the two our presenters here.

William D. Nordhaus, Dr. Nordhaus, Sterling Professor of Economics, Department of Economics at Yale.

Mr. NORDHAUS. Thank you very much, Mr. Chairman. I would like to—I have some prepared remarks which I would like to submit for the record.

Mr. HORN. That is automatically put in the record.

Mr. NORDHAUS. I will summarize those briefly. I am delighted to have the opportunity to discuss the proposal for the sharing of statistical data in H.R. 5215. I think it is an important and useful bill and urge its passage.

I am involved with a number of different groups that are involved with economic statistics. I won't go over all of those, but they do include one that is closely related to the statistical agencies that I am the chair of the advisory committee of the Bureau of Economic Analysis. That committee has—actually saw an early draft of the data sharing proposal a couple of times in its oversight meeting and discussed those.

I emphasize that the remarks I am making are only my own and don't implicate any other organizations who like to have their own views. But I did do a nonscientific sample of a number of statisticians and economists and they are highly enthusiastic of this bill.

Good economic statistics, as you know, are important because they are critical inputs into the decisions of public and private decisionmakers, the Congress in its budgetary decisions, companies on their investments, State and local governments on their infrastructures, and the private sector and households on their financial decisions. All of these issues involve and require good statistical information.

Earlier this year, the Commerce Department conducted a brainstorming session of leading academic and business economists to consider improvements in the national economic accounts. And then earlier this year, the Joint Economic Committee held some hearings where it inquired into some different proposals for improving the Federal statistical system.

I appended at the end of my testimony a summary list of the recommendations that I made to the JEC. One of the major recommendations I made was that the Congress should move ahead expeditiously with improved data sharing among statistical agencies.

Now, the Federal—I would just say one word about source data. This is one of the less romantic parts of the statistical system that people don't really know much about. We see published in the newspaper every day the data on the GDP or the inflation rate or the balance of trade.

But these are really just the visible tips of the statistical icebergs, and below the surface lies vast volumes of source data from all corners of the economy, and they are collected by the Census Bureau, the Bureau of Economic Analysis, the Bureau of Labor statistics, the IRS, the Federal Reserve and many other Federal agencies, and the quality of our economic statistics depends crucially on accurate, timely and comprehensive source data.

Now, this bill—there are many ways to improve source data, but this bill proposes one that is extremely economical and useful, and that is solving the difficulties that arise from the decentralized nature of our Federal statistical system.

The current system has many advantages, but one big disadvantage is that agencies have a variety of statistical data that they cannot share. In a sense the Government has imported data in its left pocket, but that data cannot be moved to the right pocket, and that restriction just make no sense and should be lifted.

There are many examples of how data sharing will improve the quality of Federal economic statistics. I will concentrate on the national income and product accounts, which is the system I know best. I will just mention four briefly.

One is that early estimates of our gross domestic product are subject to large revisions because the source data are sparse and often based on voluntary reporting.

With data sharing, BEA will be able to use statistical techniques to correct the data for omissions to get more accurate early estimates of quarterly GDP.

A second problem is annual revisions which come in the middle of every year. And these are often large because many monthly Census surveys are based on voluntarily supplied data. The July 2002 revisions were particularly large because some of the data that came in during this year showed revisions that were far different from the preliminary data.

By working with individual company data and comparing them with publicly available data, BEA can identify discrepancies earlier and thereby reduce annual revision.

A third example is the statistical discrepancy between the product and income side of our accounts. This is currently running at $166 billion in the last quarter, which is more than 1½ percent of total GDP.

The source of the statistical discrepancy is still a mystery. But by comparing IRS, Census, and public data, BEA may be able to sharpen its estimates of different sources, particularly of income, and narrow that discrepancy.

A final example, and one that has been very much in the news, involves data on corporate profits. These are one of the most important and hard to measure of the statistics. And accurate measures of total corporate profits are produced only with a 3-year delay, and this is because it takes that amount of time to gather all of the different tax returns and tabulate them completely.

Inaccurate profit data may well have contributed to the boom and bust cycle of stock prices in the last few years. I believe that by triangulating data from tax returns, quarterly financial reports and publicly available financial statements, BEA can develop statistics on corporate profits more accurately and in a more timely fashion, and this, of course, can help investors gauge the true movement of profits in an era when financial reports are not always reliable.

As Dr. Kroszner noted, these statistical innovations can improve the quality of Federal statistics with little, no or even negative cost.

So, in summary, I think the proposal for data sharing contained in H.R. 5215 is a small but important step toward improving the efficiency and the use of Federal statistical resources, and I support strongly its enactment.

[The prepared statement of Mr. Nordhaus follows:]

Improving Federal Statistics Through Sharing of Statistical Data

William D. Nordhaus
Yale University

Testimony Before the

Subcommittee on Government Efficiency,
Financial Management, and Intergovernmental Relations
of the
Committee on Government Reform

September 17, 2002

Summary

I am delighted to have the opportunity to discuss the current proposal for sharing of statistical data as contained in H.R. 5215. My summary view is that this proposal is a useful and inexpensive means for improving the quality and timeliness of federal economic statistics. I strongly endorse it and recommend its speedy passage.

Personal Background

I am currently Sterling Professor of Economics at Yale University. In addition, I am a member of the National Academy of Sciences and the American Academy of Arts and Sciences. I am on the research staff of the Cowles Foundation for Research in Economics and the National Bureau of Economic Research.

I might add a word of background on my interest in economic statistics. For most of the last decade, I served on the National Academy of Science's Committee on National Statistics. This body is charged "to select and study statistical topics to improve the effectiveness of the federal statistical system." The Committee has devoted significant resources to studying issues of data sharing.

I am also currently chair of the Advisory Committee of the Bureau of Economic Analysis (BEA). This Committee works with the BEA to review priorities and make

suggestions on technical issues to improve economic statistics. Additionally, I am the chairman of the newly established Committee on Economic Statistics of the American Economic Association.

