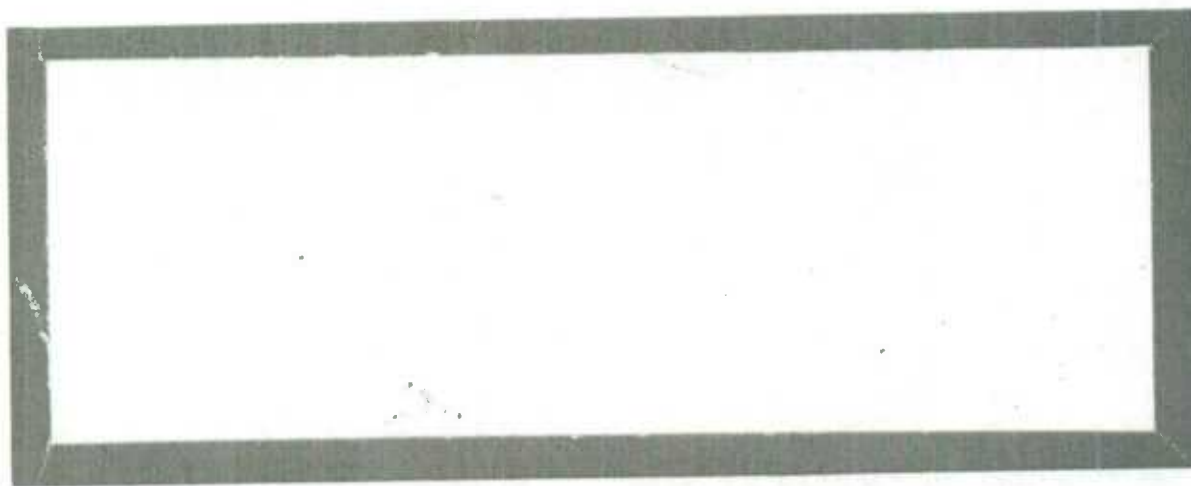




Statistics
Canada

Statistique
Canada



Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes
entreprises

11-617

no. 85-35

c.2

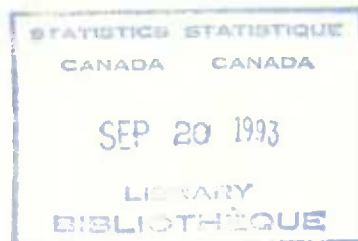
Canada

50480

A Report on the Quality of the Data in the
Business Register Master File

V. Estevao
July, 1983

Working Paper No. BSMO 85-035E



1911

Résumé

Le Fichier principal du registre des entreprises (FPRE) est un ensemble centralisé d'enregistrements portant sur les caractéristiques des entreprises. Il a été conçu dans le but d'établir des bases de sondage pouvant servir aux enquêtes-entreprises de Statistique Canada.

La qualité des données revêt une grande importance dans la construction d'une base de sondage. Avant 1983 la qualité des données que contient le FPRE n'avait jamais été évaluée de façon fiable. Une étude a donc été réalisée entre les mois d'août 1982 et de mai 1983. Les résultats sont présentés dans le rapport.

Introduction

One objective of the Business Register Management and Quality Assurance Program is to measure the quality of the data on the Business Register Master File and to establish an on-going program of quality assurance.

This report presents the results of a study conducted between August 1982 and May 1983 to determine the quality of the data on the Business Register Master File.

1. Summary of Results

This section presents a brief summary of the results of the study. Specific details are given in the following sections. Estimates at the industry group level are provided in Appendix B.

In presenting the results of this study it is important to note that they are applicable to a restricted, although rather large subset of the Business Register Master file. This subset was first determined as consisting of all active, single establishment, reference units that are classified. As of September 1982, this population totalled approximately 858,000 units on the Master File. Further restrictions noted in Section 3, reduced the population under study to about 800,000 units, classified into 13 industry groups. These groups formed the strata in the sample design and the levels for publication of the estimates.

For the population given above, it is estimated that about 17% (approx. 138,000) of the units, classified as active on the file, are in fact inactive. By industry group, Construction has the highest estimated percentage of inactive units (23% of the industry or approx. 25,000 units).

Within the estimated 662,000 active units 7% (44,000) were calculated to have undergone some type of change in their legal, financial or corporate structure, since the information for them was recorded on the BRMF. These changes are primarily associated with incorporations of existing businesses and changes of ownership. The active units are classified into 2 groups:

1. units with major legal changes (defined as changes of ownership, mergers or amalgamations)
2. units with no major legal changes (defined as incorporations, corporations with charter amendments and corporations with no legal, financial or corporate changes).

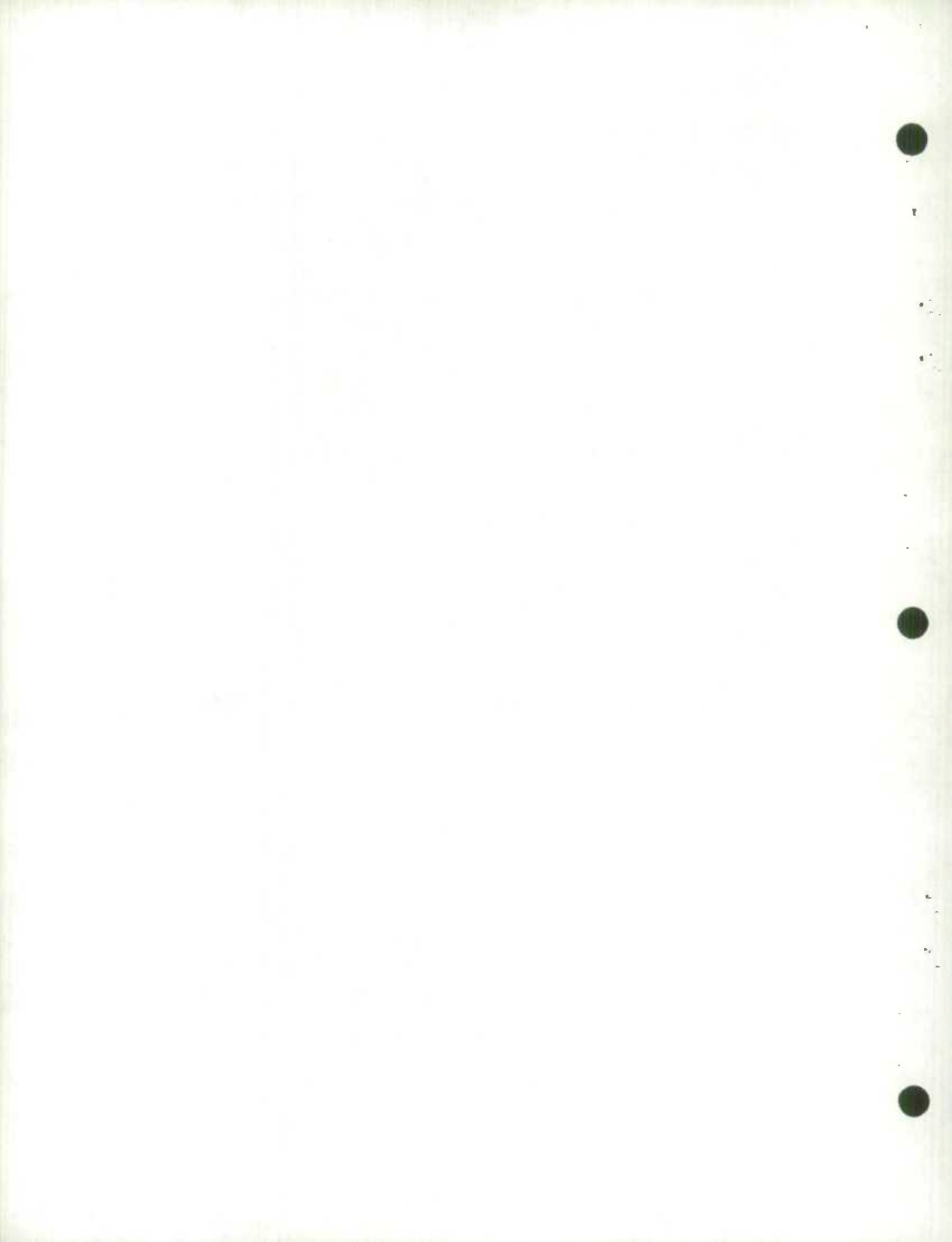
The first group was not considered for further analysis. The second group formed a domain estimated to account for over 96% (approx. 641,000) of the population of active units. It was for this domain, consisting of units with no major legal changes, that estimates were made on the quality of individual data items. Some of these estimates are given below:

1. About 23% (147,000) of the units are estimated to have a 3 digit 1970 SIC code different from that on the BRMF. This difference may be due to inaccuracy in the original BRMF code assignment, genuine change in industrial activity since the time of the assignment, insufficient information to make a 3 digit assignment or a combination of these factors. Of the 147,000 units differing in 3 digit SIC code, it is estimated that 45% (67,000) of these also differ in the industry group classification (strata).

2. Approximately 38% (246,000) of the units have an employment size code in disagreement with that on the BRMF. About 52% (128,000) of these differences involve units classified to size code 01 (maximum expected number of employees is 1 to 4) on the BRMF.

In terms of SEPH size groups, the proportion of disagreement is estimated to be about 9% (56,000).

3. Approximately 7% (42,000) of the units have a BRMF label address which represents neither the location of activity nor a valid mailing address. For an additional 86% (548,000), the BRMF label address represents the activity location, and of these about 2% (10,000) are estimated to have an incorrect BRMF Standard Geographical Classification (SGC) code.



2. Universe Definition

The universe for this study consisted of all active single reference units on the Business Register Master File as of Cycle 131 (September 1982). These units were identified by the conditions, MODE=0 (reference record), SSC=1 (active record), SUB-UNIT=0000 (unit level record) and by determining that no corresponding reference sub-units were associated with the unit records. However, these unit level records were allowed to have non-reference sub-units or sub-sub-units. This amounted to an identification of 858,413 unit records.

3. Stratification, Allocation and Sample Selection

The universe thus defined, was divided into strata based on an aggregation of 1970 SIC codes. All units identified as government, not public administration, formed a separate stratum, without regard to SIC code. They were not considered for sample selection.

The strata were labelled as industrial groups. Although 20 groups were identified, only 13 were sampled in the study. The excluded strata, for example, were made up of units with invalid or unclassified SIC codes and units pertaining to households or religious organizations. The restricted universe under study totalled 800,069 unit level records. The group definitions for the 13 sampled strata are given in Table 1.

The maximum sample size, for which data could be collected by Regional Office (RO), was set at 8200. An allocation was chosen which determined the sample size in each stratum to be proportional to the square root of the number of units in the stratum. With the given allocation a single random sample of units was drawn from each stratum. These units

were identified by an 8 digit BRID number from the BRMF. A list of these units was then sent to Business Register Division for examination and further action.

Table 2 presents the results of the allocation procedure.

Table 1. Strata Definition in Terms of BRMF SIC70 Codes

(SEPH groups also indicated.)

*need to define SEPH
& what SEPH Group are*

Stratum 1: Agriculture (Not covered by SEPH)

001, 003, 011, 013, 015, 017, 019, 021

Stratum 2: Forestry (SEPH Group 10)

031, 039

Stratum 3: Fishing and Trapping (Not covered by SEPH)

041, 045, 047

Stratum 4: Mining (SEPH Group 20)

051, 052, 057-059, 061, 064, 071-073, 079, 083,

087, 096, 098, 099

Stratum 5: Manufacturing

SEPH Group 31 101-109, 151, 153, 162, 165, 172, 174
175, 179, 181-189, 231, 239, 243-246,
248, 249, 271-274, 286-289, 365, 369,
372-379, 391-393, 397, 399

SEPH Group 35 251, 252, 254, 256, 259, 261, 264,
266, 268, 291, 292, 294-298, 301-309,
311, 315, 316, 318, 321, 323-329,
331-336, 338, 339, 351-359

Stratum 6: Construction

SEPH Group 41 404, 421

SEPH Group 45 406, 409

Stratum 7: Transportation (SEPH Group 51)

501-509, 512, 515-517, 519, 524, 527

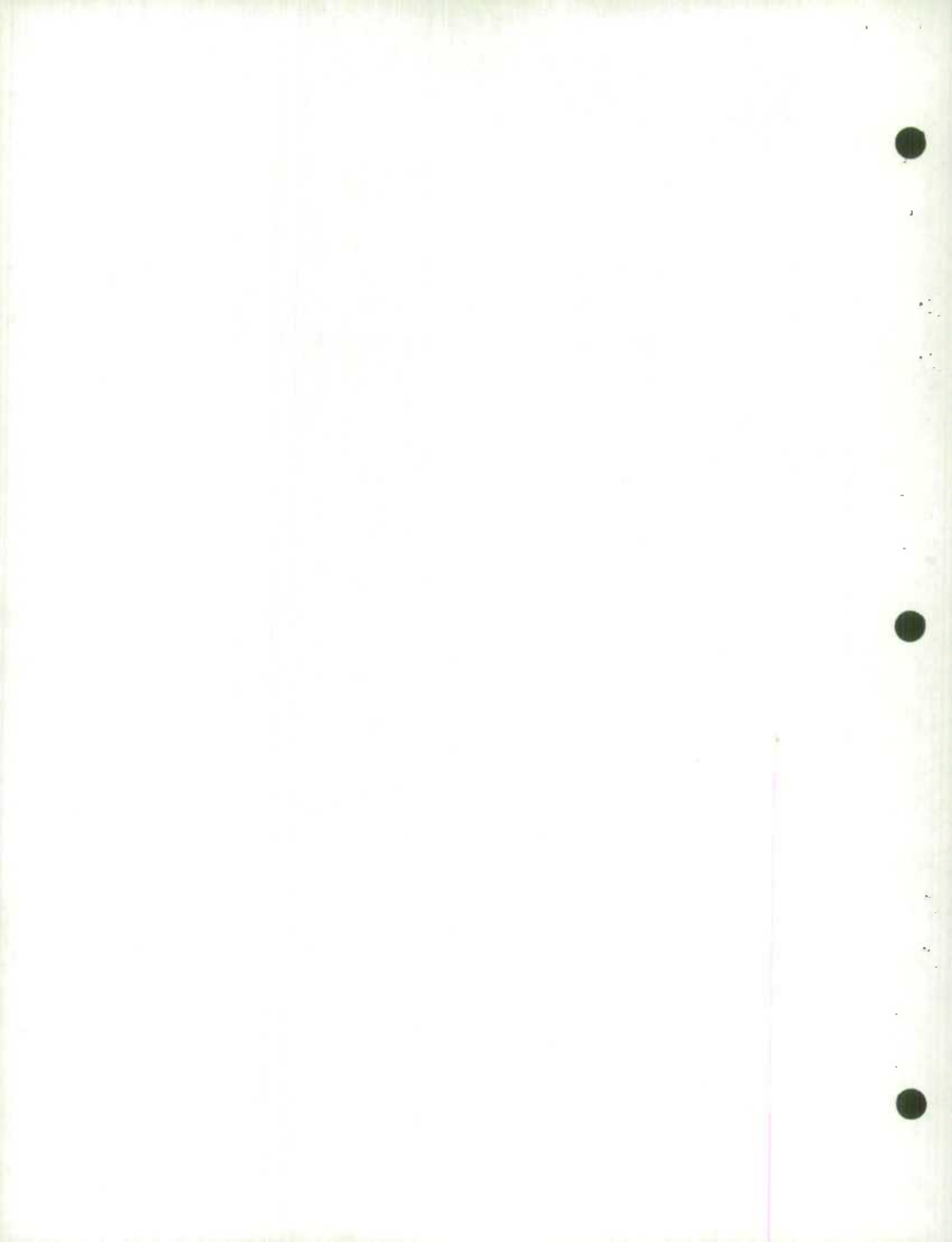


Table 1. (Continued)

<u>Stratum 8:</u>	<u>Communication and Other Utilities</u>
	<u>SEPH Group 54</u> 543-545, 548
	<u>SEPH Group 57</u> 572, 574, 576, 579
<u>Stratum 9:</u>	<u>Wholesale Trade (SEPH Group 61)</u>
	602, 606, 608, 611, 612, 614-619, 621-627, 629
<u>Stratum 10:</u>	<u>Retail Trade (SEPH Group 65)</u>
	631, 642, 652, 654, 656, 658, 663, 665, 667
	669, 673, 676, 678, 681, 691, 692, 694-697, 699
<u>Stratum 11:</u>	<u>Finance, Insurance and R. Estate (SEPH Group 70)</u>
	701, 703, 705, 707, 721, 735, 737
<u>Stratum 12:</u>	<u>Education and Health Services</u>
	<u>SEPH Group 81</u> 801-806, 809
	<u>SEPH Group 83</u> 807, 821-828
<u>Stratum 13:</u>	<u>Other Services (SEPH Group 87)</u>
	841-845, 849, 851, 853, 855, 861-864, 866, 867,
	869, 871, 872, 874, 876, 877, 879, 881, 883,
	884, 886, 891, 893-899

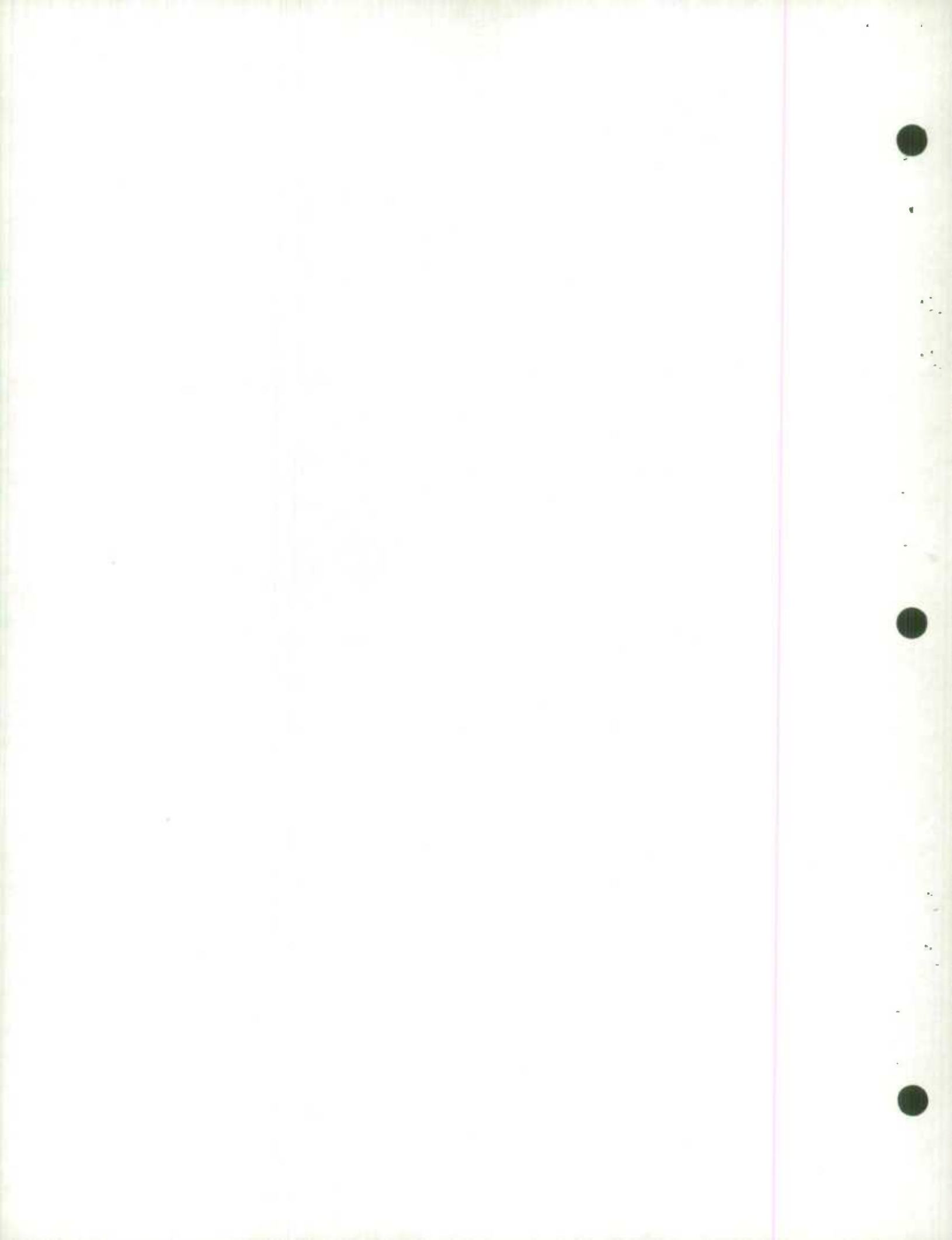


Table 2. Allocation For a Sample Size of 8200 Units

Strata: defined by industry group- 13 strata

Method: $n_h \propto \sqrt{N_h}$ $h = 1, 2, \dots, 13$

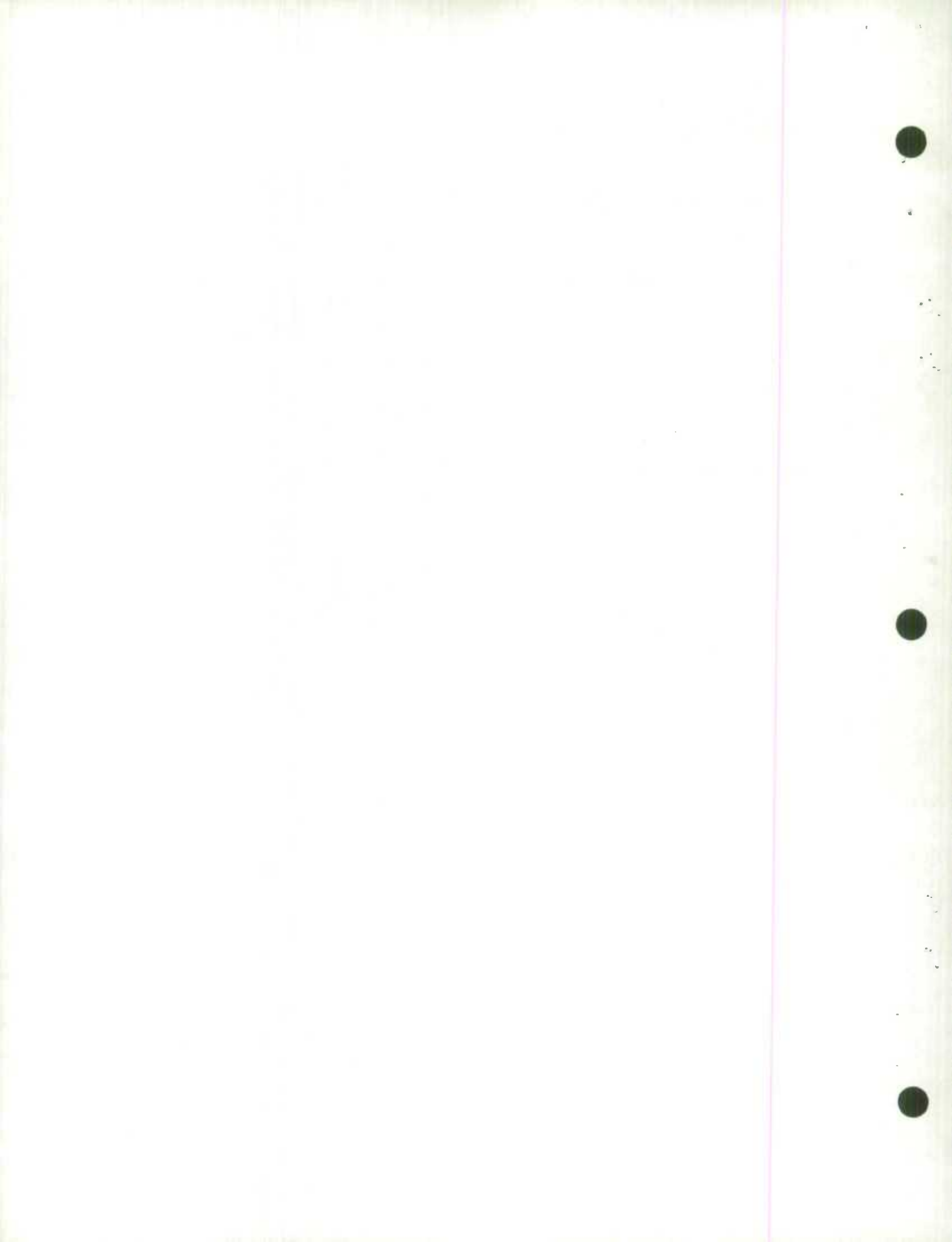
where n_h = allocated sample size of hth stratum

