

政府統計マイクロデータの試行的提供の概要

一橋大学経済研究所附属社会科学統計情報研究センターは、総務省統計局統計調査部からの依頼に応じて、秘匿処理を施した政府統計マイクロデータを、全国の大学研究者に、学術研究のため提供する試行的システムを構築・運用している。

試行的提供は、政府統計マイクロデータ提供の在り方について検討する目的で、平成16年11月から20年9月までの4年間の予定で行っており、その概要は以下のとおりである。

1 提供する調査の範囲

- 就業構造基本調査（平成4年、9年、14年）
- 社会生活基本調査（平成3年、8年、13年）
- 全国消費実態調査（平成元年、6年、11年、16年）
- 住宅・土地統計調査（平成5年、10年、15年）

2 提供するデータの範囲

- 社会生活基本調査・・・生活時間と生活行動のいずれかしか調査されていない世帯員のデータは提供しない
- 全国消費実態調査・・・購入先、曜日別の品目分類のデータは提供しない
世帯票又は家計簿のいずれかが調査されていない世帯のデータも提供しない
- 住宅・土地統計調査・・・平成10年、15年調査の調査票乙のデータは提供しない

3 提供するデータ

提供するデータセットの数は1種類とする。

社会生活基本調査については、生活時間編と生活行動編を組み合わせた分析をできるように、リサンプリングにおいても、生活時間、生活行動ともに同じ世帯（・世帯員）のデータを提供する。

全国消費実態調査については、用途分類、品目分類、貯蓄・耐久財、年間収入、家計資産のデータを結合して、1つのレコード（世帯単位）として提供する。

4 秘匿処理

- (1) 地域区分は、就業構造基本調査、社会生活基本調査及び全国消費実態調査では、「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」、「九州・沖縄」の6地域、住宅・土地統計調査は47都道府県とする。ただし、平成13年社会生活基本調査の調査票Bについては地域区分を付与しない。
- (2) 全データではなく、全体から約8割の世帯を無作為抽出したデータを提供する。ただし、住宅・土地統計調査は約1割の住宅又は住宅以外の建物とする。
- (3) 世帯員が9人以上の世帯は、その世帯全体を削除する。
- (4) 年齢については、80歳以上はすべて80歳に変換する。
- (5) 住宅・土地統計調査について、住宅の居住室数や畳数など、建物の階数はトップコーディングし、建物の建築面積と敷地面積は階級化する。
- (6) 全国消費実態調査と住宅・土地統計調査の乗率は、元の数値ではなく乗率の大きさ

階級別の平均値に変換する。

(7) データの配列順は、乱数を付してソートする方法で無作為化する。

5 利用の条件

- (1) 利用の目的は、学術研究に限定している。教育目的での利用は許可しない。営利企業からの委託研究については利用を許可しない。
- (2) 申請者は、私立大学も含めた全国の大学の講師以上の専任教員とする。申請者は、利用目的に係る研究協力者を共同利用者にすることができる。
- (3) 共同利用者には、大学の講師以上の専任教員のほか、大学の助教・助手、公的研究機関の研究者、大学や独立行政法人の「科学研究費補助金研究者」なども認めている。ただし、大学院生並びに営利企業又は民間（社団、財団等を含む）研究機関等の職員を共同利用者とするは認めていない。

6 提供方法

- (1) センターで年4回原則として1月、4月、7月、10月（それぞれ募集期間は1か月間）に利用者を募集する。
- (2) 利用者はセンターのホームページから申請書をダウンロードして、必要事項を記入し、センターに郵送で申請する。
- (3) センターでまとめて目的外使用申請を行い、承認が得られたら、秘匿処理済マイクロデータ（CD）を作成する。
- (4) 利用者はセンターが開催する説明会に出席し、秘匿処理済マイクロデータ（CD）を受け取る。説明会では、試行的提供の趣旨、データの使用条件、データ利用上の注意事項などを説明する。
- (5) 利用者は使用期間終了後、パソコン内のデータの消去を行い、その消去報告、秘匿処理済マイクロデータ（CD）及び回答を記入したアンケート用紙をセンターに返送する。
センターでは返却されたCDをまとめて裁断処置する。