I have also served as a user of economic statistics both as a scholar and in an advisory capacity to the federal government. I served as a member of the President's Council of Economic Advisers from 1977 to 1979 and am currently a member of the Congressional Budget Office's Panel of Economic Advisers.

I emphasize that the remarks that I am making today are my own and in no way implicate any of the organizations just listed. I will, however, attempt to convey the broad consensus of professional economists on the importance of high-quality and timely statistics.

The Importance of Good Economic Statistics

Good economic statistics are important because they are critical inputs into the decisions of public and private decision makers. Without good economic statistics, the Congress cannot make budgetary decisions informed by economic trends and the long-term outlook for surpluses or deficits; companies cannot plan their investments without good data on prices and quantities in their own markets; state and local governments cannot plan for roads, hospitals, and environmental quality without up-to-date demographic data; and households cannot make sound financial decisions without reliable information on the earnings of companies and the yields on alternative investments. Conducting the fiscal affairs of state without good statistics is like flying blind.

A recent report from the National Academy of Sciences summarized the central economic role of our national economic statistics:

> The modern national income and product accounts are among the great inventions of the twentieth century. Among other things, they are used to judge economic performance over time, to compare the economies of different nations, to measure a nation's saving and investment, and to track the business cycle. Much as satellites in space can show the weather across an entire continent, the national accounts can give an overall picture of the state of the economy.[1]

[1] *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment*, National Academy Press, Washington, D.C., 2000.

Earlier this year, the Commerce Department conducting a "brainstorming session" of leading academic and business economists to consider improvements in the national economic accounts.[2] I discussed many of the proposals that emerged from that session along with other ideas at a recent hearing of the Joint Economic Committee (JEC).[3] I append at the end of this testimony a summary list of the recommendations that I made to the JEC. One of the major recommendations I made to the JEC was the Congress should move ahead expeditiously with improved data sharing among statistical agencies.

Strengthen Source Data for Our Statistical System

Federal statistical agencies produce a wealth of statistics each month on output, productivity, incomes, foreign trade, inflation, the labor market, and many other facets of our economic life. The regular production of our national economic statistics gives a misleading impression about how easy it is to produce reliable, comprehensive, and timely data.

But the fact is that the numbers we read about each month - the GDP, the inflation rate, and the balance of trade - are just the visible tips of the statistical icebergs. Below the surface lie vast volumes of source data from all corners of the economy. This source data is collected by the Census Bureau, the Bureau of Economic Analysis, the Bureau of Labor Statistics, the Internal Revenue Service, the Federal Reserve, and by other federal agencies. Our federal economic statistics depend crucially on accurate, timely, and comprehensive source data.

The United States has made major investments in improved source data, and these investments have paid off in more timely and reliable statistics. But at present there are major gaps. This hearing is devoted to data sharing, which is an important step in improving the quality of federal economic statistics.

The Administration has made a proposal, which was announced in July 2002, to improve procedures for sharing statistical data. This proposal is very closely aligned with HR 5215, the "Confidential Information Protection and Statistical Efficiency Act of 2002".

---

[2] Bureau of Economic Analysis, *Survey of Current Business*, May 2002, "BEA's Strategic Plan for 2001-2005," with discussion by the BEA Advisory Committee and others.

[3] William D. Nordhaus, "An Economist's View of he Statistical State of the Nation," Testimony Before the Joint Economic Committee, July 24, 2002, available at http://www.econ.yale.edu/~nordhaus/homepage/recent_stuff.html .

This bill is a step to solving some of the difficulties that arise from our decentralized federal statistical system. The current decentralized system has many advantages. But one major disadvantage is that agencies have a variety of statistical data that they cannot share. In a sense, the government has important data in the left pocket but that data cannot be moved to the right pocket. This restriction makes no sense and should be lifted.

The proposal makes a major step by allowing data sharing. It combines two elements: First, it allows BEA, Census, and BLS to share business data for statistical purposes. Second, it clarifies and strengthens safeguards on confidentiality of information provided to government agencies.

Both of these provisions are important, but I will address only the first element, the guidelines on data sharing. It has long been recognized that the decentralized nature of the U.S. federal statistical system contains hurdles to the most efficient use of the statistical information that the government collects. Data sharing among the three statistical agencies will provide the opportunity to get both more timely and more accurate data on production, sales, employment, and industry.

There are many examples of how data sharing will improve the quality of federal economic statistics. I will concentrate on improvements to the National Income and Product Accounts, which is the statistical system that I know best. Some examples of benefits of data sharing are the following:

- Early estimates of GDP are subject to large revisions because the source data are sparse and often based on voluntary reporting. BEA will be able to derive more accurate early estimates of quarterly GDP by comparing publicly available company data with microdata on shipments and other series available from Census surveys.
- Annual revisions in mid-year are often large because many monthly Census surveys are based on voluntarily supplied data. For example, the July 2002 revisions were particularly large. By working with individual company data and comparing them with publicly available data, BEA can identify discrepancies earlier and thereby reduce annual revisions.
- The statistical discrepancy between the product and income side of the national accounts is extraordinarily high, running at around $166 billion in the last quarter. By comparing Census, IRS, and public data, BEA may be able to sharpen its estimates of different sources of income and narrow the statistical discrepancy.
- Data on corporate profits are one of the most important and hard-to-measure statistics, and accurate measures of total corporate profits are produced with a three year delay. Inaccurate profit data may well have contributed to the boom-

and-bust cycle of stock prices in the last few years. By triangulating data from tax returns, quarterly financial reports, and publicly available financial statements, BEA can develop statistics on corporate profits in a more timely fashion. Better BEA profits data can help investors gauge the true movement in profits in an era when financial reports are not always reliable.

These statistical innovations can improve the quality of federal statistics with little, no, or even negative cost.

In summary, the administration's proposal for data sharing as contained in HR 5215 is a small but important step toward improving the efficiency in the use of federal statistical resources. I strongly support its enactment.