N_h = population size of hth stratum

and $\sum_h n_h = 8200$

Results:

<u>Industry Group</u>	<u>N_h</u>	<u>n_h</u>
Agriculture	74236	795
Forestry	7633	255
Fishing and Trapping	5049	207
Mining	4983	206
Manufacturing	46270	627
Construction	109329	964
Transportation	33315	532
Communications	2785	154
Wholesale Trade	52272	667
Retail Trade	158198	1160
Fin., Ins. and R. Estate	61370	723
Educ. and H. Services	41891	597
Other Services	202738	1313
Total	800069	8200



4. Data Collection, Capture and Control

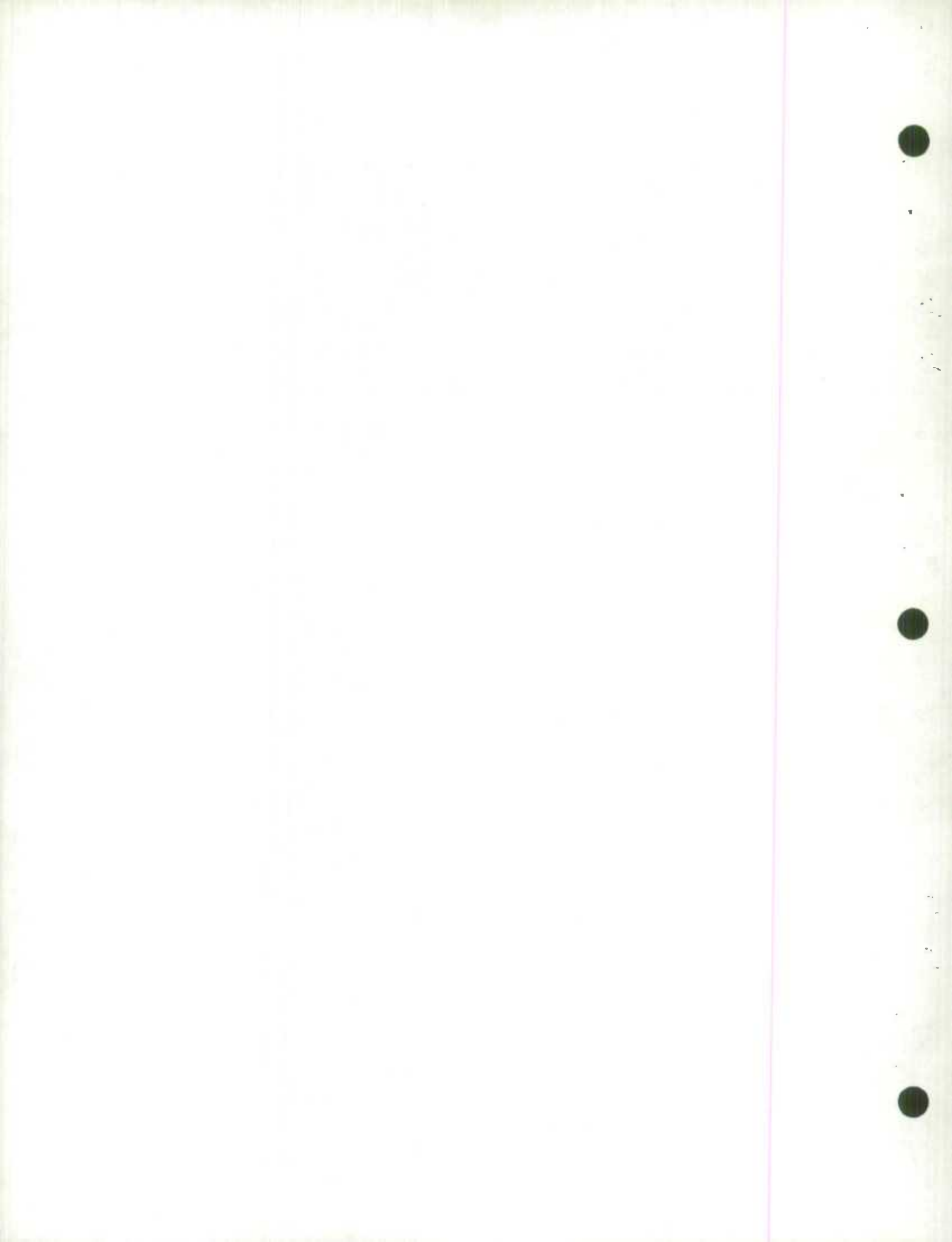
The selected units were examined and processed through the Company Contact System. One of the functions of this system is to identify those units which have already been contacted within the previous 12 months. Units may then be excluded from further contact in order to alleviate response burden. From the 8200 units selected, 611 were blocked by the system. At this stage, the members of the study group decided to exclude these units from data collection. It was agreed that for the purpose of the study, the information on the BRMF for these 611 units would be regarded as correct. The justification for this course of action is presented in Section 6. For the remaining units, a label was produced for identification and they were then sent to Regional Office for data collection using the CBS-718 form.

A data control sheet was designed to capture the information on the CBS-718 form that was pertinent to the study. Various fields were established and a set of instructions was developed to assist the clerks in the coding of the fields. A quality control scheme was designed to monitor the coding of the various fields.

The information on the data control sheets was then entered through Key-edit and stored on a file. Subsequently, the data was processed through an edit system which checked for duplication of BRID, inadmissible entries in various fields and inconsistencies in the relationships between fields. These errors were examined and the file was updated accordingly. For those fields where data was required and missing, every effort was made to obtain the necessary information. In the final version of the file only two fields, employment size code and current number of employees, had some missing entries for records where data was required. However, it is important to distinguish the



the above situation from that arising from the design of the coding procedure whereby, according to the type of record, certain fields did not require an entry or code. For example, if it was found that a unit was involved in a change of ownership, merger or amalgamation, then the employment size, 1970 SIC and various other fields were not coded. The reason for this action is that, as a result of the changes noted above, a unit may no longer exist in its perceived or original form. Therefore, it may not be appropriate to attempt to determine such characteristics as SIC without a detailed analysis of the actual corporate structure or structures resulting from such changes.



5. Estimation

5.1 Preliminary Notes

The coding procedure and the data collection process itself generate a hierarchical structure on the determination of the various fields and characteristics. For example, in table 3, we note that at the top of the data collection process, detailed information may only be obtained for those units which are first determined to be active. If a unit is found to be inactive then no further information is collected. This leads to the establishment of particular domains of study within which specific fields or characteristics may be studied. Using the example from the previous section, employment size and 1970 SIC codes are only analyzed for that domain consisting of active units which have not been involved in changes of ownership, mergers or amalgamations.

Table 3 presents a classification of the sample units according to specific observed characteristics. The first level of classification is the activity status. If a unit is active, then data for it may or may not have been collected. This is then the second level of classification, etc. Information for a specific field or characteristic is obtained at a particular level of classification. Those units captured by the Company Contact System are assumed to be active, with no legal, financial or corporate changes. Furthermore, it is assumed that the BRMF address for these units refers to the location of activity (it may also be the mailing address). These assumptions are made in accordance with those given in the previous section.

For the most part, each of the levels leads directly to a domain of study, as given in Table 4. However, there are 2 exceptions, which differ somewhat in nature. The first exception is the additional domain consisting of active units with no changes of a legal, financial or corporate nature. It is nested within the domain comprised of units with no major legal changes.

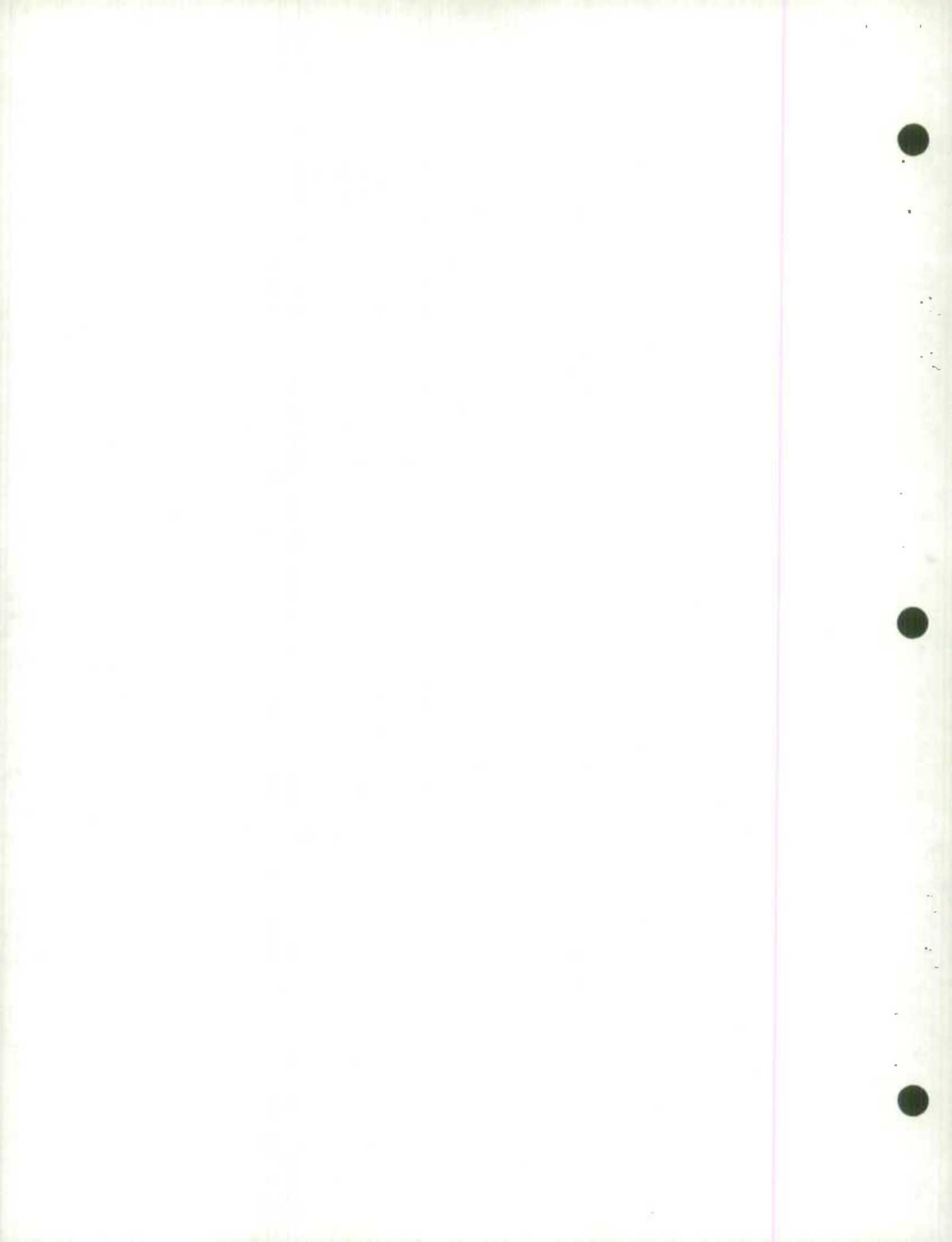
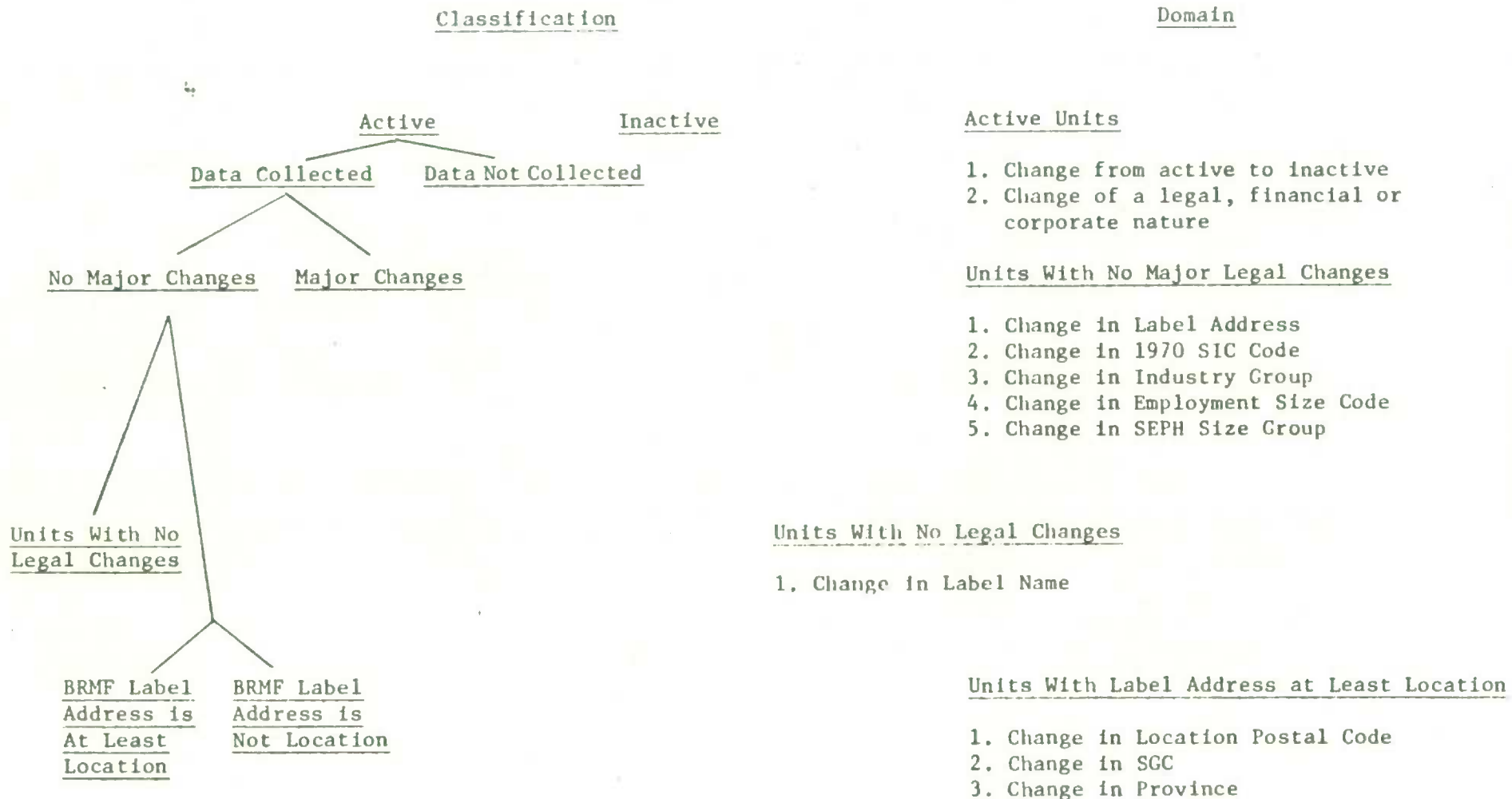
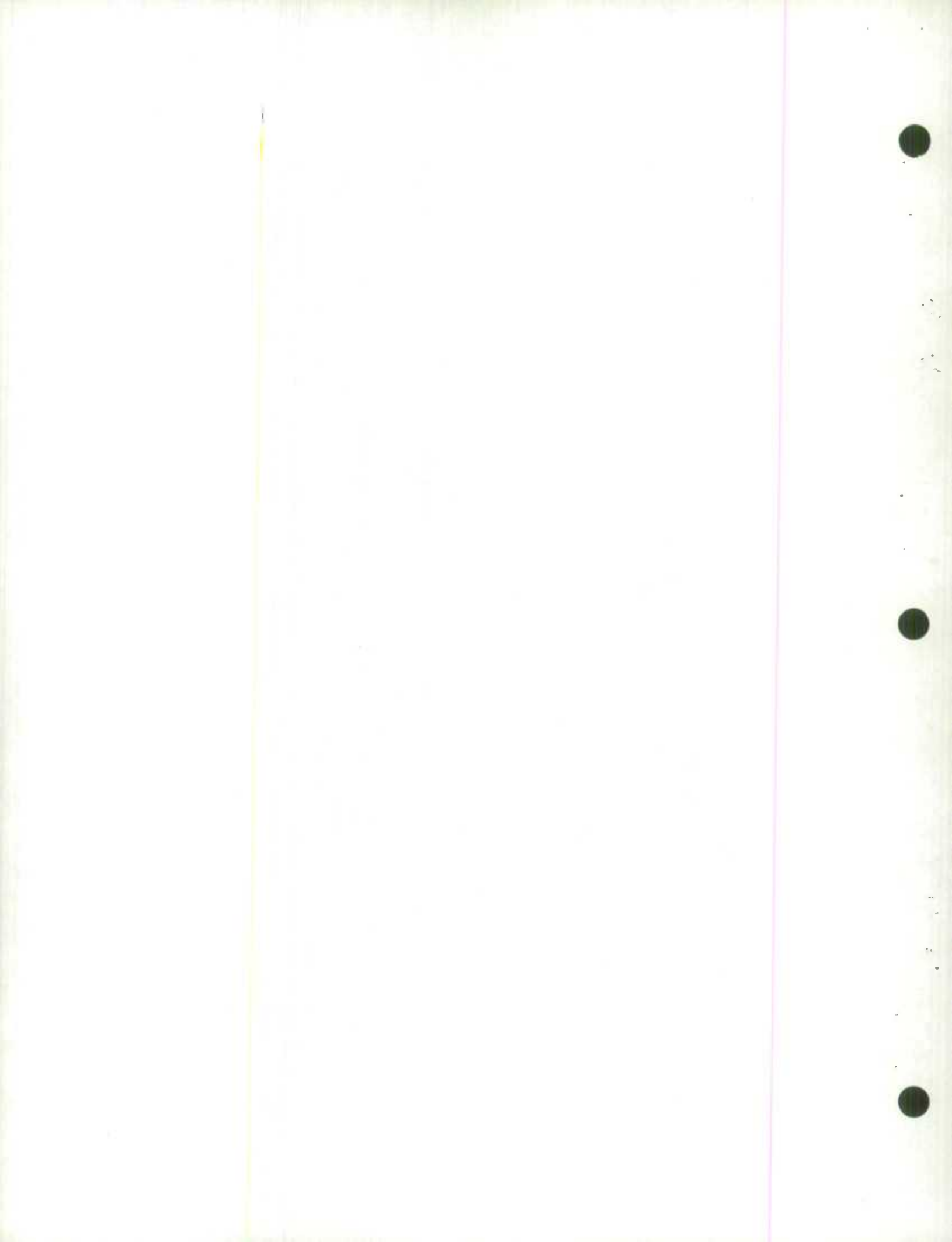
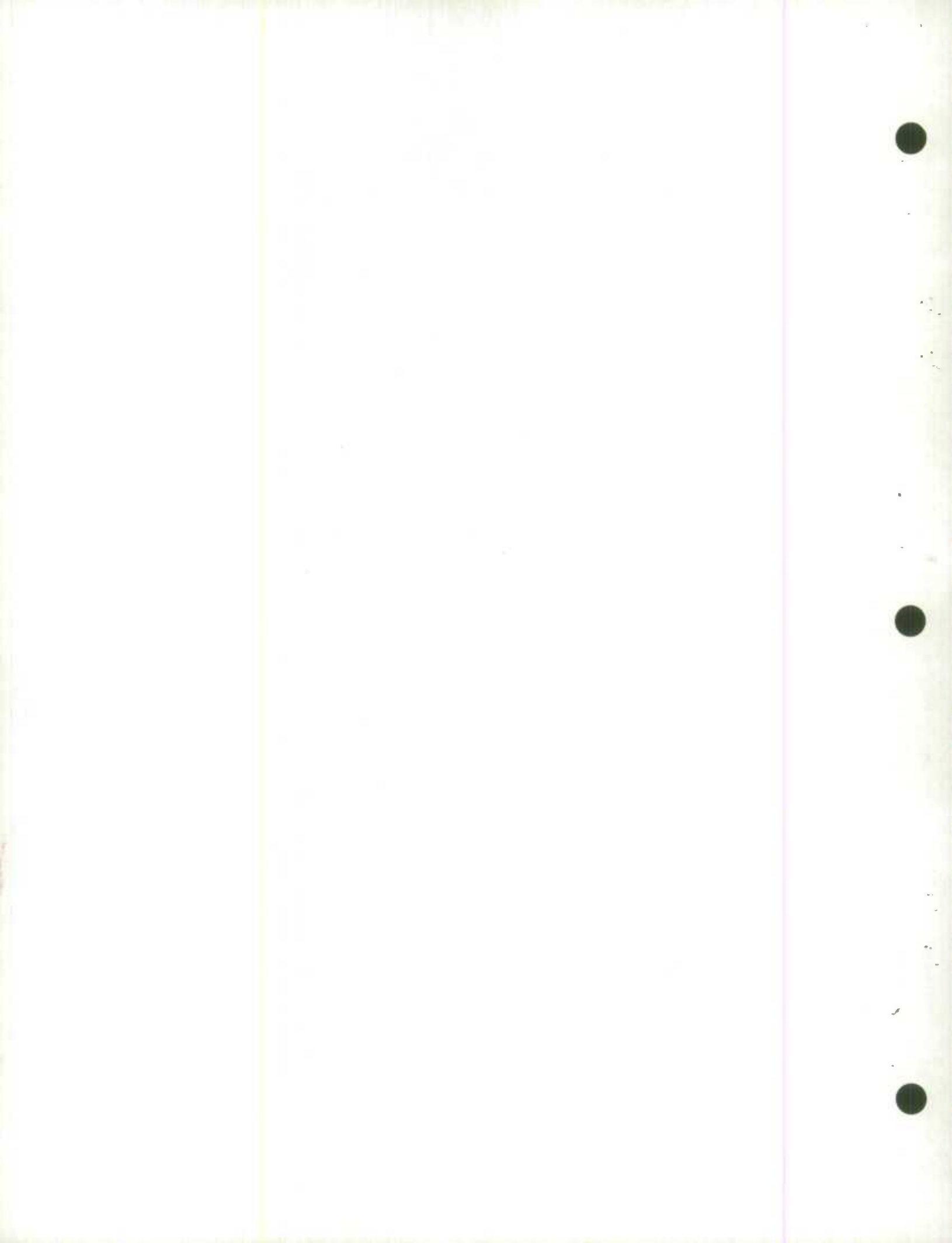