7 提供するデータの形式

データは、CSV形式で提供する。

符号表でデータの内容を明示し、先頭のレコードには符号表に記載している変数名を入れる。

8 データの使用条件

- (1) データの使用期間は6か月とする。
- (2) マイクロデータを取り扱う際には、ネットワークに接続していないパソコンをなるべく使用する。

9 研究論文

利用者は作成したマイクロデータによる研究論文をセンターに2部提出する。提出された研究論文については、随時、試行的提供のホームページ上で公開する。

政府統計ミクロデータの試行的提供における利用状況

平成16年11月から平成19年7月までの募集における申請件数は計66件で、利用者数は総計92名（うち共同利用者は26名）である。

	16年11月募集	17年4月募集	17年10月募集	18年4月募集
申請件数	6件	4件	12件	9件
(利用者別)				
申請者	6人	4人	12人	9人
共同利用者	1人	-	4人	5人
(調査別)				
就調	1件	2件	4件	7件
社会調	1件	-	1件	-
全消	4件	3件	8件	3件
(分析方法)				
表のみ	2件	-	3件	-
表+多変量解析	2件	3件	5件	8件
多変量解析のみ	2件	1件	4件	1件
(大学)				
国公立大学	1件	2件	5件	6件
私立大学	5件	2件	7件	3件

	18年11月募集	19年4月募集	19年7月募集	合計
申請件数	20件	5件	10件	66件
(利用者別)				
申請者	20人	5人	10人	66人
共同利用者	8人	1人	7人	26人
(調査別)				
就調	11件	-	4件	29件
社会調	2件	2件	3件	9件
全消	8件	3件	2件	31件
住調	2件	-	1件	3件
(分析方法)				
表のみ	6件	2件	4件	17件
表+多変量解析	12件	2件	5件	37件
多変量解析のみ	2件	1件	1件	12件
(大学)				
国公立大学	13件	2件	5件	34件
私立大学	7件	3件	5件	32件

(申請書様式)

秘匿処理済マイクロデータ使用申請書

一橋大学経済研究所附属社会科学統計情報研究センター長 殿

平成____年____月____日

申請者(所属・職名) _____

(氏名:自署) _____ 印

私(及び共同利用者)は、この度、一橋大学経済研究所附属社会科学統計情報研究センター(以下、「センター」という。)を通して秘匿処理済マイクロデータの使用を申請します。申請が承認されデータを使用するときには、下記使用条件を厳守することを誓約します。

なお、私にあって、下記使用条件に違反があった場合、今回いただく承認が取り消され、さらに、私(及び共同利用者)が今後センターを通しての秘匿処理済データの使用を一切許可いただけなくなっても不服は申しません。また、私(及び共同利用者)の氏名と違反の内容をセンターのホームページ等に掲載されても結構です。

秘匿処理済マイクロデータの使用条件

- 1 統計法の規定を遵守し、使用する秘匿処理済マイクロデータの管理は厳正に行うこと。
 - ・提供されたCD等は、使用期間中、鍵のかかる安全なところに保管すること。
 - ・提供されたCD等は、申請した使用場所以外では使用しないこと。
 - ・提供されたデータは本申請書に署名した者だけが使用し、第三者に再提供若しくは使用させないこと。
 - ・提供されたデータは本申請書に記載した集計目的だけに使用し、他の目的では使用しないこと。
 - ・パソコン等へのマイクロデータの蓄積は、一時的な場合を除いて行わないこと。
- 2 使用期間終了後、提供されたデータを直ちにセンターに返却すること。また、パソコン内のコピー等をすべて速やかに消去し、消去した旨の報告書をセンターに提出すること。
- 3 調査に関する質問等はセンターに対して行うこと。
- 4 分析結果を発表する際には、センターを経由して秘匿処理済マイクロデータの提供を受けたことを明記すること。
- 5 作成した論文等を2部センターに提出すること。
- 6 その他センターの指示に従うこと。
- 7 提供された秘匿処理済マイクロデータの使用により何らかの不利益を被ったとしても、総務省統計局統計調査部及びセンターの責任は一切問わないこと。