Summary Table of Recommendations from Testimony Before the Joint Economic Committee [4]

Recommendation 1. The first priority for the BEA is continuing to improve the coverage, detail, quality, and timeliness of the core accounts. The BEA strategic plan contains many elements that are essential for continued improvements in the NIPAs.

Recommendation 2. BEA should work to develop an experimental monthly GDP series.

Recommendation 3. Reliable statistics on the economy depend upon improvements in the source data that underlie the statistics.

Recommendation 4. Enhanced data sharing among statistical agencies will improve the timeliness and accuracy of federal economic statistics.

Recommendation 5. The new American Time Use Survey (ATUS) of the BLS will fill a critical statistical gap by providing more accurate data on hours worked as well as a broad perspective on how the population spends its time. This survey deserves strong Congressional and public support.

Recommendation 6. BLS and BEA should continue their efforts to improve the price data underlying the consumer price index, the producer price index, and the national accounts. Special efforts should be made to capture in price indexes the full range of new and improved goods and services.

Recommendation 7. BEA should work with the Federal Reserve to develop a full set of asset and wealth accounts.

Recommendation 8. BEA should develop a full set of linked National Economic Accounts that include production, income, consumption, accumulation, and wealth.

Recommendation 9. In order to improve the quality of information about publicly help corporations, corporations should publish their tax returns and reconcile their financial reports with their tax returns.

[4] This list is drawn from testimony before the Joint Economic Committee, July 24, 2002.

Mr. HORN. Dr. Ralph Rector, is the research fellow and project manager at the Center for Data Analysis at the Heritage Foundation. Welcome.

Mr. RECTOR. Thank you. Chairman Horn, thank you for inviting me to testify today. My name is Ralph Rector. I am the Project Manager of the Heritage Foundation's Center for Data Analysis. It should be noted that the following testimony is my own view, not necessarily that of the Heritage Foundation or any other organization.

Today, I would like to discuss three standards I believe should guide any proposal to improve America's statistical system. These standards are, first, protection of individual identity for the respondents who provide data; second, production of useful, timely information for data users; and third, independent evaluations of data for decisionmakers. I think of these as the three I's of statistical policy: Identity protection, information value, and independent evaluation.

The sections concerning statistical efficiency contained in title II of H.R. 5215 are examples of measures that can enhance information value by improving the accuracy and timeliness of economic data.

My testimony will focus primarily on the issues related to title I, Standard 1, identity protection. Those who provide data to statistical agencies should not have to worry that the data they provide to the Government will be used against them. In addition, statistical agencies must protect the identity of individuals who provide data that may eventually be released to the public. Provisions for protecting individual identities can be found in plans such as H.R. 5215, which clearly distinguish between statistical and nonstatistical data.

Also, these prohibit the release of data in a form that could reasonably be expected, either directly or indirectly, to yield the identity of the respondent.

Standard 2, information value. Although necessary, procedures that protect confidentiality also tend to reduce the amount and the value of data that can be released. It is not, however, necessary to adopt such extreme forms of data suppression as those found in H.R. 5215.

As currently written, this bill states that agencies cannot disclose data in identifiable form. The bill further defines identifiable form to mean representation of information that permits information about a respondent to be reasonably inferred through either direct or indirect means.

This method of protecting confidentiality precludes the disclosure of all individual level information that respondents would provide, despite the use of safeguards that protect the identify of the respondents. The problem with H.R. 5215 arises because it does not clearly distinguish between the identity of the individual respondent and the information they provide.

Denying researchers access to all individual level data would drastically reduce the value of publicly available information and would undermine the quality of important research conducted in the United States.

Standard 3, independent evaluation. Although valuable, it is not enough for Government statisticians to approach data availability solely in terms of the amount of data they provide. In addition, the data should be sufficient so that researchers outside the Government can respond effectively to evaluate the proposals, either to validate them or challenge them.

My written testimony includes two examples that may help clarify why the distinction between the amount and the form of data accessibility is so important to nongovernment researchers who provide public policy analysis.

To implement three statistical standards described in my testimony, I believe Congress should, with regard to identity protection and information value, provide guidelines similar to H.R. 5215. However, the guidelines should be modified to clearly indicate that confidentiality applies to the identity of the respondent. The current version of H.R. 5215 is not sufficiently clear in this respect.

With regard to independent evaluation, Congress should require that whenever possible Federal agencies provide data to independent researchers in a form that permits them to conduct complete and independent evaluations.

Thank you.

[The prepared statement of Mr. Rector follows:]

The Heritage Foundation 214 Massachusetts Avenue, N.E. Washington, D.C. 20002-4999 (202) 546-4400

Congressional Testimony

Before the
United States House of Representatives
Committee on Government Reform
Subcommittee on Government Efficiency, Financial Management and
Intergovernmental Relations

September 17, 2002

Testimony on the "Confidential Information Protection and Statistical Efficiency Act of 2002," H.R. 5215

By
Ralph Rector, Ph.D.
Research Fellow and Project Manager
Center for Data Analysis
The Heritage Foundation
Washington, DC 20002

Mr. Chairman, Members of the Subcommittee, thank you for inviting me to testify on the "Confidentiality Information Protection and Statistical Efficiency Act of 2002," H.R. 5215. I ask that my written testimony be entered into the record.

I am the Project Manager of The Heritage Foundation's Center for Data Analysis (CDA). I help direct the work of researchers who routinely use a wide variety of data supplied by the federal government. In addition, the CDA has entered into licensing agreements with a few federal agencies that permit our analysts to use data that are not generally available to the public.

Although The Heritage Foundation is recognized as a conservative public policy research institution, our analysts work with those from diverse ideological perspectives on issues involving access to quality data. This is the reason why The Heritage Foundation is a member of broad-based organizations such as the Association of Public Data Users (APDU) and is an affiliate member of the Council of Professional Associations on Federal Statistics (COPAFS). It should be noted that the following testimony is my own view and does not necessarily reflect that of The Heritage Foundation, or any other organization.