Table 4. Relationship Between Classification of Sample Units and Domains of Study
Showing the Fields or Characteristics Considered at Each Domain





The other exception concerns the classification of each unit as to whether or not data has been collected from it. For the purpose of the study, no domain is established from this classification. The estimation procedure, in effect, considers the data obtained (for the active sample units), as representative of the data of all active sample units in the stratum, including those units for which no data was collected. Rather than to present in detail the estimation procedure for each characteristic or field of interest, the basic concepts are illustrated through the following example.



5.2 Example

Consider the estimation of the total number of units in the Manufacturing stratum with changes of a legal, financial or corporate nature.

The sample results can be table as follows:

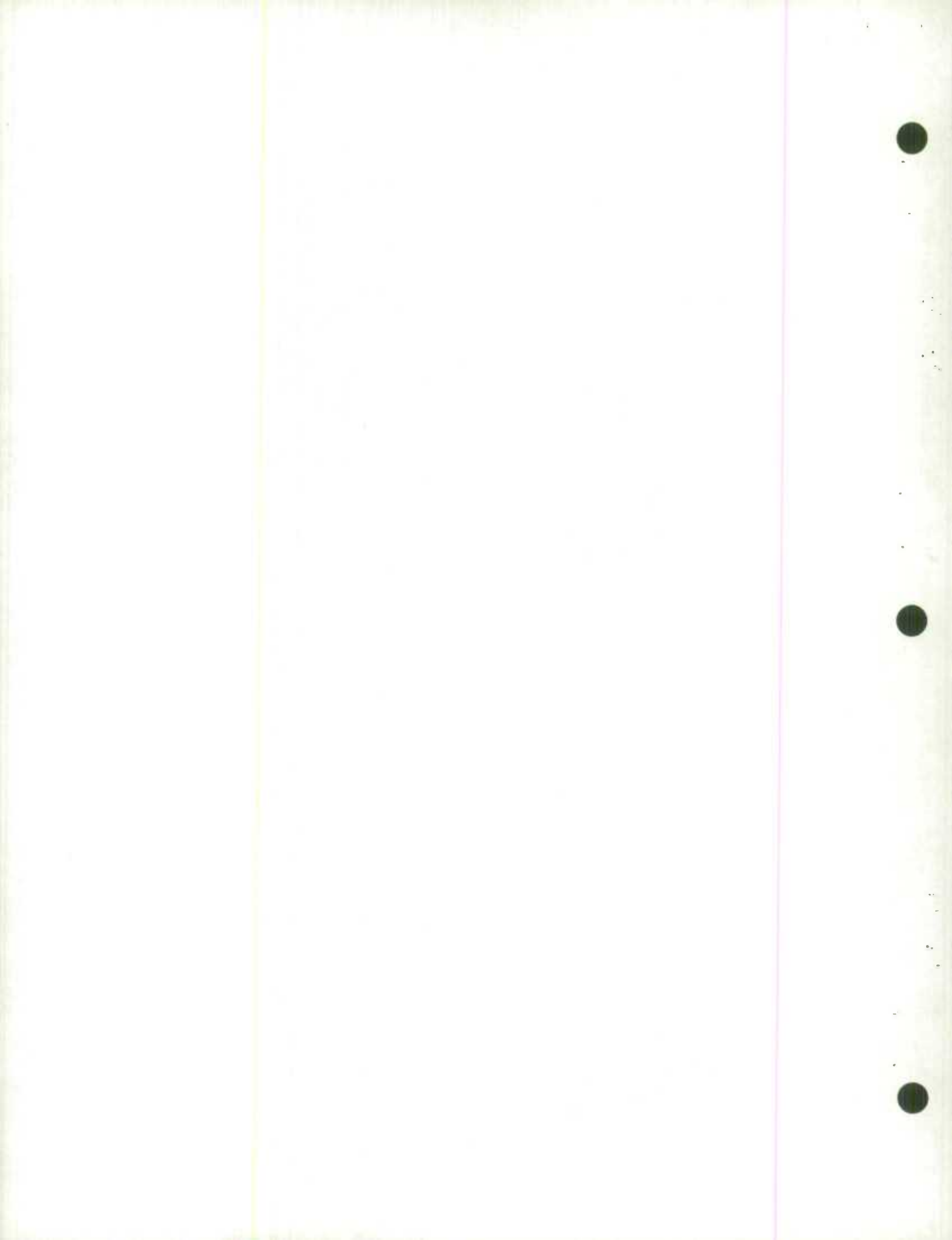
	(No. of units)		
	<u>Changes</u>	<u>No Changes</u>	<u>Total</u>
<u>Data Collected</u>	(X_{11}) 31	(X_{12}) 527	558
<u>Active</u>			
<u>Data Not Collected</u>	(X_{21}) unknown	(X_{22}) unknown	13
<hr/>			
<u>Inactive</u>			56
<hr/>			
<u>Total</u>		(n) 627	

where n is the sample size of the stratum.

The entries above are obtained from the tables of the sample results given in Appendix A.

The following notes are made:

1. The number of active units (and hence the number of inactive units) in the sample is known. By definition, a unit is inactive if it is out of business, permanently inactive or if Regional Office is unable to locate it.

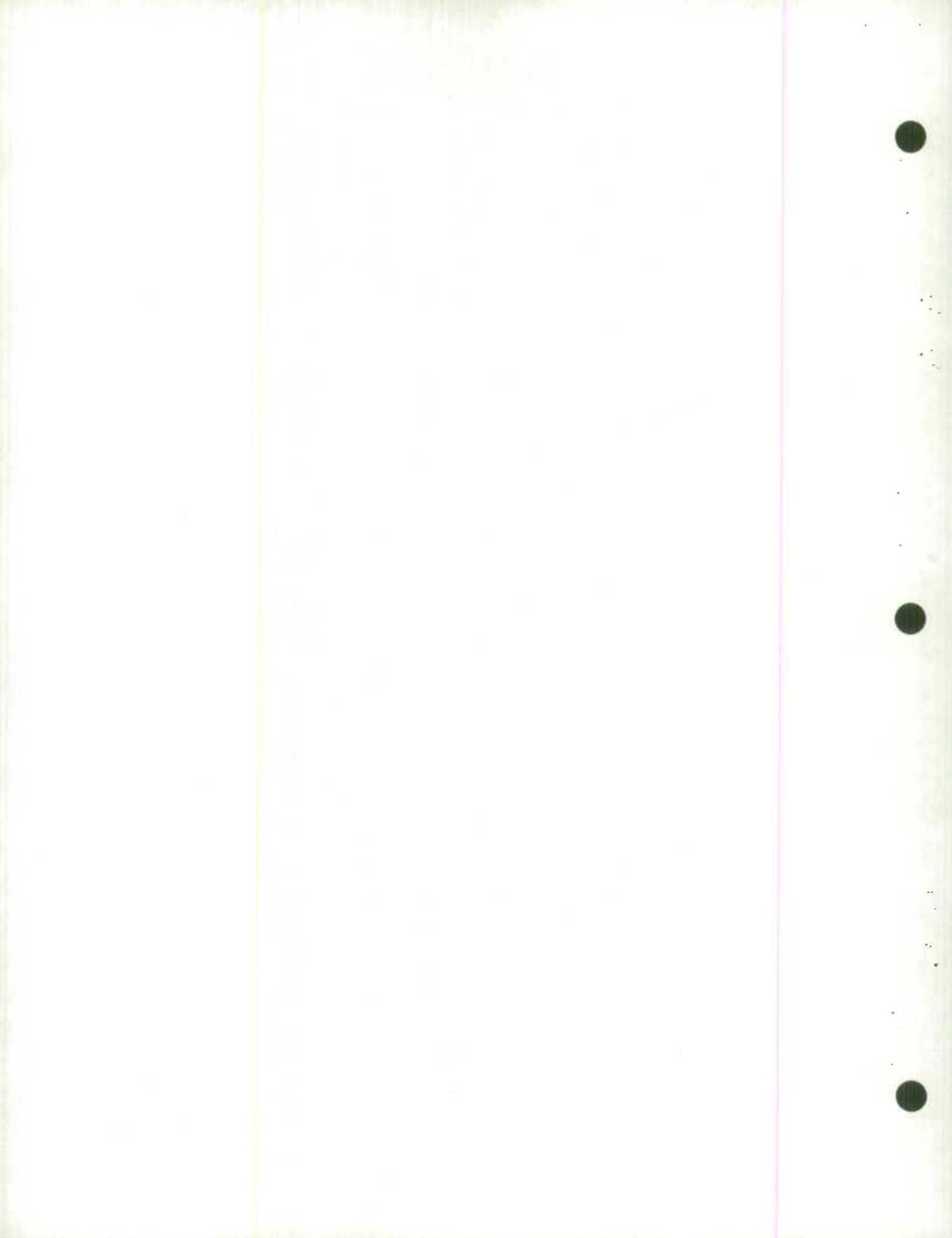


2. If data has been collected from the unit then it can be classified as having a change or no change in its legal, financial or corporate structure. Hence the values X_{11} and X_{12} are known from the sample.
3. The class given by "Data Not Collected" includes refusals, units in receivership, units requiring special contact, units which Regional Office is unable to contact and units which have not returned from Regional Office. These units are assumed to be active but no further information is known. We may suppose that there are some units in this group with changes in their legal, financial or corporate structures. This number, X_{21} , is unknown, although the total number of units for which data has not been collected, is known.

The population within the stratum can be thought as being partitioned in a manner similar to that given for the sample units:

(No. of units)			
	<u>Changes</u>	<u>No Changes</u>	<u>Total</u>
<u>Data Collected</u>	N_{11}	N_{12}	$N_{11} + N_{12}$
<u>Active</u>			
<u>Data Not Collected</u>	N_{21}	N_{22}	$N_{21} + N_{22}$
<u>Inactive</u>			$N - N_{11} - N_{12} - N_{21} - N_{22}$
<u>Total</u>			N

where N is the known stratum size



This classification of the population can be regarded as existing in principle had we conducted a census instead of a survey.

The parameter of interest is $N_{11} + N_{21}$.

If all active units had responded the usual estimate would be given by

$$\frac{N}{n} \times (x_{11} + x_{21})$$

However, x_{21} is not known so that the above estimate cannot be considered.

An alternative estimate is given by

$$\frac{1}{f} = \left(\frac{x_{11}}{x_{11} + x_{12}} \right) \times \left(\frac{x_{11} + x_{12} + x_{21} + x_{22}}{n} \right) \times N$$

where $x_{21} + x_{22}$ is known although x_{21} and x_{22} are not individually known.

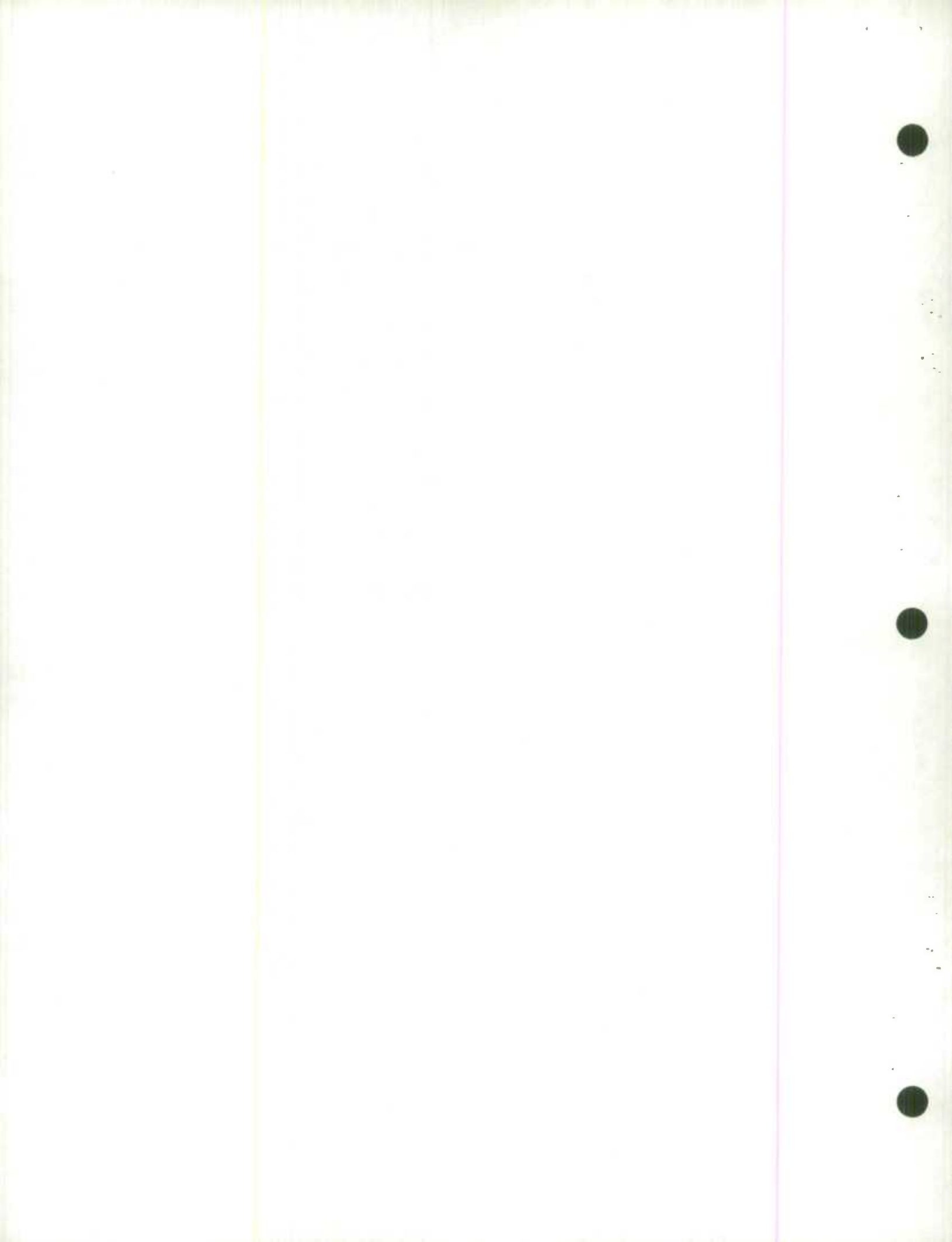
We note that

$$\left(\frac{x_{11} + x_{12} + x_{21} + x_{22}}{n} \right) \times N$$

is an estimate of the number of the proportion of active units in the stratum and,

$$\frac{x_{11}}{x_{11} + x_{12}}$$

is an estimate of the proportion of active units with changes (based on those units for which data was collected).



The estimate is appropriate if:

1. there is little or no interaction between the classification, ("Data Collected/Not Collected" and "Change/No Change").
- and 2. the observed value $\chi_{21} + \chi_{22}$ is small relative to the sample size N .

The implication of the above comments are as follows. If there is little or no interaction, it can be proven that the estimate is (approximately) unbiased for the parameter of interest. The second comment relates to the variance of the estimate. It can be shown, that the variance of the estimate is dominated by the terms involving χ_{11} and χ_{12} . Therefore even if there is a mild interaction present, the variance and the estimate of variance will not be distorted significantly if $\chi_{21} + \chi_{22}$ is relatively small.

The form of the estimate is non-linear in $(\chi_{11}, \chi_{12}, \chi_{21} + \chi_{22})$ so that approximate methods are used to compute its variance.

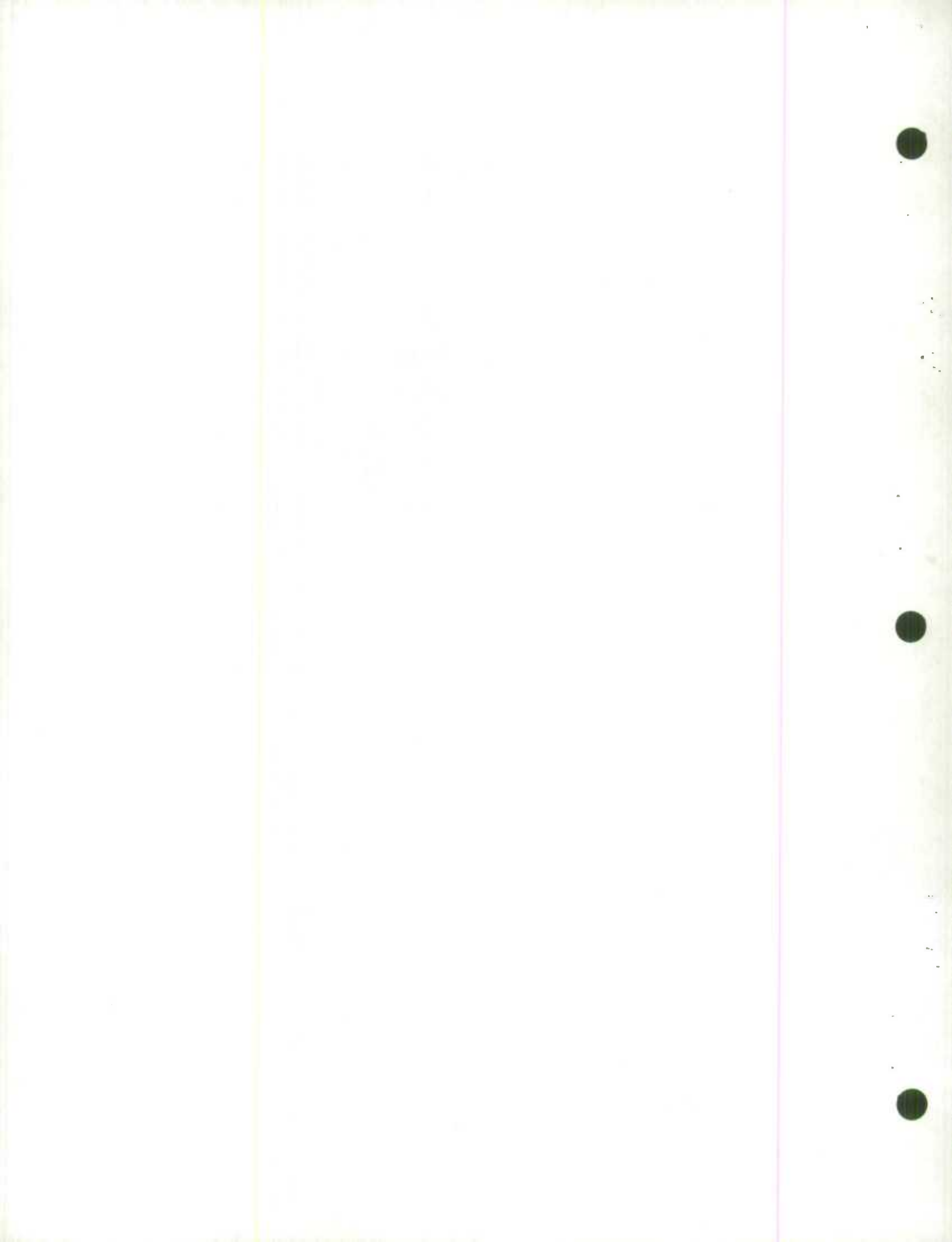
An estimate of the variance is given by the expression in Appendix C. No further details are given in this report.