私は、この度、当学部(又は研究科、研究所、センター等)の専任教員である申請者が、貴センターに秘匿処理済マイクロデータの使用を申請することを承認します。

平成____年____月____日

所属長(所属・職名) _____ (氏名) _____ 公印

申請事項

1 指定統計調査の 名称及び年次	
2 調査票の使用目的	<p>研究計画の名称（20字以内にまとめてください。）</p> <p>_____</p> <p>結果の発表の仕方 報告書等 論文 研究会等で発表</p> <p>研究目的、研究方法等を以下に簡潔に記載してください。</p>
3 集計様式	別添のとおり。
4 研究資金の出所	<p>資金なし 資金交付を受けた（出所・名称を記載してください。）</p> <p>（出所：_____ 名称：_____）</p>
5 共同利用者 （氏名は自署）	<p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p> <p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p> <p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p>
6 使用場所	
7 申請者連絡先	<p>メール ：_____</p> <p>電話番号：_____</p>

注) この用紙に入りきらないときは、適当な用紙を添付して記入してください。

(申請書記入要領)

申請書記入要領

申請書の書き方で分からないことがあれば、本センターにメールで質問してください。

申請するときの注意

使用条件と使用条件に違反したときのペナルティーを確認した上で申請してください。申請書の記載に事実でないことがあった場合、使用条件に違反したものとみなすことがあります。

申請者及び共同利用者の資格には十分に注意してください。なお、利用者全員の署名(自署)捺印と、所属長(学部長、研究科長、研究所長、センター長等)の公印の押印が必要です。

1 指定統計調査の名称と年次

秘匿処理済マイクロデータの調査名と使用する年次を記載してください。使用するファイルの種類が分かるように、全国消費実態調査の場合は二人以上の一般世帯(又は普通世帯)と単身世帯の別を、社会生活基本調査の場合は生活時間と生活行動の別を記載してください。

なお、住宅・土地統計調査については、平成5年の調査名は住宅統計調査と記載してください。

2 調査票の使用目的

研究計画の名称を20字以内で記載し、予定している結果の発表の仕方すべてをまるで囲んでください。発表を予定していない場合、使用は認めません。さらに、提供されたデータでどのような研究を行うかが分かるように研究目的と研究方法を簡潔に説明してください。

3 集計様式

集計様式の書き方を参考に明確に記載してください。明確に記載されていない場合、使用を許可しないことがあります。また、本センターから修正を求めることがあります。

4 研究資金の出所

特に研究資金がない場合は、「資金なし」をまるで囲んでください。この項目は、営利企業からの委託研究でないことを確認するためのものですので、研究資金の提供者と資金の名称が分かるように明確に記入してください。大学から交付

された研究費の場合、「大学からの研究費」と書いてください。なお、営利企業からの研究費であっても、純粋な学術奨励のための研究補助金の場合は使用を認めます。

5 共同利用者

共同利用者の所属・職名と氏名を記入してください。氏名は自署であることを要します。「科学研究費補助金研究者」に該当する場合は、研究者番号も記入してください。

6 使用場所

集計を行う場所を、例えば、「 研究室」というように記入してください。

7 申請者連絡先

本センターからの問い合わせに使用させていただきます。申請者ご自身が所属する大学のメールアドレス及び研究室の電話番号にしてください。なお、センターからの連絡は原則としてこのメールアドレスに送りますので、使用期間中はいつも確認するようにしてください。