Three standards for improving federal statistical policy

Government statistics are an indispensable component to much of the work done by policy makers. Obvious examples include economic indicators such as inflation and unemployment and budgetary estimates involving taxes and the overall level of spending. Crime, education and health care are just a few of the other public policy areas in which statistics are regularly used to better understand social problems and evaluate programs that may affect them.

Today, I would like to discuss three standards that should guide any proposal to improve America's statistical system. These standards are: (1) protection of individual identity for the respondents who provide original data, (2) production of useful, timely information for data users, and (3) independent evaluations of the data for decision-makers. These are the three I's of statistical policy: Identity protection, Information value, and Independent evaluation.

The need to improve federal statistical policy is directly related to our nation's dependence on high quality statistics. Data sharing provisions, such as those contained in H.R. 5215, can improve the quality of economic statistics produced by the government. In addition, with appropriate modifications, the identity of those providing data can be better protected by confidentiality policies such as those in H.R. 5215. However, as I will explain later, it is crucial that the language used to protect confidentiality not inadvertently and unnecessarily eliminate the type of data access that is currently available. After allowing for reasonable adjustments to protect the identity of respondents, the public should have access to the greatest amount of data possible. In addition, data should be provided in a form that allows nongovernment researchers to

provide alternative interpretations of information produced by the government's statisticians.

Two of the principles cited above have been applied in H.R. 5215. The sections concerning statistical efficiency contained in Title 2 are examples of measures that can enhance the value of information by improving the accuracy and timeliness of economic data. I have left the more detailed discussion of these issues to the economists and information providers who work daily with these data. My testimony will focus primarily on the identity protection aspects of Title 1. I will also discuss the importance of data access to nongovernment researchers.

Standard 1: Identity protection

Given the importance of numbers to government decision-makers, it is perhaps surprising that the federal statistical system is so fragmented and confusing. Individual agencies have been added to the U.S. statistical system over a period of many years and for different legislative reasons. Over 70 agencies participate in the collection, preparation, and dissemination of data collected from administrative records, surveys and censuses. While some agencies routinely generate wide-ranging products (e.g., the Bureau of the Census) others focus on more specific areas. In addition, statistics are produced as by-products in data collection associated with administrative tasks (e.g., the Internal Revenue Service).

The growth of America's statistical system has produced not only a confusing set of statistical agencies, it has also created an inconsistent set of laws and policies designed to protect the confidentiality of respondents who supply the government with data.[1] Some of the interagency coordination problems between the Department of the Census, the Bureau of Labor Statistics, and the Bureau of Economic Analysis would be reduced by changes such as those in Title 2 of H.R. 5215.

In addition, the legislation provides a new set of definitions and protections of confidentiality that would apply throughout the government. Protections such as these are important because the federal statistical system faces a serious problem of declining public trust in government, specifically trust that a respondent's identity will be kept confidential and that respondents will not be harmed by the information they supply. A uniform policy to protect the confidentiality of data providers is basic to the development of high-quality data. Unless respondents can be assured that the data they provide to the government for statistical purposes will not be used against them through regulations or other enforcement efforts, they will either not provide data or they will report inaccurate information. In either case, the effect is to create measurement biases and errors.

Unfortunately, Congress is not actively considering any proposal that would replace the current system with a coherent and comprehensive set of rules for the protection of confidentiality. Nevertheless, standards such as those in H.R. 5215 provide a framework for resolving these differences in the future. An important first step is to clearly distinguish between statistical and administrative data.

The government collects a vast amount of administrative data in conjunction with federally funded programs. With appropriate safeguards, these data can be used for research purposes. For example, administrative data can be used to determine whether federal job training programs are effective in raising the incomes of workers. However, data collected for statistical purposes should rarely, if ever, be used for administrative reasons.

Those who provide data to statistical agencies should not have to worry that the government will use their individual responses to decrease a monthly benefit check, increase their tax liability, or impose a fine for violating a government regulation. Confidentiality protections that clearly distinguish between statistical and nonstatisitcal purposes, such as those found in H.R. 5215, will help reinforce this important difference.

Statistical agencies must also protect the identity of individuals who provide data that may eventually be released to the public. Agencies protect confidentiality by modifying or suppressing data that could be used to directly or indirectly identify an individual respondent. Items such as names, addresses and identifying codes such as social security numbers are removed from publicly available databases.

In addition, reasonable steps are taken to ensure that *statistical disclosure* does not occur. Statistical disclosure can occur if the information that is released is so detailed, analysts can, with a high degree of probability, associate the information with a specific person or business. Statistical agencies use procedures to alter data in order to reduce the chance that this type of disclosure will occur. Examples of these adjustments include cell suppression, the random modification of data, and the use of topcoding.[2] The effect is to produce a database that is similar to the original file but with anonymous information. Data in this form limits the risk that the identity of respondents can be exposed through indirect means. Provisions for protecting individual identities can be found in plans such as H.R. 5215, which prohibit the release of data in a form that could reasonably be expected to either directly or indirectly yield the identity of a respondent.

Standard 2: Information value

Although necessary, procedures that protect confidentiality also tend to reduce the amount and the value of data that can be released. Technical adjustments to the data by statistical agencies reduce the usefulness of data that is available to the public and researchers. It is vital that the methods adopted to protect individual identity do not inadvertently or unnecessarily reduce the amount of information available to the public.[3]

It is important that a distinction be made between a respondent's identity and the data they provide. Individual-level data are often referred to as microdata files because they contain information about individual persons, families, business entities or some other individual unit. They include items such as age, race, sex, education levels, income and expenses. Examples of these files include the Current Population Survey, the Consumer Expenditure Survey, and the Survey of Consumer Finance. These files provide the basis

for much of the social and economic research conducted by analysts in academic institutions and in public policy organizations. This research depends on convenient access to individual-level data.[4]

Provisions to protect confidentiality are intended to shield the identity of the respondent but not suppress all data at the individual level. It is not necessary to adopt such extreme forms of data suppression as those found in H.R. 5215. As currently written, this bill states that agencies cannot disclose data that are in "identifiable form." The bill further defines data in "identifiable form" to mean the representation of information that permits information about a specific respondent to be reasonably inferred through either direct or indirect means. This method of protecting confidentiality precludes the disclosure of all individual-level information that respondents provide despite the use of safeguards that protect the identity of the respondents. Denying researchers access to all the individual-level data would drastically reduce the value of publicly available information and undermine the quality of important research performed in the United States.