For the given example, the estimated number of active units in the Manufacturing stratum, with changes of a legal, financial or corporate nature, is

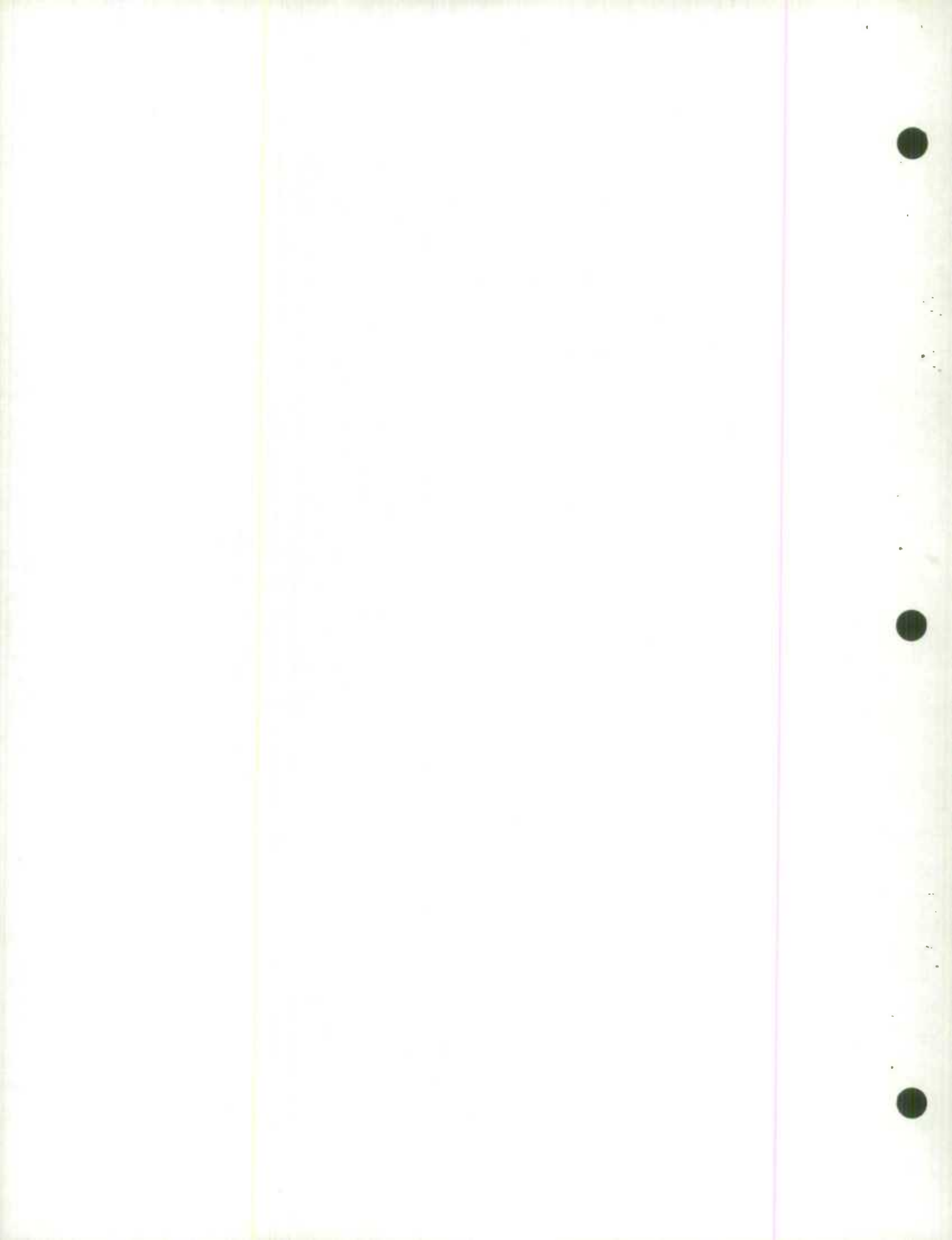
$$\left(\frac{31}{31+527}\right) \times \left(\frac{31+527+13}{627}\right) \times 46270 \approx 2341$$

with estimated variance 165806 obtained by the formula in Appendix C. The corresponding estimates of standard error and coefficient of variation are

407.2 and .1739 respectively.



A similar technique and assumptions are used to estimate other characteristics. The estimates for the various characteristics or fields are given in Appendix B. The estimates are presented by stratum (industry group) and also at the universe level, which is the aggregation of the strata. However, each estimate may only relate to a specific domain within each stratum. It is important to note the particular domain under consideration for each of the estimates produced. The domain is the population of reference for which the estimate is produced and hence, proper inferences can only be made within the domain. Thus for example, the estimates produced for the quality of the BRMF 1970 SIC code, refer only to that domain consisting of active units with no major legal changes and not to the domain of active units. A definition or description of each of the various characteristics is also given in Appendix B.

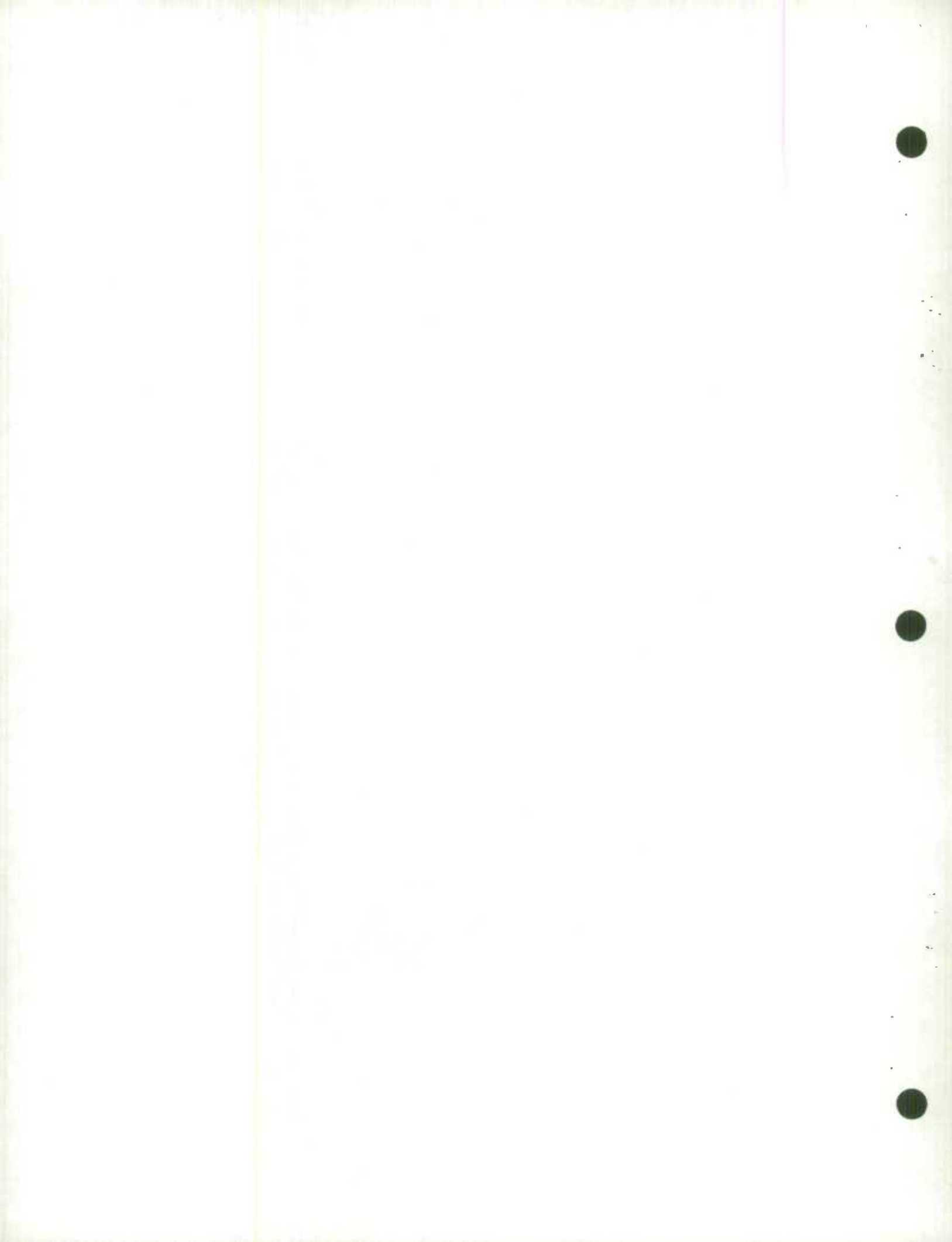


6. Further Notes

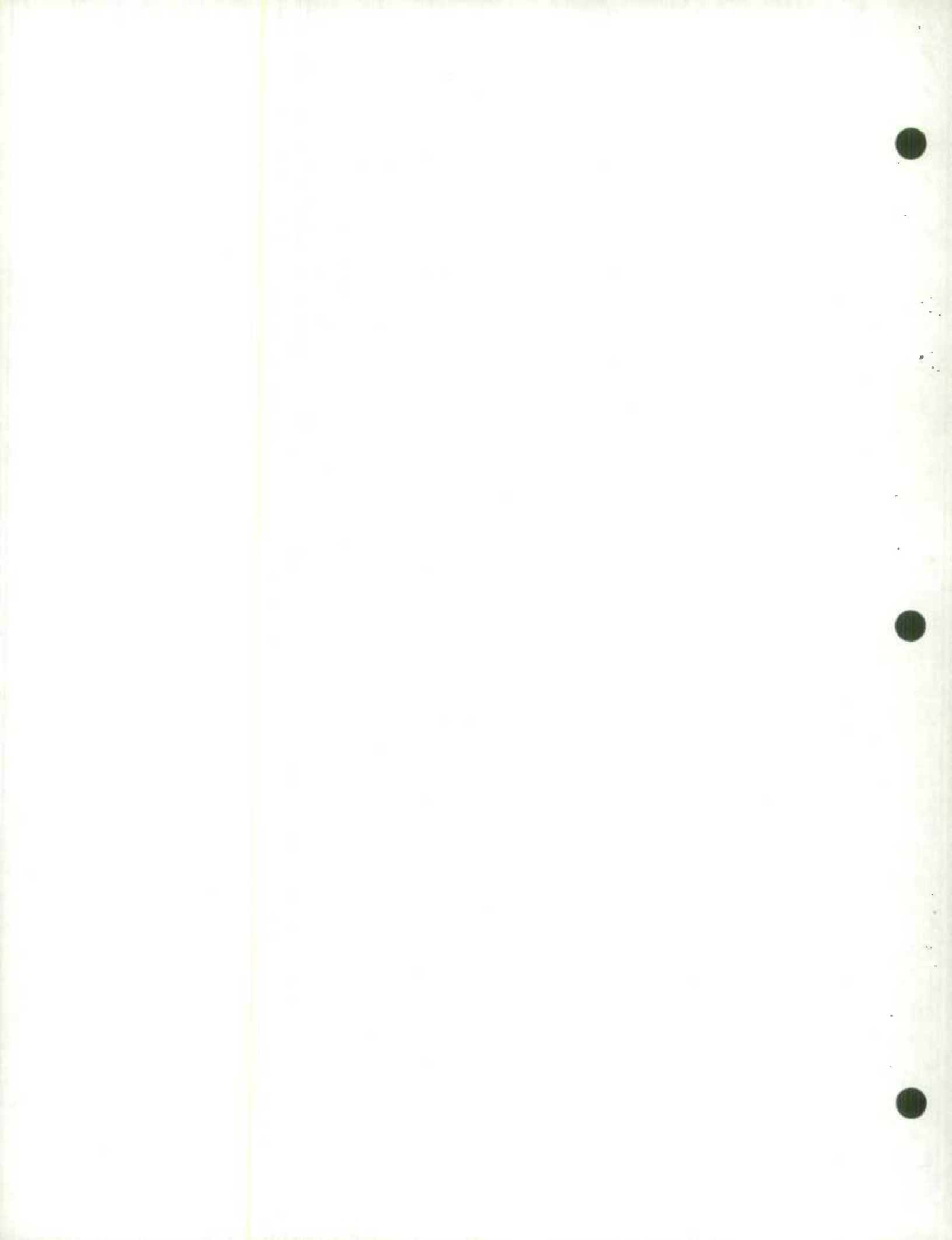
In this section, a few comments are made regarding an earlier assumption. Also, a brief note is made about future work.

An important assumption that is made throughout this study is that the 611 units captured by the Company Contact System, have up-to-date information on the BRMF. This assumption was required when it was decided that no mail-out would be carried out for these units. While this assumption is not totally valid, the rationale for such course of action was as follows. First, there was some concern about response burden on these units, which have already been contacted within the previous 12 months. This reason was further reinforced by a second consideration. It was felt that for most of these 611 units, the information captured within the previous 12 months, would still remain fairly indicative of their status at the time of selection. This was built on the premise that the information would have been used to update the BRMF and that, in the interim period, few changes would have occurred to these units. These assertions are reasonably valid, although it is known for instance, that there is some backlog in the processing of the units. Nevertheless, it was further noted that, with the exception of Forestry and Manufacturing, these units formed a small percentage of each stratum sample (2%-11%). Therefore, even if some changes had occurred in the characteristics of some units, the corresponding estimates of change would not be significantly distorted. For the purposes of the study, this was considered to be satisfactory.

As a follow-up to this study, future work will involve an investigation of the feasibility of incorporating the data from the SEPH survey into an on-going program of quality assurance. Currently, data is



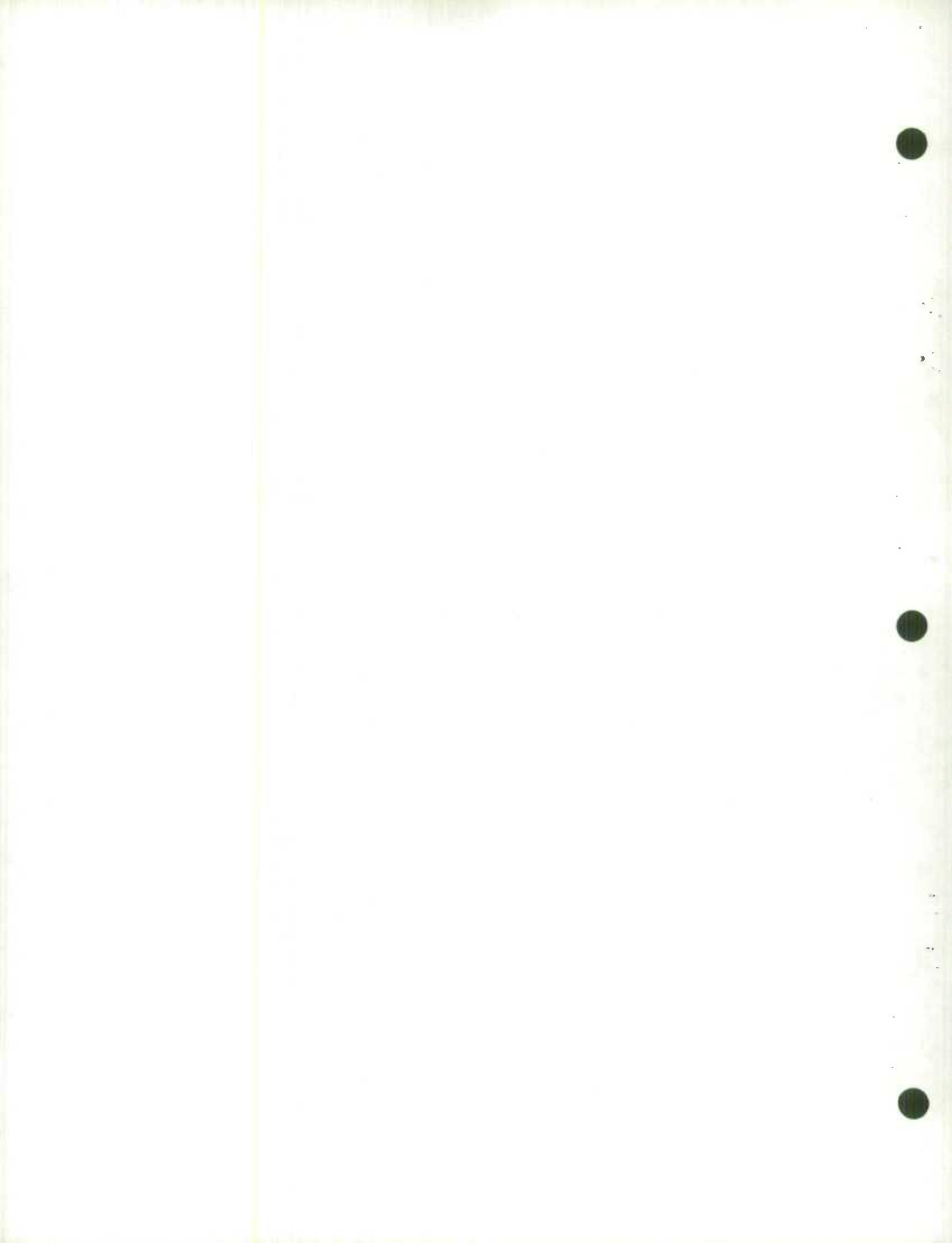
available from the SEPH survey on a monthly basis, for a fairly large number of sample units. However, the SEPH survey is primarily concerned with employment related characteristics. As such, it does not cover certain SIC groups such as Agriculture. Also, it does not systematically capture information on certain fields of interest such as changes in label name. A supplementary sample could be drawn to accommodate these considerations and the combined sample could be used to produce the necessary estimates.