所属長の承認

申請に当たっては、所属長の承認が必要です。所属長の所属・職名、氏名を記載の上、公印を押印してください。

平成 年 月 日

総務大臣 殿

大学 学部 職名

氏 名 印

誓 約 書

「平成 年 調査」調査票の統計法第 15 条第 2 項の規定に基づき、目的外使用に当たっては、統計法第 14 条の規定を遵守し、調査票を適正に管理するとともに、秘密漏洩のないよう厳重に注意します。

見本

共同利用者がいる場合は、一人一人別に作成

誓約書を作成した日

平成18年11月27日

総務大臣 殿

一橋大学経済研究所附属社会科学統計情報研究センター 助教授

所属・職名

山口 幸三

自署

捺印

印

誓約書

調査年次、調査名

「平成元年、6年、11、16年全国消費実態調査」調査票の統計法第15条第2項の規定に基づく目的外使用に当たっては、統計法第14条の規定を遵守し、調査票を適正に管理するとともに、秘密漏洩のないよう厳重に注意します。

符 号 表

別紙4

作成日	2006/7/25	訂正日	2007/7/27	調査名	平成16年全国消費実態調査
調査年次	2004年(平成16年)			データ名	Zenshou2004F(二人以上の一般世帯),Zenshou2004T(単身世帯)
備考					

	項目名	変数名	符号	備考
1	調査ID	SurveyID	Zenshou	
2	調査年	SurveyYear	2004	
3	センター使用欄	Centre		
4	レコード連番号	RecNo	0000001~	
集計区分別データの有無				
5	家計収支編(用途分類・品目分類)	KakeiDataUmu	1:あり 0:なし	
6	貯蓄編	ChochikuDataUmu		
7	耐久財編	TaikyuuDataUmu		
8	家計資産編	ShisanDataUmu		
9	地域符号	District	1:北海道・東北 2:関東 3:北陸・東海 4:近畿 5:中国・四国 6:九州・沖縄	
10	世帯符号	SetaiCode	0000001~	
11	乗率	Weight	nnnnn.n(プランク補足,小数点以下1桁まで表章)	
12	一般・単身の別	SetaiBetsu	F:二人以上の世帯 T:単身世帯	
調査票の有無				
13	家計9月	Kakeibo_Sep	1:有り :無し	
14	家計10月	Kakeibo_Oct		
15	家計11月	Kakeibo_Nov		
16	年収票の有無	Nenshuuhyou	1:有り :無し	
17	年間収入	NenshuuFushou	V:年間収入不詳有り :年間収入不詳無し,年収票無し	
18	貯蓄	ChochikuFushou	V:貯蓄に不詳あり :貯蓄に不詳無し,年収票無し	
19	耐久財票の有無	Taikyuuzaihyou	1:有り :無し	
20	調査月数	ChousaTsukisuu	1~3:1~3か月	
21	世帯区分	SetaiKubun	1:勤労 2:勤労及び無職以外 3:無職	
22	世帯人員	SetaiJinin	1~ 8:1~8人	
23	就業人員	ShuugyouJinin	0~ 8:0~8人	
家計を主に支える人				
24	世帯主との続柄	FuzaiTsunami	2:配偶者 3:子(未婚) 4:子(夫婦であるもの) 5:父母 6:兄弟(未婚) 7:兄弟(夫婦であるもの) 8:孫(未婚) 9:孫(夫婦であるもの) 0:祖父母 X:2~0以外の親族 :3か月以上の不在者なし	
25	不在理由	FuzaiRiyuu	1:単身赴任 2:出稼ぎ 3:入院 4:その他 :3か月以上の不在者なし	