The problem with the approach taken in H.R. 5215 arises because it does not clearly distinguish between the identity of the individual respondent and the information they provide. Protection of confidentiality requires that the identity of the individual be kept confidential. However, other information that is currently available to researchers should remain accessible. Confidentiality protections such as those in H.R. 5215 should be modified so it is clear that they protect the identify of respondents.

Data providers often refer to a tension between the protection of individual identity and the degree of information usefulness. On the one hand, government statisticians want to reassure respondents who provide data. On the other hand, they would like to fulfill legitimate requests for data by users. The tension is often depicted by statisticians in a graph where the risk of disclosure is measured on one axis and the amount of information provided is measured on the other axis.[5] The graph shows a trade-off in which a lower level of disclosure risks leads to a reduction in the amount of information that can be provided. The goal is to strike a balance that provides reasonable protections for confidentiality and the greatest amount of useful data. Although helpful, graphs that only plot disclosure risks and the usefulness of data omit the role that data plays in protecting our form of government.

Standard 3: Independent evaluation

Although providing valuable data is a very important standard, it is not enough for government statisticians to view data access solely in terms of the amount of data they provide to the public. In addition, the data should be sufficient so that researchers outside the government can respond effectively to government proposals – either to validate or to challenge them. To function properly, the U.S. government depends on the ability of potentially opposing interests to influence the decision-making process and thereby reach a more informed and reasoned outcome. The U.S. system of government was designed with checks and balances, and depends for its effectiveness on the free flow of information.

There is a subtle but critical difference between a standard for the quality of information that is provided and a standard that deals with the form in which it is provided. Government statisticians may supply the public with a large quantity of valuable data but this information typically comes packaged in numerical aggregations and generalized categories. If nongovernment researchers are to provide an independent evaluation of official government data, they must have access to information that is similar to that used by government statisticians. Without this access, a basic U.S. principle of open government, reflected in the U.S. Constitution and in many laws, most notably the Freedom of Information Act (FOIA), will be violated. The U.S. government was designed to be of and for the people, not to be run by an elite with the unique ability to choose how data are to be categorized, processed, and released.

A few examples may help clarify why the distinction between the amount and form of data accessibility makes a difference. I have selected two studies conducted by Heritage's data center and ask that they be included in the record.[6] Although these are Heritage publications, I must point out that public policy analysts commonly produce this type of research and I could have selected from a large number of studies from individuals associated with universities and nonprofit organizations.

The first report is an analysis of the distribution of income in the United States. The authors of this study identify four weaknesses with the official measurements of income inequality used by the Census Bureau. For example, the quintiles that Census uses to divide income do not contain an equal number of people. In addition, the conventional Census figures do not take into account the effects of taxation and omit many types of cash and non-cash income. Because the underlying Census data are publicly available, Heritage analysts were able to make the adjustments they believed were appropriate to recompute the distribution of income. The revised analysis shows a more even distribution of income than that contained in official Census reports.

A second Heritage report asked what share of child poverty can be attributed to the growth of single parenthood since the 1960s. As with the previous study, analysts used data in a form similar to that available to Census statisticians. The report notes that "The March 2001 [Current Population Survey] supplement, also known as the annual demographic file, includes extensive questions on family demographic characteristics and previous year income that make it useful for social analyses, such as this one."[7] Heritage analysts utilized the Census data to estimate the effects that marriage rates have on poverty. They were also able to use an expanded definition of income that counts the Earned Income Tax Credit and food stamps as part of a family's resources for determining whether the family is poor.

Examples of similar research can be found in Heritage reports on education, taxation and the Social Security system. And, more important, other public policy analysts who have divergent political perspectives rely on the same type of data. Although statistical agencies often state that they are committed to providing access that allows for independent evaluations there are few regulations or laws that require them to do so.

Authors and sponsors of federally funded program evaluations seem particularly reluctant to release their data sets to independent researchers.[8] Requiring public access to program evaluation data encourages government evaluators to apply more rigorous methods than would otherwise be the case. If we are to have open and informed debate on public policy issues it is vital that all researchers have access to data that permit them to challenge the government's official reports and to offer alternative perspectives.

What Congress Should Do

To implement the three statistical standards described in this testimony, Congress should:

- Provide guidelines, such as those in H.R. 5215, that clearly distinguish between data that are used for statistical and nonstatistical purposes. In addition, the guidelines should specify, as they do in H.R. 5215, that reasonable measures be implemented so that respondent identities cannot be determined either directly or indirectly.

- Provide guidelines that clearly indicate that confidentiality applies to the identify of the respondent. The current version of H.R. 5215 is not sufficiently clear in this respect. The protection of a respondent's identity does not require that all the information about the respondent be suppressed.

- Require that, whenever possible, federal agencies provide data to independent researchers in a form that permits them to conduct complete and independent evaluations.

---

[1] Joe Cecil, Senior Research Associate at the Federal Judicial Center, notes that "Records maintained by U.S. federal agencies are governed by a web of federal statutes that are 'inconsistent at best and chaotic at worst" (Commission on Federal Paperwork, 1977). The exchange of statistical information must conform to standards that often were designed to guard against administrative abuses, standards that may be inappropriate for records used only for statistical purposes. As a result, researchers who seek information maintained by federal agencies often must recast their request for access in terms of a regulatory scheme that does little to anticipate the special characteristics of statistical data." See Joe S. Cecil, "Confidentiality Legislation and the United States Federal Statistical System," Journal of Official Statistics, Vol. 9, No. 2, 1993, p. 519.