Appendix A

Distribution of Sample Units

By Selected Characteristics

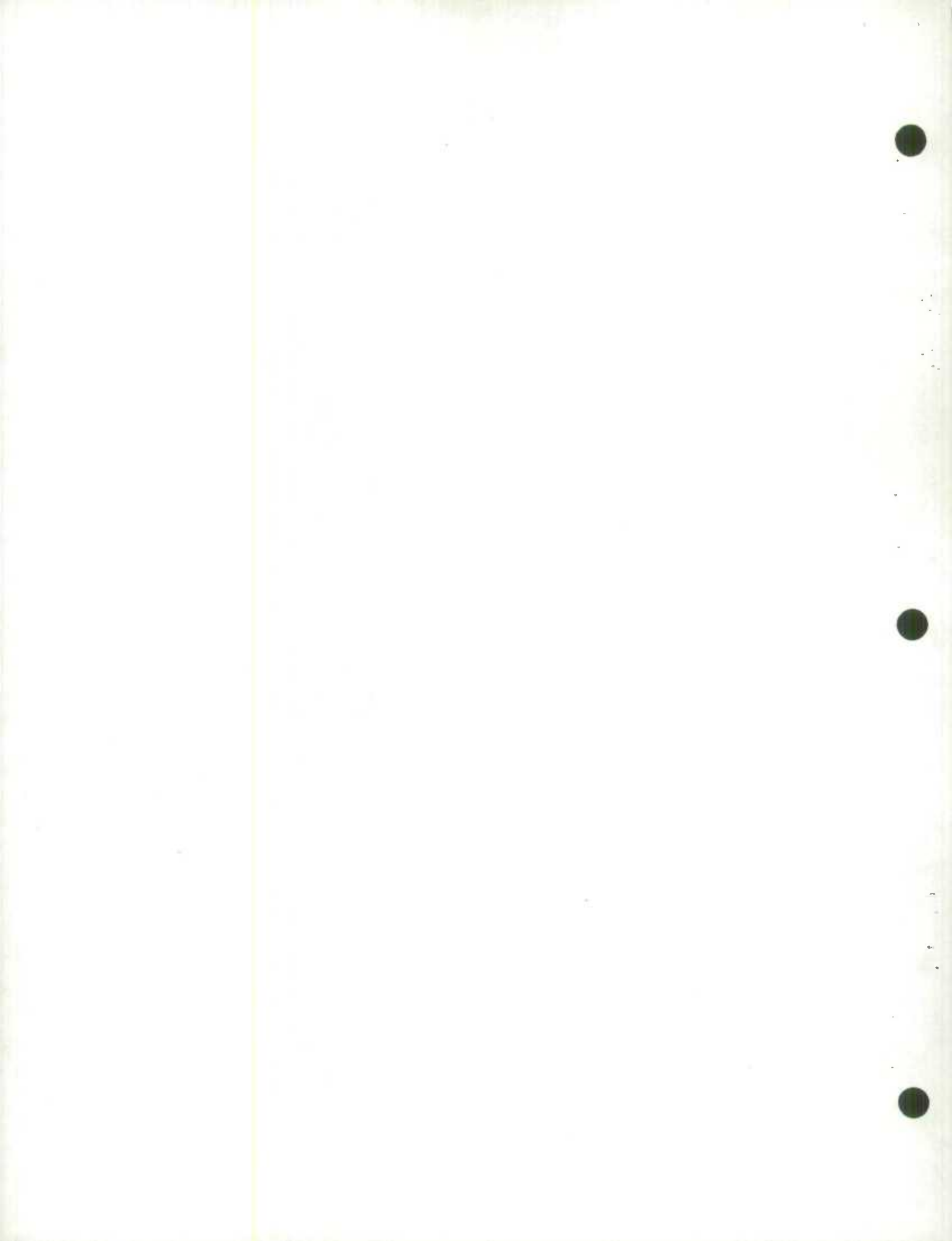


Classification by Sample Characteristics

(First 2 levels of classification)

ActiveInactive

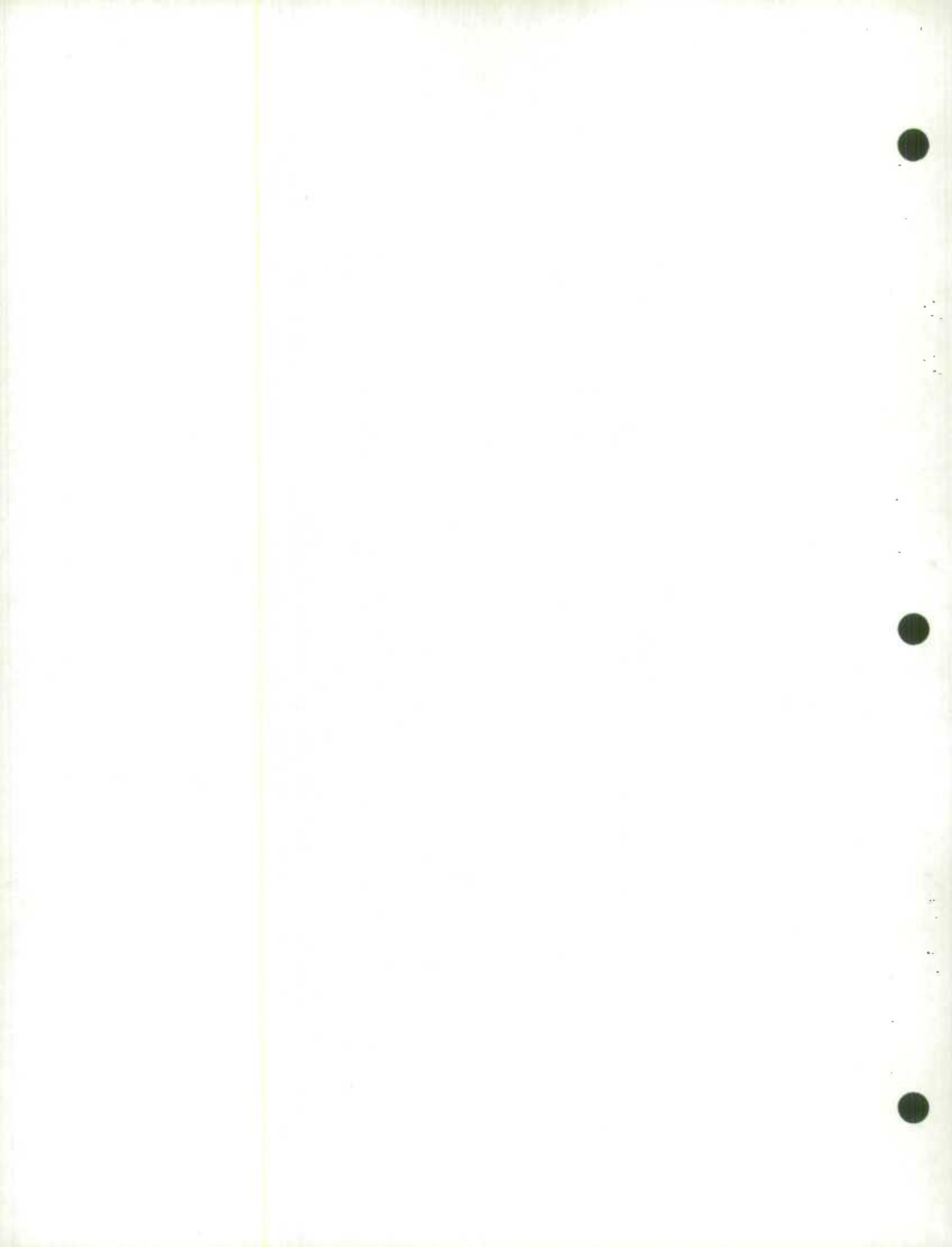
INDUSTRY GROUP	<u>Data Collected</u>		<u>Data Not Collected</u>					TOTAL			
	CAPTURED BY CCS	OTHER SAMPLE UNITS	UNABLE TO CONTACT	REFUSAL	IN RECEI VERSHIP	SPECIAL CONTACT	NOT RETURNED FROM RO		UNABLE TO LOCATE	OUT OF BUSINESS	PERM INACTIVE
AGRICULTURE	25 3.14	668 84.03	26 3.27	4 0.50	2 0.25	1 0.13	0 0.00	29 3.65	33 4.15	7 0.88	795
FORESTRY	77 30.20	113 44.31	12 4.71	1 0.39	0 0.00	1 0.39	0 0.00	17 6.67	31 12.16	3 1.18	255
FISHING AND TRAPPING	5 2.42	153 73.91	11 5.31	1 0.48	0 0.00	0 0.00	0 0.00	11 5.31	23 11.11	3 1.45	207
MINING	9 4.37	149 72.33	7 3.40	2 0.97	1 0.49	2 0.97	0 0.00	17 8.25	13 6.31	6 2.91	206
MANUFACTURING	106 16.91	452 72.09	5 0.80	5 0.80	2 0.32	1 0.16	0 0.00	21 3.35	34 5.42	1 0.16	627
CONSTRUCTION	92 9.54	615 63.80	27 2.80	1 0.10	5 0.52	2 0.21	1 0.10	85 8.82	100 10.37	36 3.73	964
TRANSPORTATION	42 7.89	389 73.12	11 2.07	6 1.13	2 0.38	2 0.38	0 0.00	34 6.39	39 7.33	7 1.32	532
COMMUNICATIONS AND OTHER UTILITIES	17 11.04	115 74.68	0 0.00	0 0.00	1 0.65	0 0.00	0 0.00	7 4.55	13 8.44	1 0.65	154
WHOLESALE TRADE	38 5.70	493 73.91	16 2.40	3 0.45	0 0.00	2 0.30	1 0.15	47 7.05	52 7.80	15 2.25	667
RETAIL TRADE	55 4.74	820 70.69	23 1.98	6 0.52	7 0.60	3 0.26	0 0.00	56 4.83	169 14.57	21 1.81	1160
FINANCE INSURANCE AND REAL ESTATE	62 8.58	524 72.48	10 1.38	5 0.69	0 0.00	3 0.41	0 0.00	54 7.47	45 6.22	20 2.77	723
EDUCATION AND HEALTH SERVICES	33 5.53	494 82.75	11 1.84	7 1.17	1 0.17	0 0.00	0 0.00	9 1.51	32 5.36	10 1.68	597
OTHER SERVICES	50 3.81	972 74.03	36 2.74	11 0.84	3 0.23	0 0.00	0 0.00	100 7.62	118 8.99	23 1.75	1313
TOTAL	611	5957	195	52	24	17	2	487	702	155	8200



Classification by Sample Characteristics

(For active units for which data was collected)

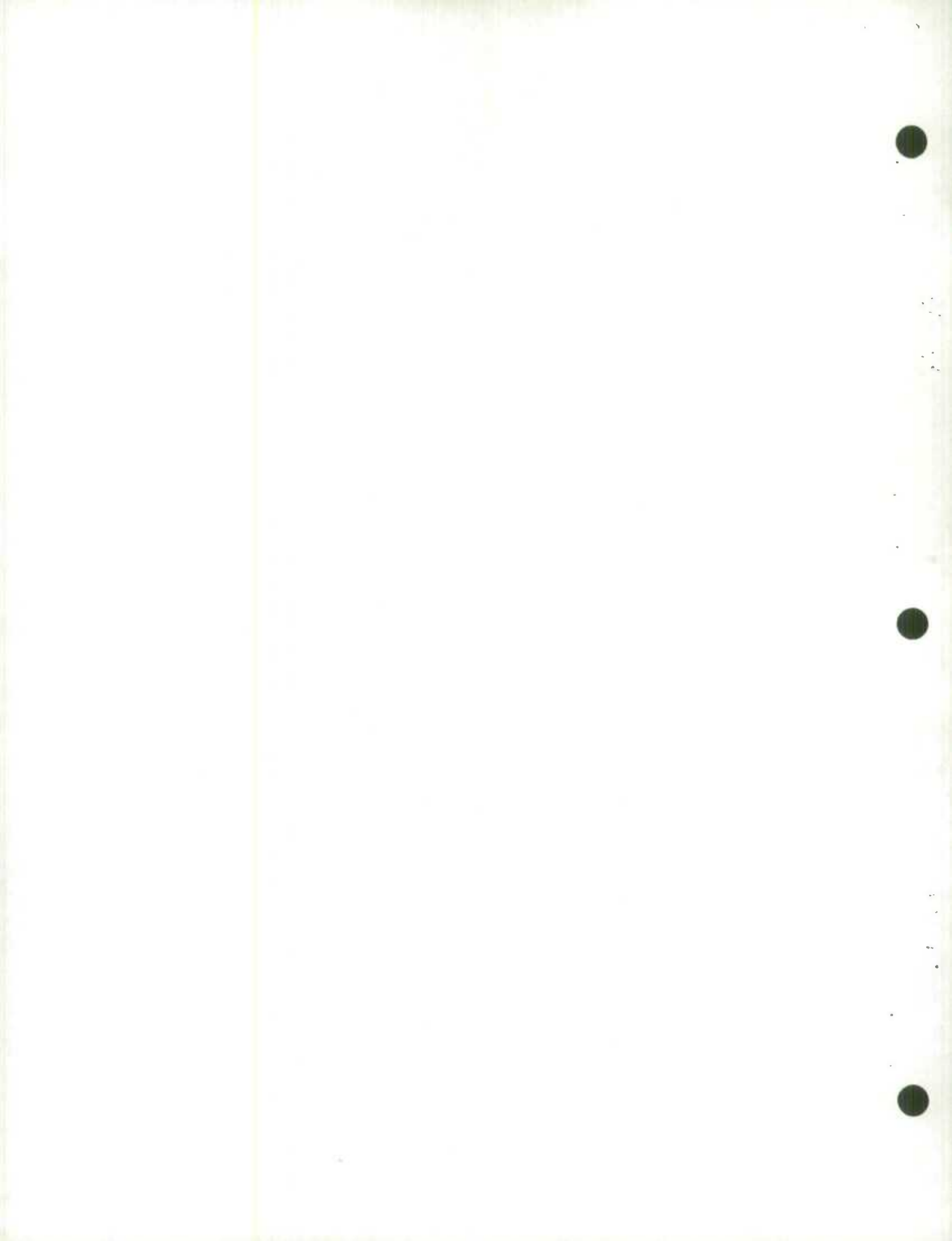
INDUSTRY GROUP	<u>No Major Change</u>			<u>Major Change</u>			TOTAL
	NO LEGAL CHANGES	NEW INCORPORATIONS	CHARTER AMENDMENTS	CHANGES OF OWNER SHIP	AMALGAMATIONS	OTHER	
AGRICULTURE	652	16	2	13	0	0	693
	95.53	2.31	0.29	1.88	0.00	0.00	
FORESTRY	185	4	1	0	0	0	190
	97.37	2.11	0.53	0.00	0.00	0.00	
FISHING AND TRAPPING	154	3	1	0	0	0	158
	97.47	1.90	0.63	0.00	0.00	0.00	
MINING	147	2	2	3	3	1	158
	93.04	1.27	1.27	1.90	1.90	0.63	
MANUFACTURING	527	8	10	10	1	2	558
	94.44	1.43	1.79	1.79	0.18	0.36	
CONSTRUCTION	678	17	6	5	0	1	707
	95.90	2.40	0.85	0.71	0.00	0.14	
TRANSPORTATION	405	13	4	8	0	1	431
	93.97	3.02	0.93	1.86	0.00	0.23	
COMMUNICATIONS AND OTHER UTILITIES	124	4	2	1	1	0	132
	93.94	3.03	1.52	0.76	0.76	0.00	
WHOLESALE TRADE	487	13	16	13	0	2	531
	91.71	2.45	3.01	2.45	0.00	0.38	
RETAIL TRADE	788	30	9	47	0	1	875
	90.06	3.43	1.03	5.37	0.00	0.11	
FINANCE INSURANCE AND REAL ESTATE	558	4	12	5	4	3	586
	95.22	0.68	2.05	0.85	0.68	0.51	
EDUCATION AND HEALTH SERVICES	515	6	2	1	0	3	527
	97.72	1.14	0.38	0.19	0.00	0.57	
OTHER SERVICES	938	19	16	45	1	3	1022
	91.78	1.86	1.57	4.40	0.10	0.29	
TOTAL	6168	139	83	151	10	17	6568



Classification by Sample Characteristics
(For active units with no major legal changes)

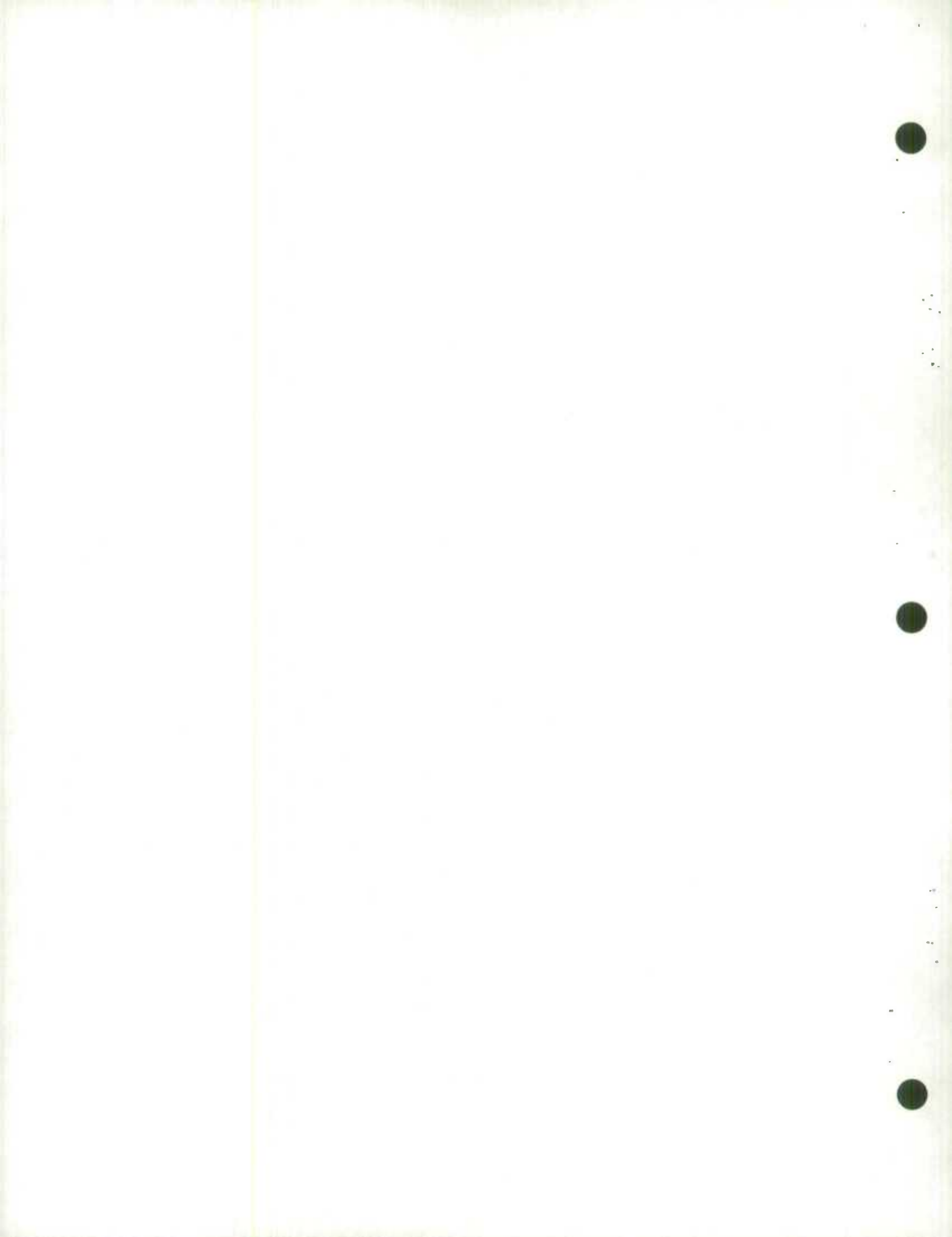
BRMF Label Address

INDUSTRY GROUP	LOCATION ONLY	LOCATION AND MAILING	MAILING ONLY	NEITHER LOCATION NOR MAILING	CAPTURED BY CCS	TOTAL
AGRICULTURE	12 1.76	576 84.71	38 5.59	29 4.26	25 3.68	680
FORESTRY	5 2.63	93 48.95	9 4.74	6 3.16	77 40.53	190
FISHING AND TRAPPING	5 3.16	118 74.68	24 15.19	6 3.80	5 3.16	158
MINING	8 5.30	112 74.17	9 5.96	13 8.61	9 5.96	151
MANUFACTURING	22 4.04	354 64.95	39 7.16	24 4.40	106 19.45	545
CONSTRUCTION	18 2.57	483 68.90	58 8.27	50 7.13	92 13.12	701
TRANSPORTATION	19 4.50	299 70.85	34 8.06	28 6.64	42 9.95	422
COMMUNICATIONS AND OTHER UTILITIES	3 2.31	91 70.00	14 10.77	5 3.85	17 13.08	130
WHOLESALE TRADE	23 4.46	381 73.84	33 6.40	41 7.95	38 7.36	516
RETAIL TRADE	31 3.75	618 74.73	74 8.95	49 5.93	55 6.65	827
FINANCE INSURANCE AND REAL ESTATE	53 9.23	365 63.59	38 6.62	56 9.76	62 10.80	574
EDUCATION AND HEALTH SERVICES	11 2.10	406 77.63	43 8.22	30 5.74	33 6.31	523
OTHER SERVICES	57 5.86	709 72.87	88 9.04	69 7.09	50 5.14	973
TOTAL	267	4605	501	406	611	6390



Appendix B

Tables of Estimates for
Various Fields and Characteristics



Description of Fields and Corresponding Estimates

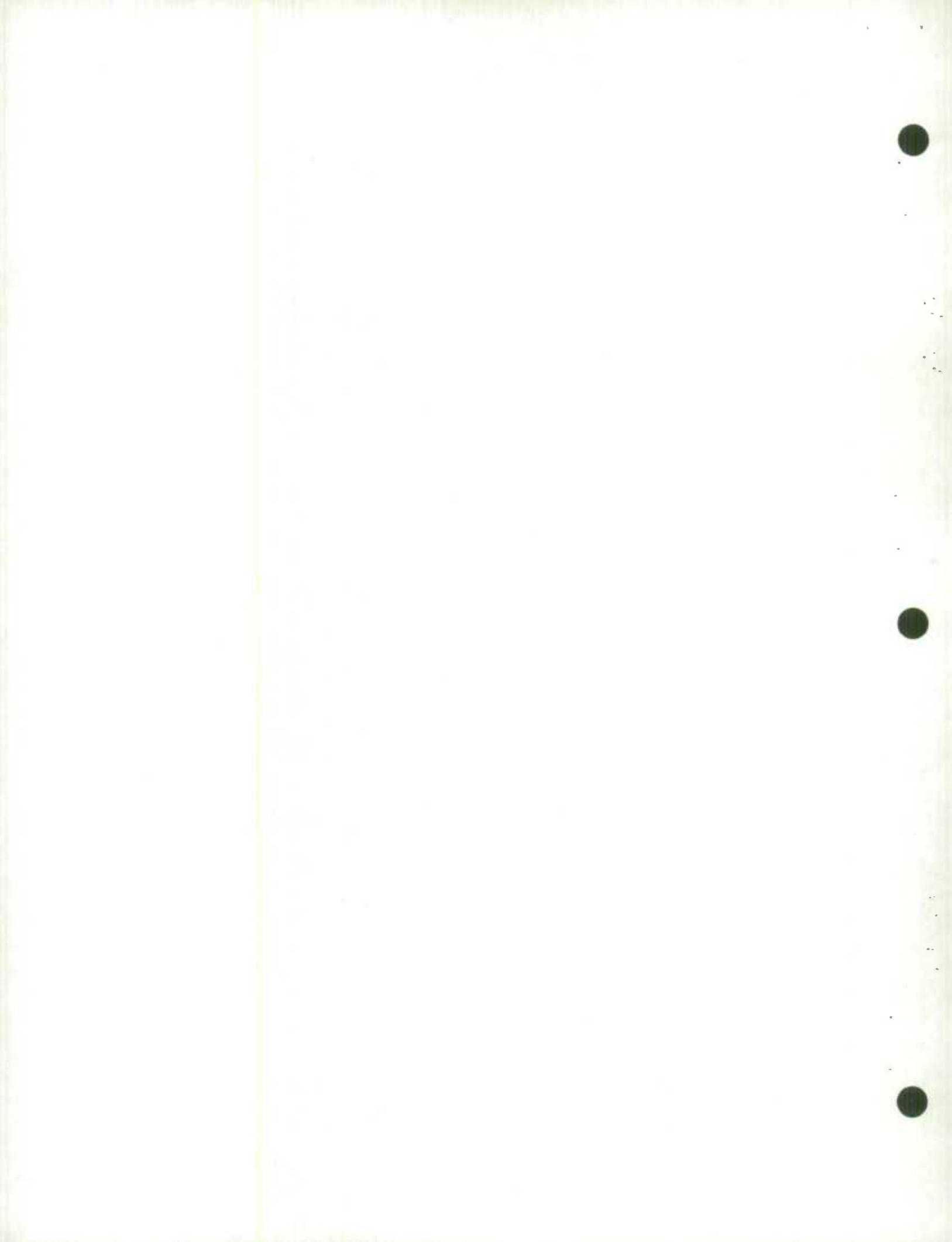
A. Classification and Identification Fields

1970 SIC Code:

- for those sample units classified as active with no major legal changes, a 3 digit 1970 SIC code was assigned based on all available information (CBS-718 data on industry activity, research and knowledge of the BRMF SIC code).
- this assigned code was then compared with that on the BRMF to establish estimates of change in 1970 SIC and industry group classification.
- the comparison of the SIC codes leads to the classification of the units into 4 mutually exclusive groups:
 - (i) no change in SIC
 - (ii) change at the first level (first digit of the code)
 - (iii) change at the second level (second digit)
 - (iv) change at the third level (third digit)
- the domain (active units with no major legal changes) estimates for the above groups are given in Table B.4.
- tables B.5.1, B.5.2 and B.6 present domain estimates of the number of units with a change in 1980 SIC code and industry group classification.