家計を主に支える人以外の不在者の不在理由別人数				
26	入院	FuzaiNyuuin	1~9:1~9人 :3か月の不在者なし	
27	学業	FuzaiGaku		
28	その他	FuzaiTa		
29	子の住んでいる場所	KoDoukyo	1:同居(生計同一) 2:同居(生計別) 3:同敷地内 4:近く(徒歩5分程度) 5:片道1時間以内 6:片道1時間以上 7:子はいない	
30	単身世帯の形態	Tanshin	1:単身赴任 2:出稼ぎ 3:その他 :一般世帯	
31	住宅の所有関係	Shoyuu	1:持ち家(世帯員名義) 2:持ち家(世帯員以外の名義) 3:民営賃貸住宅(設備専用) 4:民営賃貸住宅(設備共用) 5:市区町村営賃貸住宅 6:公団公社賃貸住宅 7:社宅・公務員住宅(借り上げふくむ) 8:借間 9:寮・寄宿舎	
32	住宅の構造	Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:ブロック造 5:その他	
33	住居の建て方	Tatekata	1:一戸建 2:長屋建 3:共同住宅(1~2階建て) 4:共同住宅(3~5階建て) 5:共同住宅(6~10階建て) 6:共同住宅(11階建て以上) 7:その他	
34	水洗式トイレの有無	Suisen	1:有り 2:無し	
住宅面積				
35	延べ床面積	NobeYuka	10~99999:1.0~9999.9㎡	下一桁が小数第一位
36	うち業務用面積	GyoumuyouYuka	:業務用無し 10~99999:1.0~9999.9㎡	下一桁が小数第一位
37	持家の場合	Chidai	1:世帯員名義の家で無し 2:世帯員以外の名義で無し 3:有り :持ち家以外	
38	住宅の敷地面積	Shikichi	10~99999:1.0~9999.9㎡	下一桁が小数第一位
39	建築年	KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 V:不詳 :昭和20年以前又は持ち家以外	
40	持ち家以外の入居年	NyuukyoYear	1~11:平成元~11年 :昭和63年以前又は持ち家	
41	現住居以外の住宅の有無	TanoJuutakuUmu	1:有り 2:無し :「単身赴任」または「出稼ぎ」(単身世帯)	
42	用途	J1_Youto	1:親族居住用 2:賃貸用 3:その他 :入力無し	
43	建築年	J1_KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 V:不詳 :昭和20年以前又は入力無し	
44	延べ床面積	J1_NobeYuka	10~99999:1.0~9999.9㎡	下一桁が小数第一位
45	構造	J1_Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:ブロック造 5:その他 :入力無し	
現住居以外の住宅及び土地について				
46	用途	J4_Youto	1:親族居住用 2:賃貸用 3:その他 :入力無し	
47	建築年	J4_KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 V:不詳 :昭和20年以前又は入力無し	
48	延べ床面積	J4_NobeYuka	10~99999:1.0~9999.9㎡	下一桁が小数第一位
49	構造	J4_Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:その他 :入力無し	