[2] For review of the adjustments that statistical agencies employ and the possible effects they may have on the usefulness of data see articles in Pat Doyle, Julia I. Lane, Jules J.M. Theeuwes, and Laura V. Zayatz, editors, Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies (New York: Elsevier Science, 2001).

[3] Some agencies are allowed to provide data to external researchers through data licensing or use agreements. These licenses extend the legal responsibilities for handling confidential data to the external researcher. They can be an effective means of preserving respondent confidentiality without significantly affecting the quality of research that can be performed off-site by nongovernment analysts. For a review of licensing arrangements see: Marilyn M. Seastrom, "Licensing," pp. 279-289, in Doyle, et. al., editors, Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies. Other alternatives, such as making researchers special sworn employees, are much less effective in providing data access. The access provided is time-consuming to obtain, costly, temporary and must be carried out at a remote site. In addition, special requirements often limit the research to those subjects that further the mission of the statistical agency.

[4] This issue was considered by the members of The Panel on Confidentiality and Data Access of the Committee on National Statistics. They warn that efforts by statistical agencies to protect confidentiality could significantly reduce the value of the data. "Because of legitimate concerns about the possibility of disclosure of individual information, statistical agencies have limited the amount of detailed data provided to nongovernment users in tabulations and public-use microdata files. This lack of detail restricts the ability of users to do analyses that could contribute to the understanding of significant economic, social, and health problems." The panel recommended that "Statistical agencies should continue widespread release, with minimal restrictions on use, of microdata sets with no less detail than currently provided." See George T. Duncan, Thomas B. Jabine, and Virginia A. de Wolf, editors, Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics (Washington, D.C.: National Academy Press, 1993), p. 7.

[5] See, for example, various papers in: Doyle, et. al., Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies.

[6] See the attached reports: Robert Rector and Rea S. Hederman, "Income Inequality: How Census Data Misrepresent Income Distribution," The Heritage Foundation, Center for Data Analysis Report, September 29, 1999, and Robert Rector, Kirk A. Johnson, and Patrick F. Fagan, "The Effect of Marriage on Child Poverty," The Heritage Foundation, Center for Data Analysis Report, April 15, 2002.

[7] Rector, Johnson, Fagan, "The Effect of Marriage on Child Poverty," p. 3

[8] For example, The National Job Corps Study, funded by the Department of Labor (DOL) and authored by Mathematica Policy Research (MPR), was published in July 2001. The DOL and MPR have denied requests to release the data used for the study. In addition, the Community Oriented Policing Services (COPS) refused a FIOA request by The Heritage Foundation to release data from the National Evaluation of the Effect of COPS Grants on Crimes from 1994 to 1999.

Mr. HORN. Thank you. Let's go to a few questions here. Where in the bill would you like, Dr. Rector, to solve that problem?

Mr. RECTOR. I believe that the problem exists in the definition of identifiable form, and I have provided specific line numbers to the staff so that they can see exactly what the language would, I think—it would need to be in order to very clearly distinguish between the individual respondent, their identity and the information that they provided.

Mr. HORN. And my staff has provided you a copy of the amendments I intend to offer at today's markup on H.R. 5215. Does this amendment satisfy your concerns with the original language of the bill?

Mr. RECTOR. I have read the substitute amendment to H.R. 5215 provided by your staff, and I do believe that the new version does correct this problem.

Mr. HORN. So that solves that problem, good. Something is happening today anyhow.

Let me ask you a few questions that my colleague didn't have a chance to do it, Dr. Sawyer, who was needed elsewhere. And he and I both worry about the Attorney General, who has sought access to survey information provided by individuals. What would be the effect on economic statistics of information collected from businesses for statistical purposes if it was used in legal proceedings by the Government against the businesses and the executives that provided that information? How do you feel about that one way or the other?

Ms. HAVER. I am only speaking for myself now, because we certainly haven't queried our members. But I would not be supportive of the Attorney General having access to information that is provided to the Government for statistical purposes.

Mr. HORN. Dr. Nordhaus.

Mr. NORDHAUS. My own view again, it depends a little bit on the context. But it seems to me that whether it is a person or a business responding to a survey for statistical purposes, I think it would be very chilling for—to get good responses to that if there were the possibility of it being used in legal proceedings. I think that is particularly applicable for voluntary surveys. For mandatory then the person would have some—would have a tug of war between which of the two provisions were more—he was more fearful of.

But in the case of voluntary surveys, people would say, well, it is voluntary, why should I fill it in if there is any chance that this would be used against me in a legal proceeding. So I think that is a serious, serious concern.

Mr. HORN. Dr. Nordhaus, the Federal Economic Statistics Advisory Committee is a unique committee in that it is charged with addressing the interaction among these three statistical agencies. Would you please comment on the role that the committee might play in advising those agencies in carrying out the intent of this legislation?

Mr. NORDHAUS. The FESAC, as it is called, is an agency or was a group that was set up a couple of years ago. The Bureau of Labor Statistics was the lead agency on that, although the Census Bureau and the Bureau of Economic Analysis were also participants.

I was—I have met with that as—in terms of my advisory capacity of the BEA.

I think all of these advisory committees can play a very useful role in terms of reviewing proposals like this. As I think I mentioned in any oral remarks, the BEA advisory committee actually discussed the proposal that Dr. Kroszner brought before it earlier this year. I think it was a very useful discussion among the different people, because people from different points of view, from business, from the research community and from Government, all had something to say.

So I think this is a useful forum, these are useful forums for discussions of those kinds of issues. It has been useful in the past, and I hope it can be again in the future.