Label Name:

- the label name refers to the name of the unit as given on the CBS-718 form and the BRMF
- changes to the label name were obtained as part of the Regional Office enquiry
- changes to the label name were only considered for units with no legal, financial or corporate changes.
- the estimates of change at the domain level (given above) are presented in table B.9



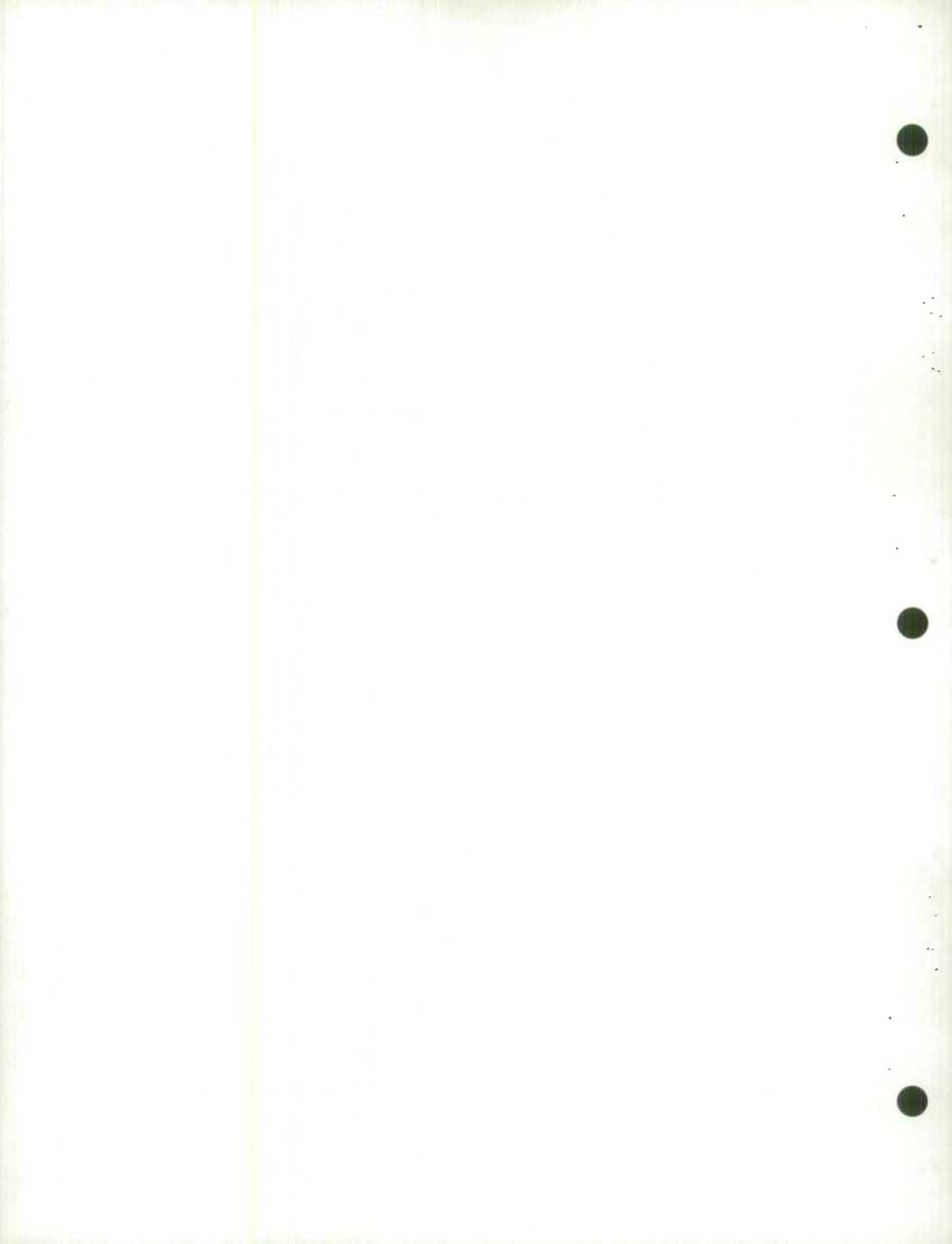
B. Size Fields

Employment Size Code

- for those sample units, active with no major legal changes, an employment size code was assigned based on the reported maximum number of employees within 1 year.
- if this number was not obtained then the code was assigned based on the current number of employees.
- if neither the expected nor the current number of employees was given then this field was set to code 99 (unknown)
- the coding was done using the standard size code values:

<u>Max. Expected Number of Employees</u>	<u>Size Code</u>
0	00
1-4	01
4-9	02
10-19	03
⋮	⋮
5000 and over	12
1-19 (but unknown)	20
Unknown	99

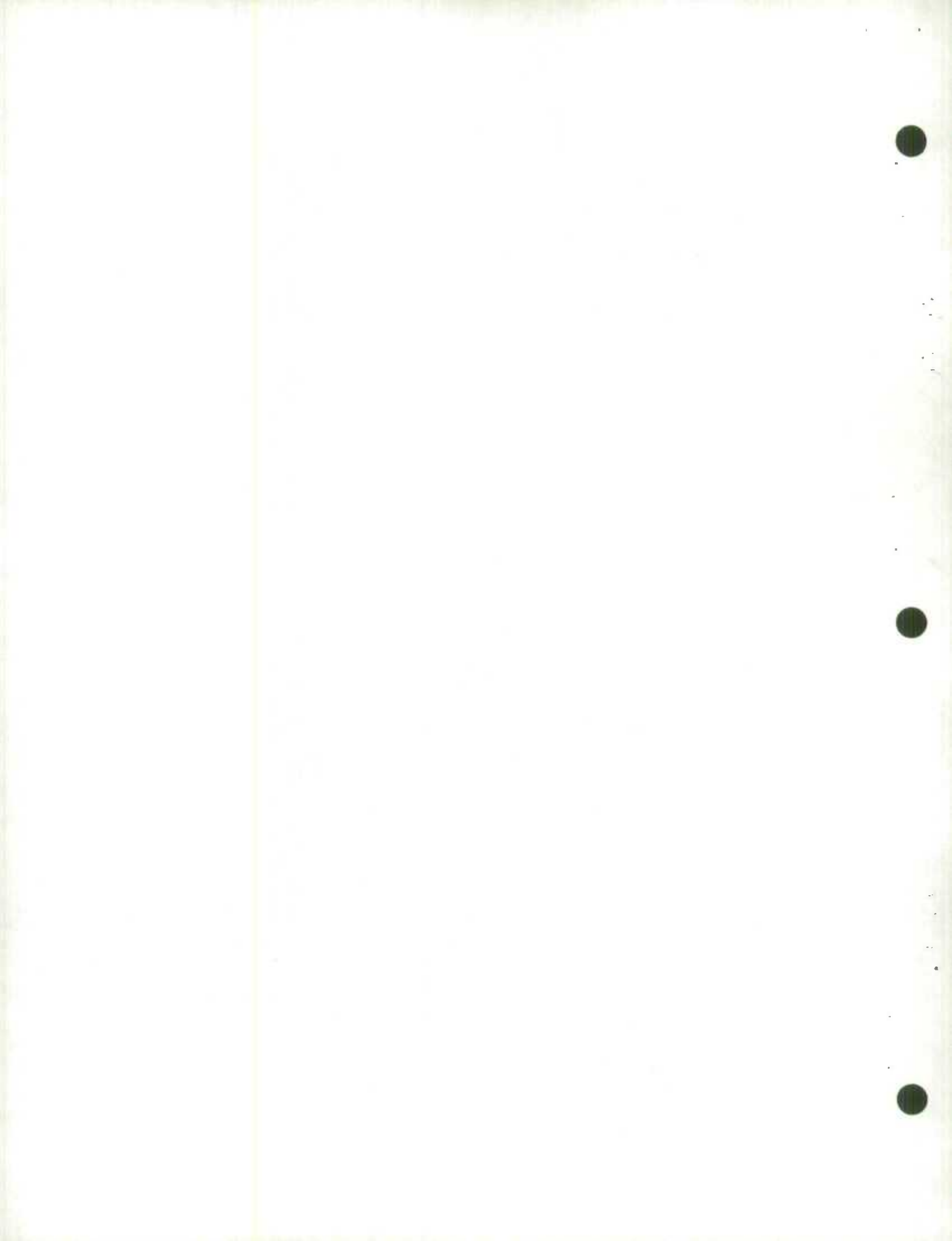
- the estimates of change given in table B.7 and B.8 are based only on those sample units with a current code other than 99 (except, if both the BRMF and the current size codes are 99)
- the reason for the above restriction is that a current size code of 99 provides no information as to whether or not the BRMF size code has changed (since the current code is unknown); however, if both codes had a value of 99 then this was considered to be consistent with no change in the size code.
- the estimates given in tables B.7 and B.8 were determined as follows: a change in the size code of a unit was considered to have occurred if the BRMF size code was different from that coded from the sample data; except if, the BRMF size code was 20 and the current code was either 00, 01, 02 or 03; again, this was considered to be consistent with no change in the size code



- for the estimates in table B.8, a change is noted to be a difference in size codes which results in a different group classification; the groups are based on the SEPH employment size classes:

<u>Group</u>	<u>Employment Size Codes</u>
1	00, 01, 02, 03 and 20
2	04
3	05 and 06
4	07, 08, 09, 10, 11, 12
(5) *	99 (on BRMF)

* (for convenience the above groups are referred as SEPH size groups, although the last group is not considered as such for the SEPH survey).



C. Address Fields

- for each sample unit, active with no major legal changes, the label address was classified into one of the following categories:
 - (i) location only
 - (ii) location and mailing
 - (iii) mailing only
 - (iv) neither location nor mailing
- the label address is the address for the unit as given on the CBS-718 label and the BRMF
- table B.3 presents the estimated number of units for which the label address is neither location of activity nor a valid mailing address.

Postal Code

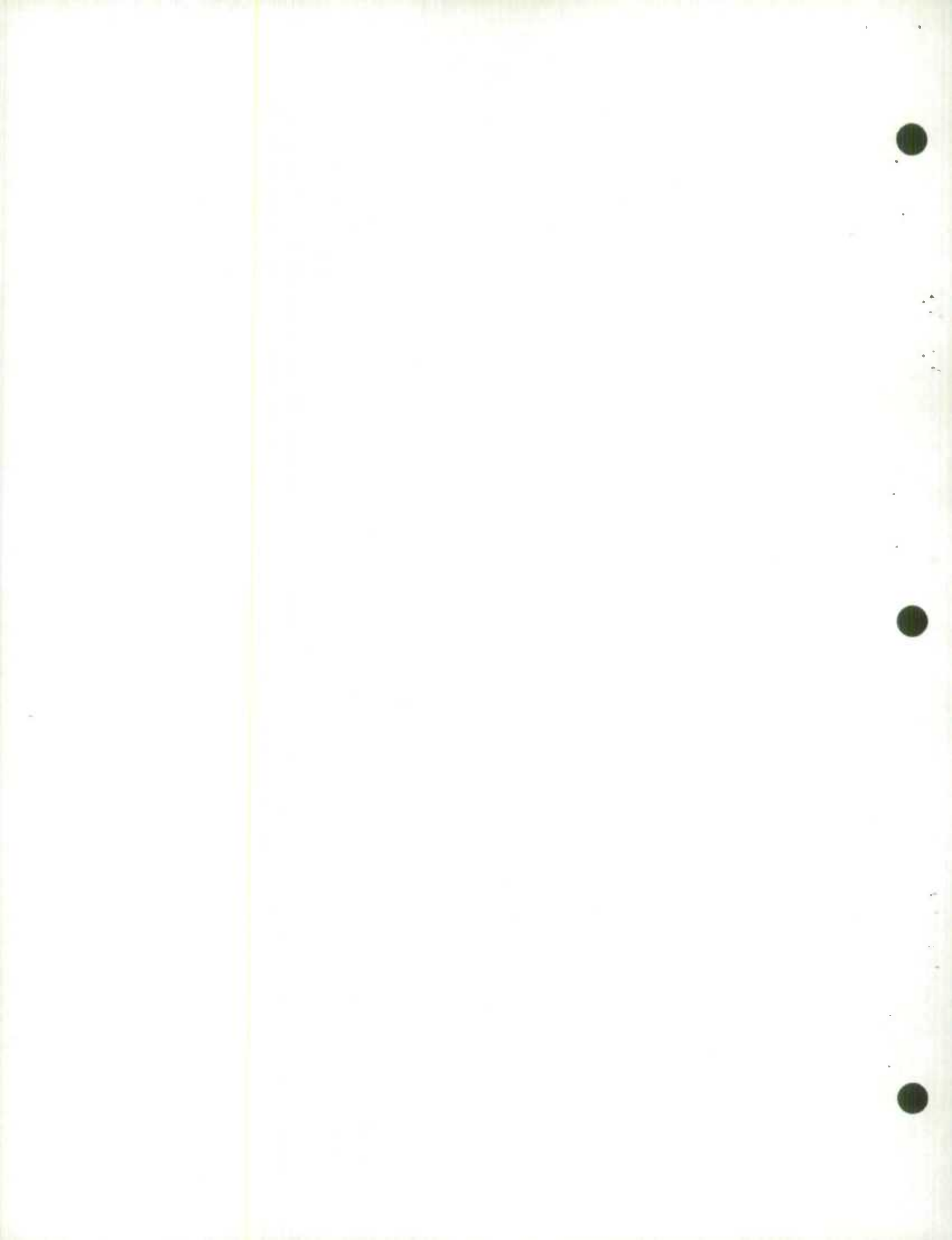
- the postal code field was only coded for those sample units with a BRMF label address representing the location of activity (it may be both the location and mailing address)
- for these units, estimates of the change in postal code were produced from a comparison of the BRMF and the reported postal code; these estimates are presented in table B.10.

Province

- the province code was also only captured for those sample units with a BRMF label address representing the location of activity.
- comparing the BRMF and the reported province codes it was found that there were only 5 sample units with a change in province code (1 in Construction, 1 in Transportation and 3 in Fin., Ins., and R. Estate)
- the estimates are given in table B.13.

Standard Geographical Classification (SGC) Code

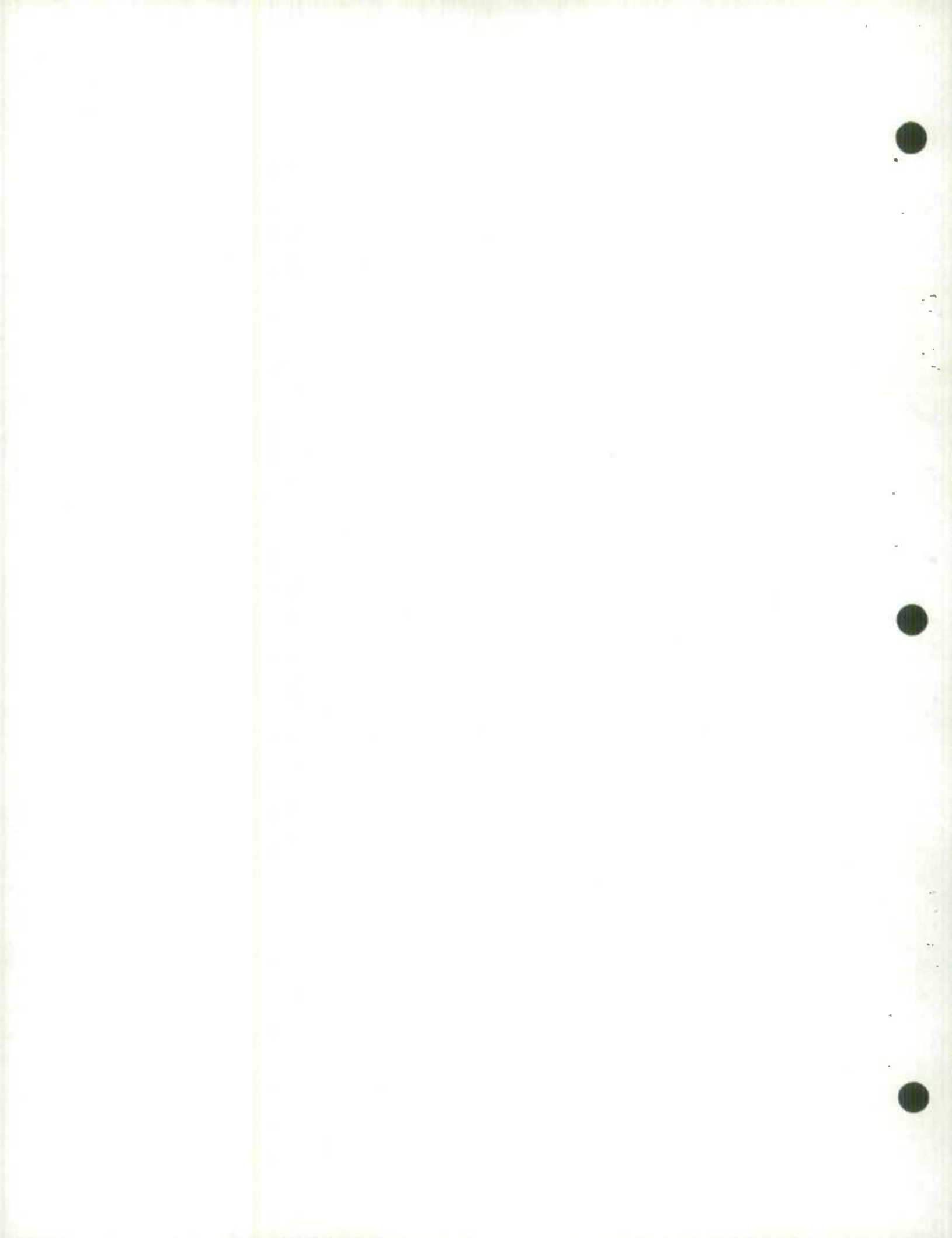
- The SGC code for the sample units was generated, by a routine, from the (location) postal code



- upon comparison with the BRMF SGC the sample units were classified into 4 groups:

- (i) no change in SGC
- (ii) change at the first level (first 2 digits of the code)
- (iii) change at the second level (next 2 digits)
- (iv) change at the third level (last 3 digits)

- the estimates at the domain level are given in tables B.11 and B.12



Index to Tables

(Presented by Domain)

Active Units (based on BRMF code)

- | | | |
|----|--|-----------|
| 1. | Estimation of the Total Number of Active and Inactive Units (By Industry Group) | Table B.1 |
| 2. | Estimation of the Total Number of Units with Changes of a Legal, Financial or Corporate Nature | Table B.2 |

Active Units With No Major Legal Changes

- | | | |
|----|---|----------------------------|
| 1. | Estimation of the Total Number of Units for which the Label Address is Neither Location Nor Mailing | Table B.3 |
| 2. | Estimation of the Total Number of Units with a Change in 1970 SIC Code (3 Tables) | Table B.4, B.5.1 and B.5.2 |
| 3. | Estimation of the Total Number of Units with a Change in Industry Group | Table B.6 |
| 4. | Estimation of the Total Number of Units with a Change in Employment Size Code | Table B.7 |
| 5. | Estimation of the Total Number of Units Changing SEPH Employment Size Group | Table B.8 |

Active Units With No Legal Changes

- | | | |
|----|--|-----------|
| 1. | Estimation of the Total Number of Units Requiring a Non-Legal Update to the Label Name | Table B.9 |
|----|--|-----------|

Active Units With No Major Legal Changes and for Which the Label Address is at Least Location

- | | | |
|----|---|----------------------|
| 1. | Estimation of the Total Number of Units With a Change in Location Postal Code | Table B.10 |
| 2. | Estimation of the Total Number of Units With a Change in SGC (2 Tables) | Tables B.11 and B.12 |
| 3. | Estimation of the Total Number of Units With a Change in Province Code | Table B.13 |