データ利用上の注意事項（全国消費実態調査）

1 分析に当たっての全般的な注意事項

- (1) 分析に当たっては、報告書で調査方法、標本設計の方法等を十分に理解するとともに、調査票様式をよく確認してください。報告書に掲載されている結果表による分析の場合にはそれほど問題ありませんが、マイクロデータを分析する場合にはこれらのことを十分に理解しておかないと、誤りをおかしてしまうおそれがあります。
- (2) この調査の調査期間は、二人以上の一般世帯の場合は9～11月、単身世帯の場合は10～11月です。したがって、ボーナス月は含まれていないなどの問題があり、年間の収支とみて分析することはできません。
- (3) 提供データには乗率が付されています。そのことを理解して、分析するようにしてください。例えば、平均値を求めるときには加重平均する必要があります。
- (4) 提供データにはトップコーディングやリサンプリングなどの秘匿処理を施してあります。このため、提供データによる集計結果を公表されている結果と完全に一致させることはできません。なお、秘匿処理の方法は「参考1」のとおりです。
- (5) 多変量解析を行うときには、外れ値を考慮するように注意してください。調査票の記入誤りなどはチェック・修正済みですが、外れ値を除くような処理は行っていません。

2 データの形式等について

- (1) データはCSV形式です。先頭のレコードに変数名を入れてあります。
- (2) 1世帯1レコードになっています。用途分類、品目分類、貯蓄・耐久財、年間収入、家計資産のデータを1レコードにまとめています。
- (3) 世帯票と家計簿（1か月分以上）の両方が提出されている世帯のデータだけを提供しています。世帯によっては他の調査票が提出されていない場合もありますが、その場合、該当項目の欄は欠損値（.）、不詳（V）、0又は空白になっています。調査票の提出状況については、「集計区分別データの有無」の欄及び「調査票の有無」欄で確認してください。
- (4) データの符号の付け方等については、符号表を参照してください。特殊な符号の付け方になっていることもありますので、必ず確認してから分析してください。
- (5) データ中の世帯属性等の欄については0補足を行うなどして桁数をそろえてありますが、用途分類等の欄については必要桁数だけにしています。
- (6) 当該マイクロデータが分析の対象であることを確認してから集計するようにしてください。必要のない欄に符号が付いていること（過剰記入）もあります。

(7) 各年次のデータ数については、「参考2」のとおりです。

3 データの内容について

- (1) 用途分類及び品目分類の数値は、調査期間中の収支金額を調査月数で除した1か月当たりの平均値にしています。提供するデータでは整数になるよう四捨五入しています。調査月数は用途分類の「調査月数」欄に入っています。なお、総務省統計局で行っている用途分類及び品目分類の集計では、一般の（普通）世帯の場合、2か月しか調査していない世帯の乗率は3分の2に、1か月しか調査していない世帯の乗率は3分の1にしています。単身世帯の場合及び集計世帯数の集計の場合も同様な考え方で集計されています。
- (2) 平成元年のデータには家計資産はありませんが、用途分類の末尾に「住宅・宅地資産額」と「耐久消費財資産額」だけが入っています。平成6年、11年及び16年のデータにおいて、耐久財・貯蓄等の調査票が未提出などの理由で資産データが作成されていない場合、家計資産額はすべて欠損値（.）にしてあります。
- (3) 年間収入調査票が提出されていない場合、世帯属性の年間収入の欄は0又は不詳（V）となっています。ただし、用途分類の欄には総務省統計局で推計した年間収入の値（世帯の合計年間収入）が入っています。
- (4) 乗率は、平成11年、16年の単身世帯の場合は比推定済乗率、他はすべて線形推定乗率になっています。乗率は小数点以下1桁まで表示しています。
- (5) 地域は「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」及び「九州・沖縄」の6区分にしています。
- (6) 平成元年の普通世帯のデータでは、「世帯員情報」の欄（の多分、最後）に不在世帯員の情報が入っていますので、「行番号」の欄で確認した上で利用するようにしてください。
- (7) 「住宅の延べ床面積」や「宅地の敷地面積」の欄は、平成元年の場合は整数の値が入っていますが、平成6年、11年及び16年の場合は10倍した値（小数点以下1桁目までの値）が入っています。ただし、小数点以下1桁目が0のとき、0ではなく空白になっていることもあります。

お願い

マイクロデータの利用において疑問が生じたときには、当センターにメールでご質問ください（micro@ier.hit-u.ac.jp）。なるべく早くお答えするようにします。

また、今回の提供データの作成に当たっては、十分に注意したつもりですが、思わぬ誤りが残っているかも知れません。データを利用して疑問が生じたときには、当センターにメールでご連絡いただくようお願いいたします。

参考1 秘匿処理の方法

- (1) 地域は「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」及び「九州・沖縄」の6区分にしています。
- (2) 全データではなく、全体から約8割の世帯を無作為抽出したデータを提供します。
- (3) 世帯員が9人以上の世帯は、その世帯全体を削除しています。
- (4) 世帯の配列順は、乱数を付してソートする方法で無作為化しています。
- (5) 年齢については、80歳以上はすべて80歳に変換しています。

参考2 データ数

年次	一般の（普通）世帯	単身世帯
平成元年	44,778	3,288
平成6年	44,803	3,772
平成11年	44,537	4,013
平成16年	44,006	4,001