Mr. HORN. Mr. Sawyer's question is to Dr. Nordhaus. While this bill provides a mechanism to bring the collective talents of these three agencies to bear to improve economic statistics, it does not provide a mechanism to draw on the expertise in universities and private research organizations.

Do you have thoughts on how we can take that next step?

Mr. NORDHAUS. Well, I—there are two separate issues here. One is the actual production of statistics. I regard producing the GDP and the CBI as production. I mean it is a very intellectual and high level production, but it is like producing cars in the sense you have got to roll them out every month.

There is a separate question, which is the research that lies behind those statistics, and there is where a very useful role can be played. Some of the agencies, particularly the Census, has taken the lead on this, have basically taken researchers as employees of those statistical agencies to help do research on behavior, on the behavior of particular series or relationships, and I think those have been very useful in bringing academic research to bear on the questions.

But for the most part, actually I think the research staffs of those agencies are very high level, and I think the data sharing will go a long way in improving some of the easily fixed problems with the statistics.

Mr. HORN. Dr. Haver, do you agree with him?

Ms. HAVER. Yes, I do. I think that certainly the production of the statistics is something that requires compromises that sometimes our academic colleagues would prefer not to see. We occasionally find the theory is wonderful, but then there is the application of that theory, and that sometimes becomes much more difficult. But I think that the academic community has a lot to contribute and is doing so, for example, through organizations like the BEA Advisory Committee which Professor Nordhaus chairs and other organizations like these.

Mr. HORN. Dr. Rector, would you agree with both of those colleagues?

Mr. RECTOR. I would, but I would like to followup on a comment just made about the compromises. As I had indicated in my testimony, I think that there is a tradeoff between protecting individual identity and information value.

The protection of identity tends to reduce the quality and the value—the amount of information that can be released. I believe

that the agencies' data disclosure review boards, the boards that make many of the compromises that were just described, they do not work closely enough with data users to produce compromises that are needed, and I would encourage Government statisticians to work with users more closely as they go through their normal effort to review these tradeoffs, make the decisions, and to involve data users earlier on in that process.

Mr. HORN. Well, I think that makes sense. As a user of Government data, what suggestions would you give to the Federal statistical agencies as they implement the provisions of this bill?

Mr. RECTOR. Well, again my focus is more on the title I aspects, and those are very wide ranging. They don't just deal with the business data. I think that there are many data bases that the Government produces, not only that have to do with statistical reports and the activities of the statistical agencies, but there are many data bases that are produced in conjunction with reports that Congress has mandated for policy evaluation, policy review.

It is difficult, sometimes impossible, for independent researchers, whether it is in a think tank, or whether it is in the Academy, to gain access to these data bases. I would encourage agencies, again not just the statistical agencies but all Federal agencies that collect data, to require as part of that, particularly the program evaluation studies, to include a mechanism for the timely release of these data bases to independent researchers.

Mr. HORN. I learned about 30 years ago that it was very difficult to get faculties involved with the politicians, now that I am one, and the problem is time, and we need it in 24 hours. They need the sabbatical every 7th year, and they will think about it. So that is a little problem that we have got there.

But what do you feel should be done by the Government side, although you touched some of it, what can be done with these data?

Mr. RECTOR. Well, specifically with regard to policy evaluations, because so many of those are mandated by Congress, I would like to see that data dissemination actually be built into the requirement when a report is released, that the data be made available at that point in time for independent review, peer review, evaluation by other researchers, that just be part of the grant process.

Mr. HORN. What do you see as the greatest hurdle in trying to improve response rates from private sector companies?

Ms. HAVER. I think it is a very big problem actually of—certainly as corporations are having more and more financial difficulties or are not performing as well as they might like, they are looking at all corners of their business and trying to make sure that every person in that company is doing something that enhances profitability.

So clearly filling out Government forms isn't on that list, or isn't high on that list. So it really is important that the data that is released by the Government really accurately reflect the industry, if it happens to be industry data, that companies might actually find useful.

I think we will go a long way toward improving response rates if in fact companies do believe that the information reported on their industry is accurate.

We had a reference earlier today to the semiconductor reporting problem. That was very simply that the companies did not think that those statistics accurately reflected their business. And so I think it is very important that we improve those statistics. And, as was explained earlier, part of improving our industry data is simply getting establishments put into the right classification. And if our—if the study that was done with 1994 data is accurate today, and I think it is probably worse rather than better, then we have to believe that only 70 percent of single establishment companies are classified correctly.

And, therefore, you know, we have a lot of mistakes going into that information, which explains why companies often don't think they truly reflect the reality that is out there.

Mr. HORN. Dr. Haver, you indicate that data users in industry are strong supporters of H.R. 5215. Do you know what industry data suppliers think of the bill?

Ms. HAVER. Well, after I was asked to testify today I called up a variety of representatives of data suppliers, the National Association of Manufacturers, the U.S. Chamber of Commerce, the NFIB, the Semiconductor Industry Association, and so on, to discuss this bill with them.

They were quick to point out that they really couldn't speak on behalf of their members because they had not asked their members their opinions on this legislation. However, they did say to me that they could see absolutely nothing in this bill that would cause their members, in their personal opinion, to not support the legislation.

In other words, it seemed to them, in one person's words, a no-brainer. This was legislation that improved assurances of confidentiality for companies and at the same time provided our fundamental general statistical agencies with the capabilities of sharing information to really improve our data and to make it more relevant.

So although I can't say that there is not some company out there that might have a problem with this legislation, I have to say, among the numerous organizations I did talk to, I did not hear anyone express that view.

Mr. HORN. For all three of you, do you see any downside to the bill from the viewpoint of industry? We have heard Dr. Haver.

Ms. HAVER. No.

Mr. HORN. Dr. Nordhaus.

Mr. NORDHAUS. No. I think it is a plus. I think that we can actually do more with the data that we have now with data sharing. So I think it is a plus.

Mr. RECTOR. I am unaware of any downside.