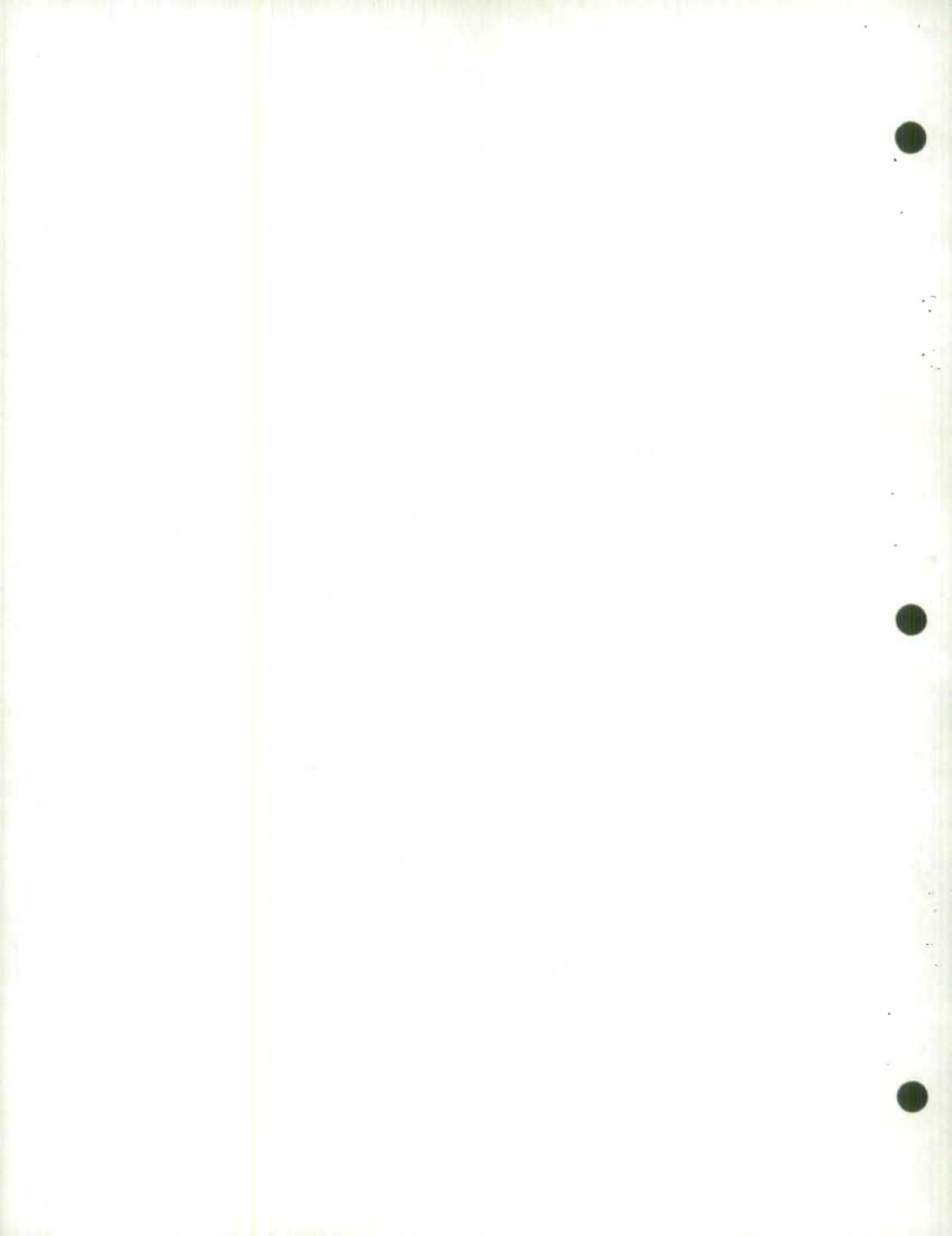


Table B.1

ESTIMATION OF THE TOTAL NUMBER OF ACTIVE AND INACTIVE UNITS
BY INDUSTRY GROUP

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		INDUSTRY TOTAL	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF INACTIVE UNITS	
	ACTIVE	INACTIVE		STANDARD ERROR	CV
AGRICULTURE	67793 91.32	6443 8.68	74236	737.7	0.1145
FORESTRY	6106 79.99	1527 20.01	7633	188.3	0.1234
FISHING AND TRAPPING	4147 82.14	902 17.86	5049	132.0	0.1463
MINING	4112 82.52	871 17.48	4983	129.4	0.1486
MANUFACTURING	42137 91.07	4133 8.93	46270	523.8	0.1268
CONSTRUCTION	84265 77.07	25064 22.93	109329	1474.4	0.0588
TRANSPORTATION	28305 84.96	5010 15.04	33315	512.6	0.1023
COMMUNICATIONS AND OTHER UTILITIES	2405 86.36	380 13.64	2785	75.1	0.1977
WHOLESALE TRADE	43338 82.91	8934 17.09	52272	757.6	0.0848
RETAIL TRADE	124649 78.79	33549 21.21	158198	1892.5	0.0564
FINANCE INSURANCE AND REAL ESTATE	51269 83.54	10101 16.46	61370	841.9	0.0833
EDUCATION AND HEALTH SERVICES	38312 91.46	3579 8.54	41891	476.2	0.1331
OTHER SERVICES	165526 81.65	37212 18.35	202738	2159.7	0.0580
TOTAL	662364 82.79	137705 17.21	800069	3617.5	0.0263

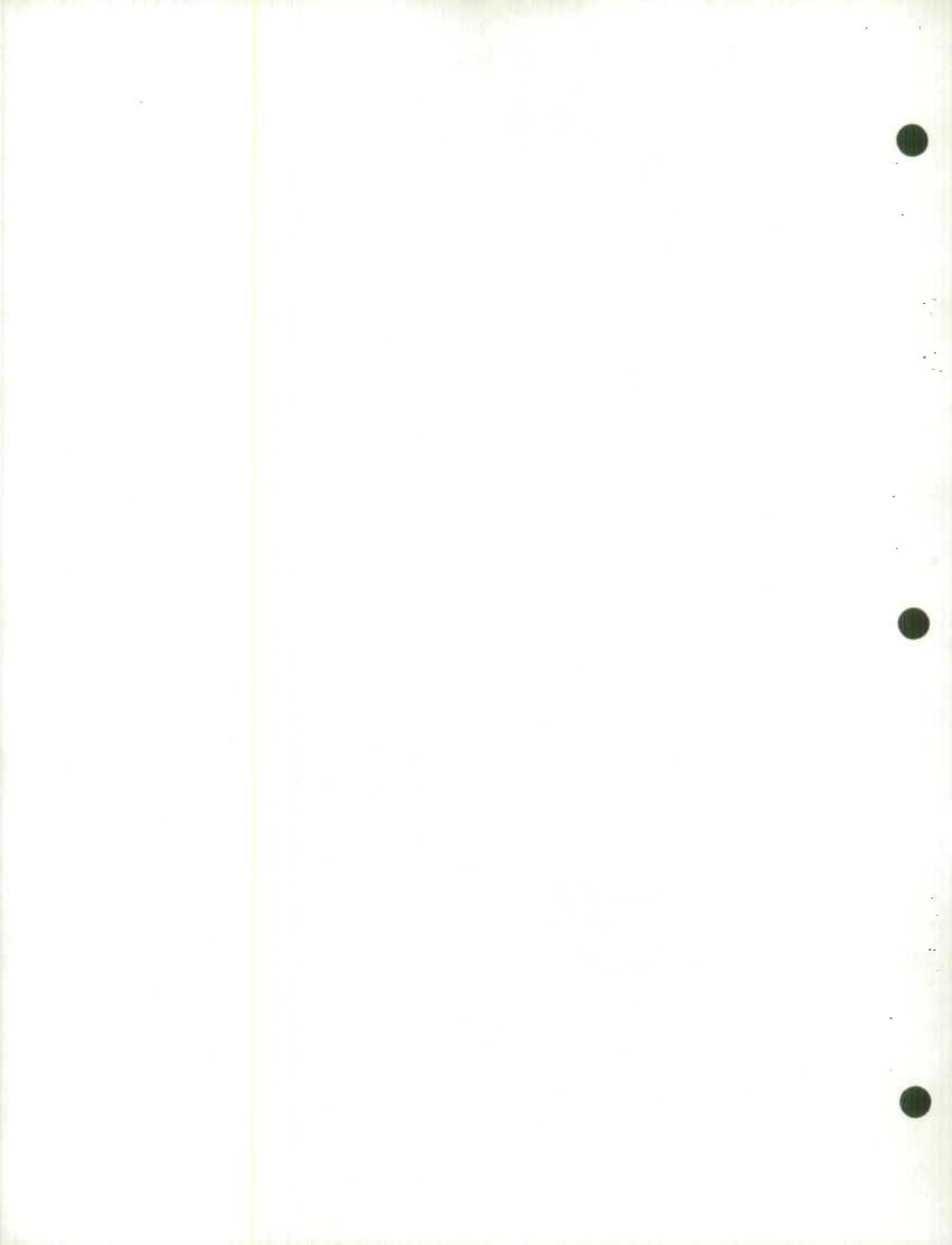


Table B.2

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH CHANGES
OF A LEGAL, FINANCIAL OR CORPORATE NATURE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH CORPORATE, FINANCIAL OR LEGAL CHANGES	
	WITHOUT CHANGES	WITH CHANGES			STANDARD ERROR	CV
AGRICULTURE	64760 95.53	3033 4.47	67793	91.32	530.8	0.1750
FORESTRY	5945 97.37	161 2.63	6106	79.99	70.0	0.4358
FISHING AND TRAPPING	4042 97.47	105 2.53	4147	82.14	51.0	0.4856
MINING	3826 93.04	286 6.96	4112	82.52	82.2	0.2872
MANUFACTURING	39796 94.44	2341 5.56	42137	91.07	407.2	0.1739
CONSTRUCTION	80809 95.90	3456 4.10	84265	77.07	629.0	0.1820
TRANSPORTATION	26598 93.97	1707 6.03	28305	84.96	323.8	0.1896
COMMUNICATIONS AND OTHER UTILITIES	2259 93.94	146 6.06	2405	86.36	48.9	0.3356
WHOLESALE TRADE	39747 91.71	3591 8.29	43338	82.91	519.3	0.1446
RETAIL TRADE	112255 90.06	12394 9.94	124649	78.79	1270.9	0.1025
FINANCE INSURANCE AND REAL ESTATE	48819 95.22	2450 4.78	51269	83.54	451.2	0.1842
EDUCATION AND HEALTH SERVICES	37440 97.72	872 2.28	38312	91.46	247.6	0.2838
OTHER SERVICES	151921 91.78	13605 8.22	165526	81.65	1429.1	0.1050
TOTAL	618217 93.33	44147 6.67	662364	82.79	2270.8	0.0514

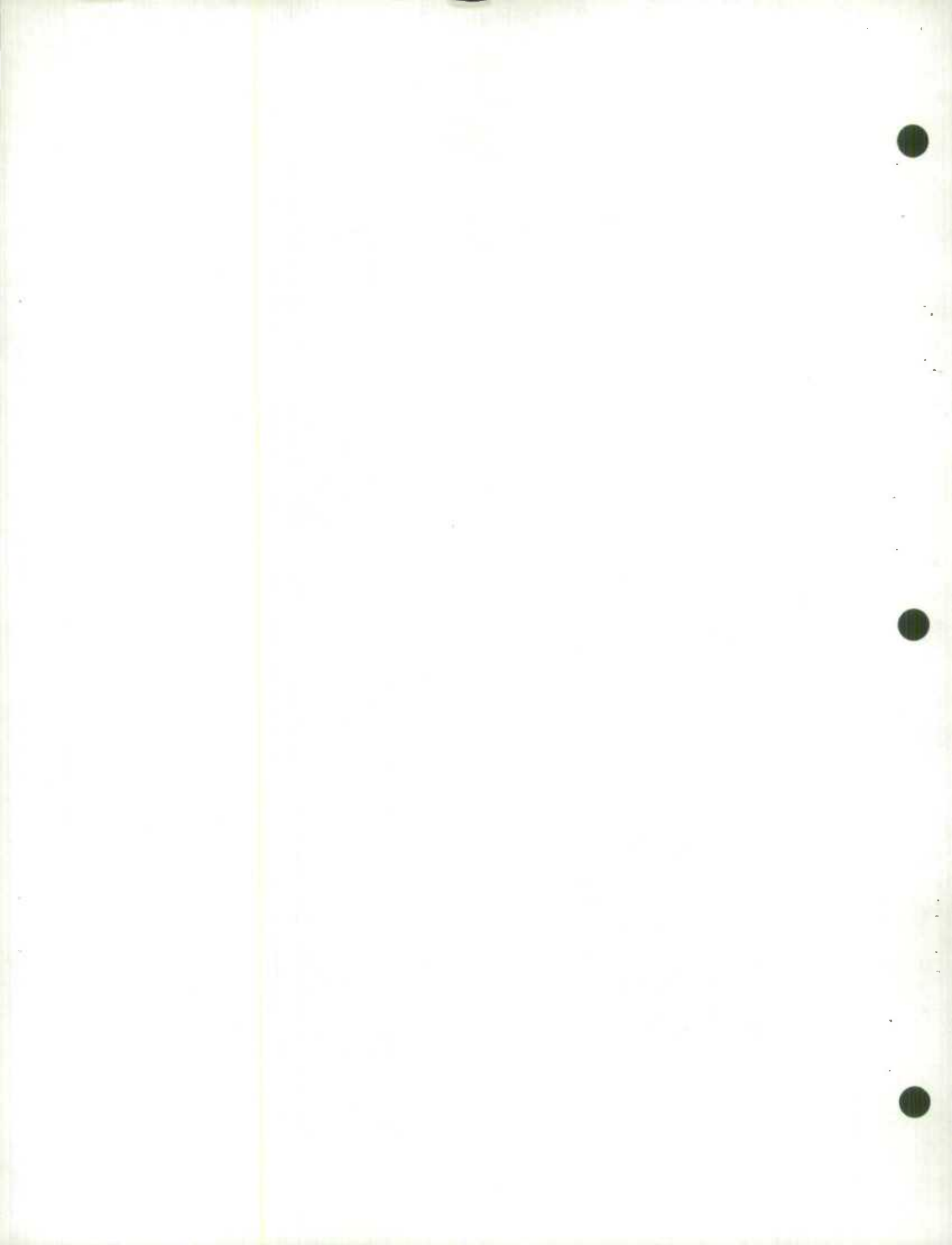


Table B.3

ESTIMATION OF THE TOTAL NUMBER OF UNITS FOR WHICH THE
LABEL ADDRESS IS NEITHER LOCATION NOR MAILING

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES
(NEW INCORPORATIONS, CORPORATIONS WITH CHARTER AMENDMENTS
OR CORPORATIONS WITH NO CORPORATE, FINANCIAL OR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS FOR WHICH THE LABEL ADDRESS IS NEITHER LOCATION NOR MAILING	
	LABEL ADDRESS IS EITHER LOCATION OR MAILING	LABEL ADDRESS IS NEITHER LOCATION NOR MAILING			STANDARD ERROR	CV
AGRICULTURE	63684 95.74	2837 4.26	66521	89.61	514.2	0.1812
FORESTRY	5913 96.84	193 3.16	6106	79.99	76.6	0.3970
FISHING AND TRAPPING	3990 96.21	157 3.79	4147	82.14	62.1	0.3944
MINING	3592 91.40	338 8.60	3930	78.87	88.9	0.2627
MANUFACTURING	39343 95.60	1812 4.40	41155	88.95	360.4	0.1989
CONSTRUCTION	77591 92.87	5959 7.13	83550	76.42	815.9	0.1369
TRANSPORTATION	25875 93.36	1839 6.64	27714	83.19	335.3	0.1823
COMMUNICATIONS AND OTHER UTILITIES	2278 96.16	91 3.84	2369	85.06	39.1	0.4289
WHOLESALE TRADE	38768 92.05	3346 7.95	42114	80.57	502.6	0.1502
RETAIL TRADE	110831 94.08	6980 5.92	117811	74.47	971.5	0.1392
FINANCE INSURANCE AND REAL ESTATE	45320 90.24	4899 9.76	50219	81.83	624.6	0.1275
EDUCATION AND HEALTH SERVICES	35840 94.26	2181 5.74	38021	90.76	385.2	0.1766
OTHER SERVICES	146415 92.91	11175 7.09	157590	77.73	1303.6	0.1167
TOTAL	599440 93.48	41807 6.52	641247	80.15	2150.8	0.0514

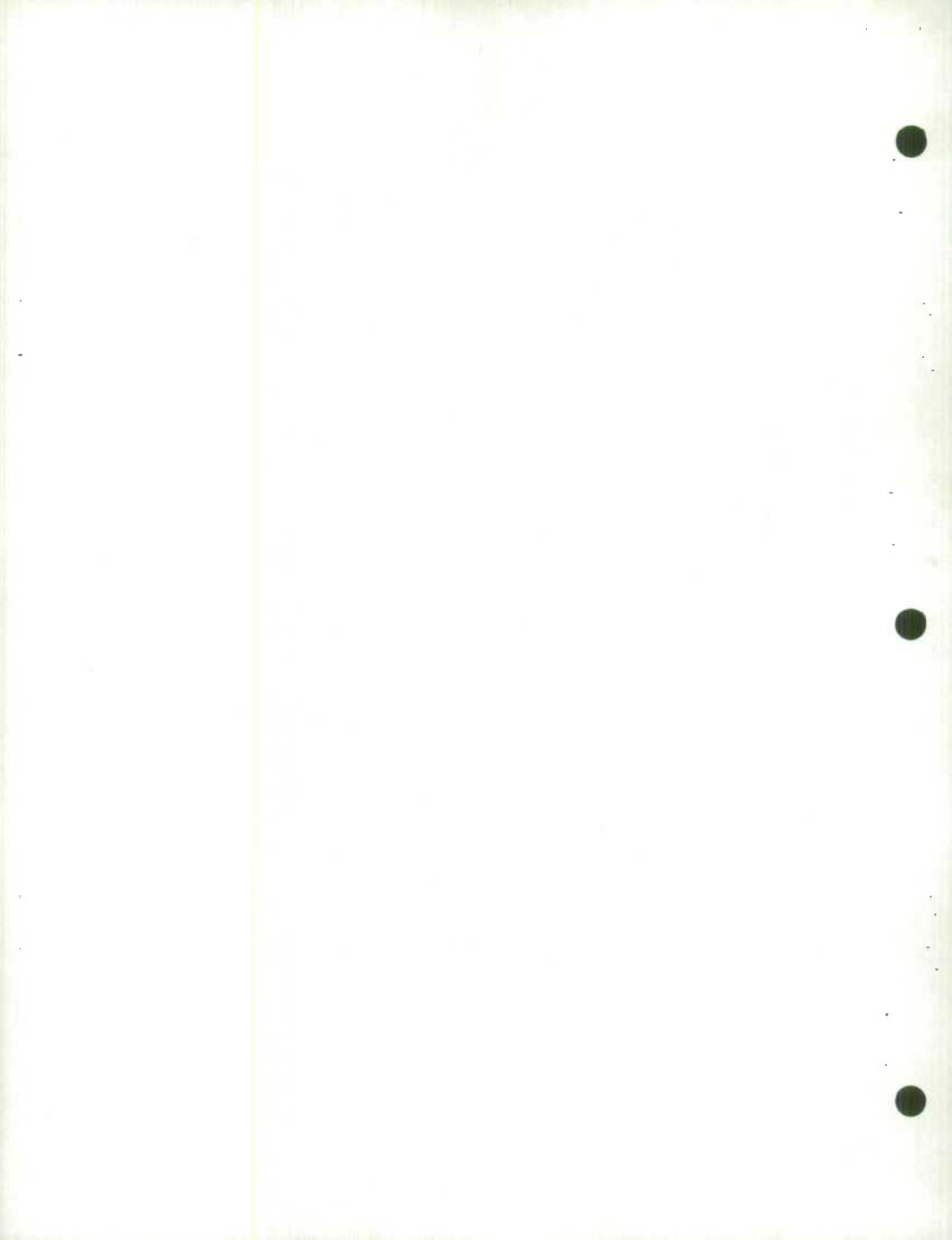


Table B.4

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE IN 1970 SIC CODE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN				ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY
	NO CHANGE IN SIC	CHANGE AT FIRST LEVEL (FIRST DIGIT)	CHANGE AT SECOND LEVEL (SECOND DIGIT)	CHANGE AT THIRD LEVEL (THIRD DIGIT)		
AGRICULTURE	40695 61.18	4304 6.47	881 1.32	20641 31.03	66521	89.61
FORESTRY	5238 85.79	643 10.53	193 3.16	32 0.53	6106	79.99
FISHING AND TRAPPING	3701 89.24	367 8.86	53 1.28	26 0.63	4147	82.14
MINING	2082 52.98	1249 31.79	260 6.62	339 8.63	3930	78.87
MANUFACTURING	33981 82.57	4380 10.64	1284 3.12	1510 3.67	41155	88.95
CONSTRUCTION	68890 82.45	7389 8.84	6794 8.13	477 0.57	83550	76.42
TRANSPORTATION	24759 89.34	2036 7.35	394 1.42	525 1.90	27714	83.19
COMMUNICATIONS AND OTHER UTILITIES	1768 74.62	474 20.00	91 3.85	36 1.54	2369	85.06
WHOLESALE TRADE	29872 70.93	5142 12.21	5060 12.02	2040 4.84	42114	80.57
RETAIL TRADE	89890 76.30	8832 7.50	11966 10.16	7123 6.05	117811	74.47
FINANCE INSURANCE AND REAL ESTATE	37008 73.69	4987 9.93	4199 8.36	4025 8.01	50219	81.83
EDUCATION AND HEALTH SERVICES	33732 88.72	872 2.29	1163 3.06	2254 5.93	38021	90.76
OTHER SERVICES	122120 77.49	14253 9.04	9394 5.96	11823 7.50	157590	77.73
TOTAL	493736 77.00	54928 8.56	41732 6.51	50851 7.93	641247	80.15