Mr. HORN. OK. What do you see as the greatest benefits of data sharing provisions for users of Government data like yourselves?

Ms. HAVER. More accurate, relevant data. I think that we have discussed this really at some length today, so I won't go through some of the points again. But there are many cases where the ability to share information among our agencies will at least give us the possibility of improving our statistics, and there are many situations today where we know that somebody is wrong.

You know, in 1997, information technology firms had a certain level of employment, but if you compare the numbers of the Bureau of Labor statistics and the Census Bureau for a year—that was

after all an economic census year—there is a rather sizable discrepancy of—I don't quite remember the number, but between 10 and 15 percent.

So these are the kind of anomalies that hopefully the agencies can start investigating. They can try to figure out why is it that BLS has more workers in this industry. So I think what we are going to have will be data that are much more consistent across our three agencies and hopefully, as I said before, a better reflection of the truth.

Mr. HORN. As a user of Government data, what suggestions would you give to the Federal statistical agencies as they implement the provisions of this bill?

Ms. HAVER. Well, I would say, first of all, they should focus their energies where they are going to have the greatest payoff, and one clear example is the business list. I would like to see one business list, but I do know that Census perhaps has its way of using its list and BLS has yet another way of using its own. So at a minimum I hope we can get to the discrepancies. But what I would like to see would be simply one list.

I think also, and it may be very difficult for our agencies right now given their budget constraints, at least the first markups on the appropriations do not look very good for these agencies. But I would like to see them investigate ways of reducing duplicate surveys so that we can really go to companies and say we are not going to ask this of you more than once, that efforts are really being undertaken in the statistical system to reduce the burden.

I think that would be a big selling point to companies when a survey arrives to be filled out or when they receive a letter asking them to participate in a survey.

Mr. HORN. Thank you, and I think that completes the presentation, unless there is something any of you want to put on.

And Mrs. Maloney, who has been an excellent ranking person, she would like to submit a statement for the record. And without objection, that will be in at this point.

And now I want to thank the staff of the subcommittee, Bonnie Heald, the staff director back there against the wall. And the gentleman doing all of the work here in many ways is the senior counsel, Henry Wray, to my left, your right, and counsel Dan Daly. Where is Dan Daly? Come on, don't be shy. Put your hand up there, fellows. And Chris Barkley, majority clerk, he is over there.

You know, when you get next to a wall, why, there is a problem there. And then minority staff member, David McMillen, professional staff, back here. He gives us a lot of advice. And Jean Gosa, minority clerk, and she is with the staff of the minority. And our court reporters, and we are delighted to have them, and that is, Desirae Jura, and Mark Stuart. Thank you very much. We appreciate all of the work you have done over the years.

So, with that, we thank you all and wish you well.

[Whereupon, at 3:10 p.m., the subcommittee was adjourned.]

[The prepared statement of Hon. Carolyn B. Maloney follows:]

STATEMENT OF THE HONORABLE CAROLYN B. MALONEY OF NY

September 17, 2002

Mr. Chairman, it is a pleasure to again be working with you on this bill. It seems like only yesterday that we introduced the administration's bill on data sharing and statistical confidentiality. The Statistical Confidentiality Act was then, and is today, the foundation for moving the federal statistical system into the 21st century.

Objective indicators point to increasing expense and declining quality of federal statistics. Survey response rates have declined steadily since the early 1980's making them more expensive and less accurate. The lack of current data results in preliminary data on the Gross Domestic Product, that must be revised when better data are available – sometimes turning what looked like economic growth into economic decline.

While the statistical system is being asked to do more with less, and criticized for declining accuracy, it is also subject to greater scrutiny than ever before. The 2000 census was notable, in part, because of the intense media coverage -- more intense than ever before. Alan Greenspan, Chairman of the Federal Reserve, pushed the Consumer Price Index to the front pages when he testified before Congress that errors in that index were costing the government billions.

This confluence of social and political currents pushes the federal statistical agencies to find new ways to measure our social and economic indicators, as well as define new measures.

In short, these agencies need to find new ways of doing business. But to do so, they need new tools.

H.R. 5215 provides the opportunity for agencies to begin charting new ground. This bill provides the framework for the research and experimentation that will define the statistical system for the new millennium.

One of the purposes of the bill is to provide uniform safeguards for the confidentiality of information acquired for exclusively statistical purposes. These provisions were embodied by the legislation introduced by Mr. Sawyer. A second goal is to improve the efficiency of Federal statistical programs and the quality of Federal statistics by permitting limited sharing of records for statistical purposes under strong safeguards.

In short, this bill allows statistical agencies to share information collected from the public to improve statistical measures. It also provides strong safeguards that the privacy of those individuals will be protected, and that the information, once drawn together, will be used only for statistics.

Unfortunately, unlike the version of this bill passed by the House in the 106th Congress, this bill applies only to data collected from businesses. The tools necessary to improve data on the quality of life for Americans has been excluded from this bill at the administration's request. Perhaps it is because of the lingering effects of the recession that we are emphasizing economic statistics. I find this short sighted. Just last weekend, *The Washington Post* reported the disparity between poverty as measured by official statistics, and the real poverty in our nation. Our official measure of poverty is sadly outdated, and the tools originally proposed in this legislation would be valuable in correcting that measure.

It is not just social statistics that will be harmed by this change. Currently, the Census Bureau is conducting innovative research where data from household surveys are being matched

to data from economic surveys. These matched data may prove valuable in understanding the relationship between human capital and productivity. Unfortunately, the provisions in this bill will not allow the full intellectual resources of our statistical system to be brought to bear on this problem, because these data will be excluded from data sharing.

The Administration has put together a bill that lays the foundation for developing new, less burdensome, and less expensive ways of developing statistical information. This bill, for the first time, begins to take a system-wide view of federal statistics. I congratulate my colleague Chairman Horn for his efforts to move this bill, and I look forward to working with him to make it law.

#

○