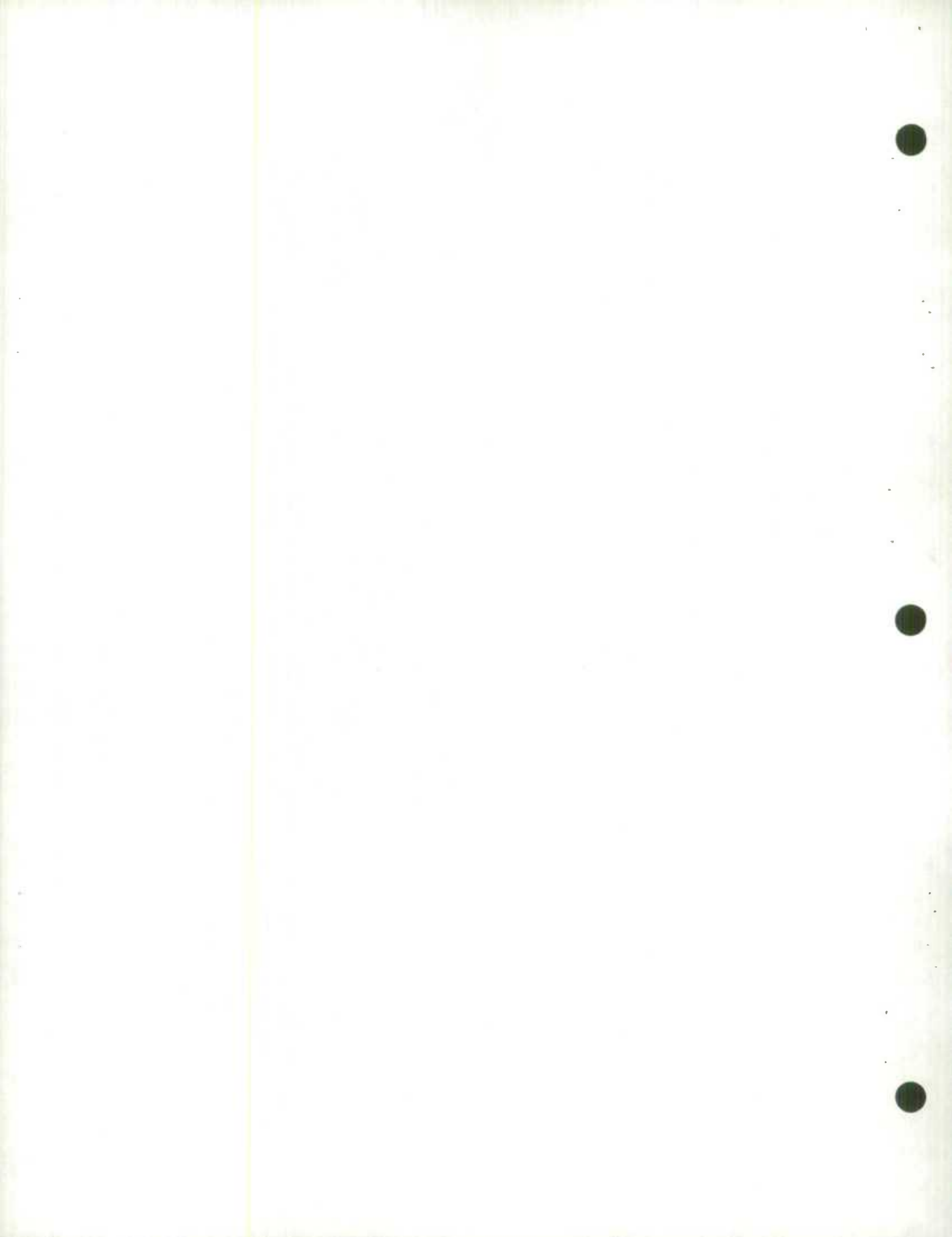


Table B.5.1 ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE IN 1970 SIC CODE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN SIC	
	NO CHANGE IN SIC	CHANGE IN SIC			STANDARD ERROR	CV
AGRICULTURE	40695 61.18	25826 38.82	66521	89.61	1276.0	0.0494
FORESTRY	5238 85.79	868 14.21	6106	79.99	154.7	0.1783
FISHING AND TRAPPING	3701 89.24	446 10.76	4147	82.14	101.3	0.2272
MINING	2082 52.98	1848 47.02	3930	78.87	170.0	0.0920
MANUFACTURING	33981 82.57	7174 17.43	41155	88.95	672.4	0.0937
CONSTRUCTION	68890 82.45	14660 17.55	83550	76.42	1223.9	0.0835
TRANSPORTATION	24759 89.34	2955 10.66	27714	83.19	417.4	0.1412
COMMUNICATIONS AND OTHER UTILITIES	1768 74.62	601 25.38	2369	85.06	90.4	0.1503
WHOLESALE TRADE	29872 70.93	12242 29.07	42114	80.57	868.7	0.0710
RETAIL TRADE	89890 76.30	27921 23.70	117811	74.47	1801.8	0.0645
FINANCE INSURANCE AND REAL ESTATE	37008 73.69	13211 26.31	50219	81.83	946.6	0.0717
EDUCATION AND HEALTH SERVICES	33732 88.72	4289 11.28	38021	90.76	525.6	0.1225
OTHER SERVICES	122120 77.49	35470 22.51	157590	77.73	2168.4	0.0611
TOTAL	493736 77.00	147511 23.00	641247	80.15	3701.2	0.0251

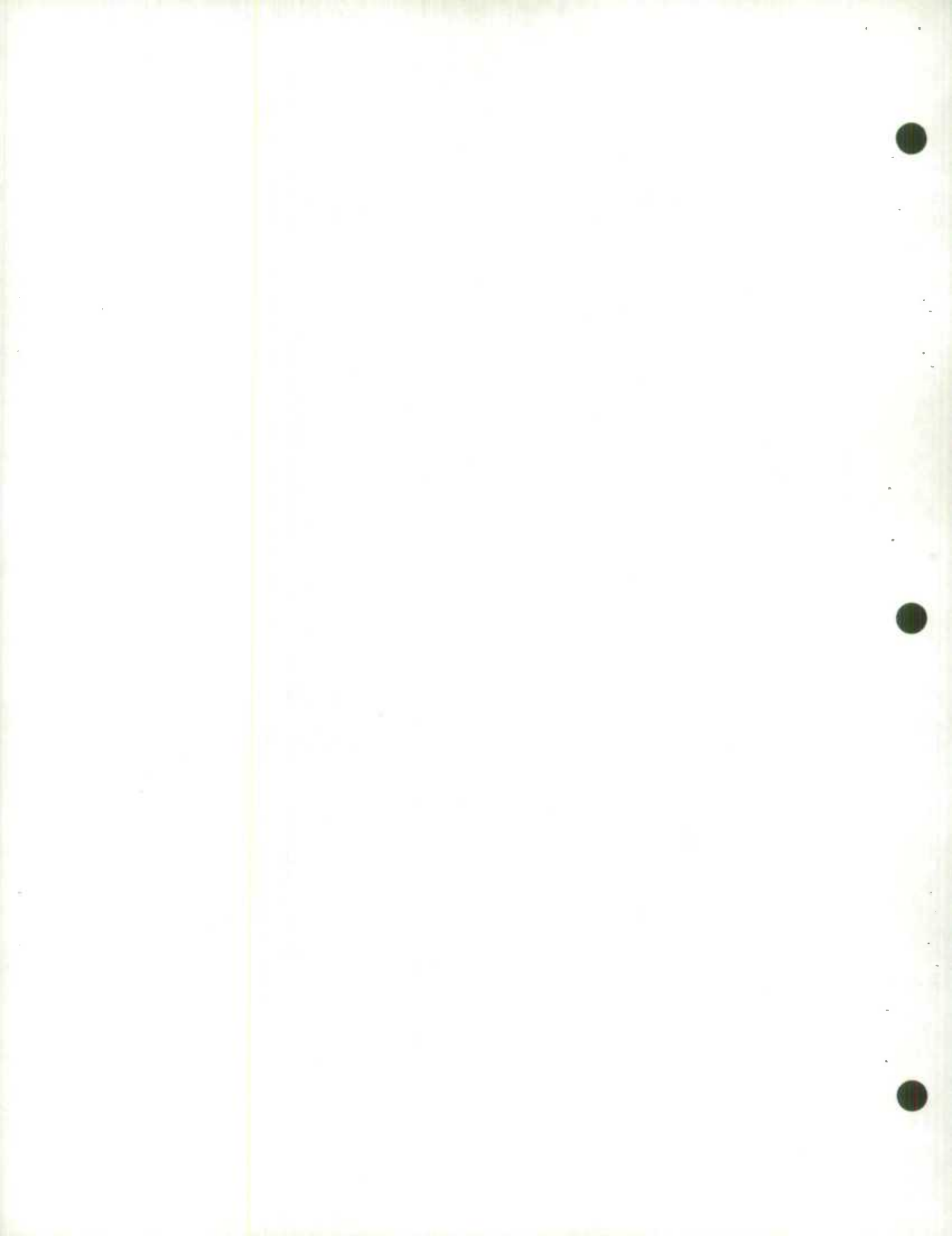


Table B.5.2

INDUSTRY GROUP	NO CHANGE IN SIC		CHANGE IN SIC		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY
	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED NUMBER OF UNITS PERCENTAGE OF THE TOTAL CHANGE IN SIC			
	NO CHANGE IN INDUSTRY GROUP		NO CHANGE IN INDUSTRY GROUP	CHANGE IN INDUSTRY GROUP		
AGRICULTURE	40695 61.18		21228 82.20	4598 17.80	66521	89.61
FORESTRY	5238 85.79		32 3.69	836 96.31	6106	79.99
FISHING AND TRAPPING	3701 89.24		26 5.83	420 94.17	4147	82.14
MINING	2082 52.98		547 29.60	1301 70.40	3930	78.87
MANUFACTURING	33981 82.57		3852 53.69	3322 46.31	41155	88.95
CONSTRUCTION	68890 82.45		7270 49.59	7390 50.41	83550	76.42
TRANSPORTATION	24759 89.34		919 31.10	2036 68.90	27714	83.19
COMMUNICATIONS AND OTHER UTILITIES	1768 74.62		36 5.99	565 94.01	2369	85.06
WHOLESALE TRADE	29872 70.93		3917 32.00	8325 68.00	42114	80.57
RETAIL TRADE	89890 76.30		14815 53.06	13106 46.94	117811	74.47
FINANCE INSURANCE AND REAL ESTATE	37008 73.69		8224 62.25	4987 37.75	50219	81.83
EDUCATION AND HEALTH SERVICES	33732 88.72		2617 61.02	1672 38.98	38021	90.76
OTHER SERVICES	122120 77.49		17330 48.86	18140 51.14	157590	77.73
TOTAL	493736 77.00		80812 54.78	66699 45.22	641247	80.15

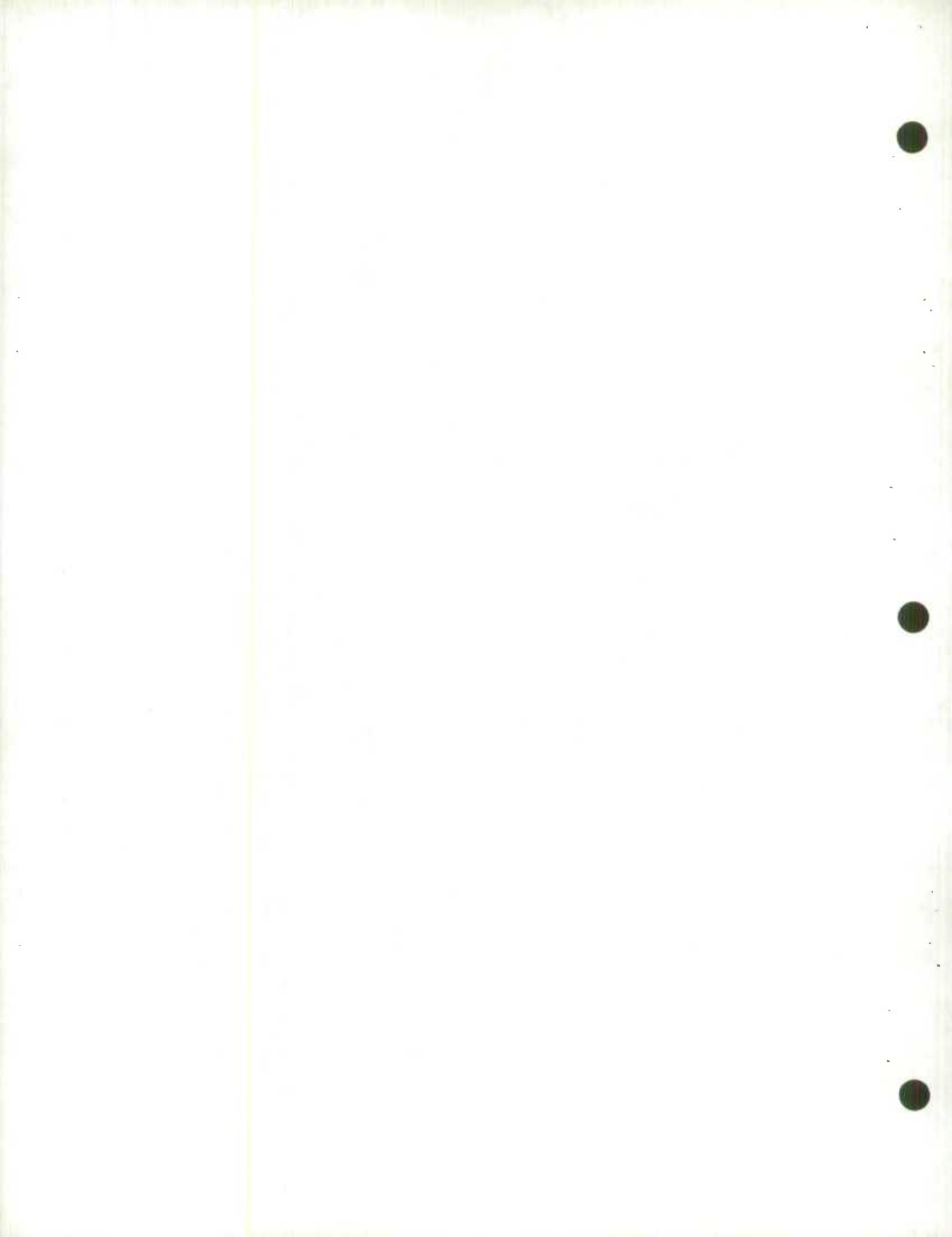


Table B.6

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE
IN INDUSTRY GROUP

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN INDUSTRY GROUP	
	NO CHANGE IN INDUSTRY GROUP	CHANGE IN INDUSTRY GROUP			STANDARD ERROR	CV
AGRICULTURE	61923 93.09	4598 6.91	66521	89.61	646.4	0.1406
FORESTRY	5270 86.31	836 13.69	6106	79.99	152.2	0.1821
FISHING AND TRAPPING	3727 89.87	420 10.13	4147	82.14	98.6	0.2348
MINING	2629 66.90	1301 33.10	3930	78.87	154.9	0.1190
MANUFACTURING	37833 91.93	3322 8.07	41155	88.95	479.7	0.1444
CONSTRUCTION	76160 91.15	7390 8.85	83550	76.42	902.2	0.1221
TRANSPORTATION	25678 92.65	2036 7.35	27714	83.19	351.7	0.1728
COMMUNICATIONS AND OTHER UTILITIES	1804 76.15	565 23.85	2369	85.06	88.3	0.1563
WHOLESALE TRADE	33789 80.23	8325 19.77	42114	80.57	750.9	0.0902
RETAIL TRADE	104705 88.88	13106 11.12	117811	74.47	1303.7	0.0995
FINANCE INSURANCE AND REAL ESTATE	45232 90.07	4987 9.93	50219	81.83	629.7	0.1263
EDUCATION AND HEALTH SERVICES	36349 95.60	1672 4.40	38021	90.76	339.4	0.2030
OTHER SERVICES	139450 88.49	18140 11.51	157590	77.73	1630.1	0.0899
TOTAL	574548 89.60	66699 10.40	641247	80.15	2661.3	0.0399

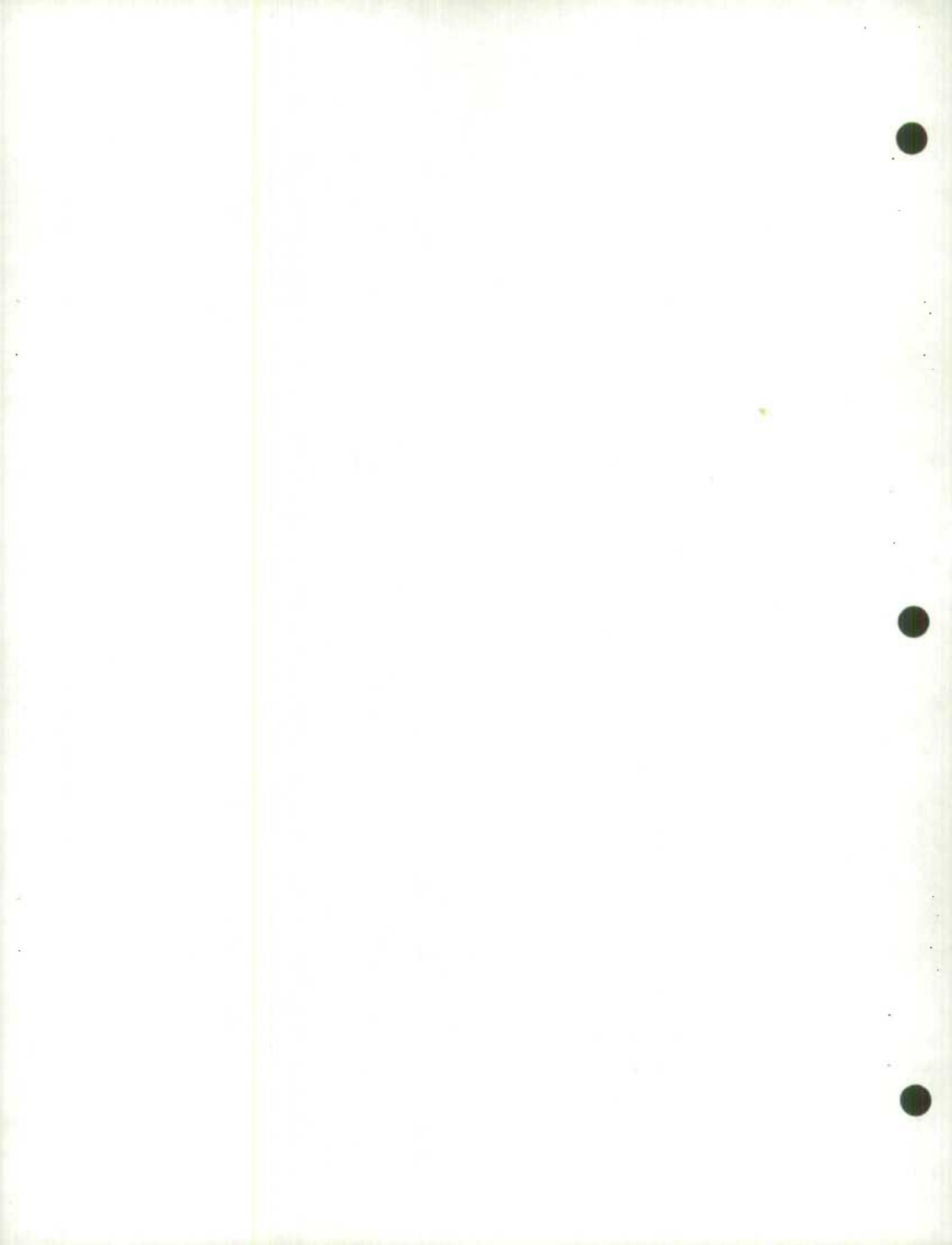


Table B.7

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE
IN EMPLOYMENT SIZE CODE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN EMPLOYMENT SIZE CODE	
	NO CHANGE IN EMPLOYMENT SIZE CODE	CHANGE IN EMPLOYMENT SIZE CODE			STANDARD ERROR	CV
AGRICULTURE	40679 61.15	25842 38.85	66521	89.61	1278.8	0.0495
FORESTRY	4426 72.49	1680 27.51	6106	79.99	202.2	0.1203
FISHING AND TRAPPING	2932 70.70	1215 29.30	4147	82.14	152.8	0.1258
MINING	2045 52.04	1885 47.96	3930	78.87	172.1	0.0913
MANUFACTURING	24481 59.48	16674 40.52	41155	88.95	893.0	0.0536
CONSTRUCTION	53927 64.54	29623 35.46	83550	76.42	1604.1	0.0541
TRANSPORTATION	17280 62.35	10434 37.65	27714	83.19	683.6	0.0655
COMMUNICATIONS AND OTHER UTILITIES	1543 65.13	826 34.87	2369	85.06	100.7	0.1218
WHOLESALE TRADE	24007 57.00	18107 43.00	42114	80.57	976.8	0.0539
RETAIL TRADE	71405 60.61	46406 39.39	117811	74.47	2157.1	0.0465
FINANCE INSURANCE AND REAL ESTATE	30836 61.40	19383 38.60	50219	81.83	1073.4	0.0554
EDUCATION AND HEALTH SERVICES	26418 69.48	11603 30.52	38021	90.76	777.0	0.0670
OTHER SERVICES	95504 60.60	62086 39.40	157590	77.73	2640.7	0.0425
TOTAL	395483 61.67	245764 38.33	641247	80.15	4462.5	0.0182

Table B.8

ESTIMATION OF THE TOTAL NUMBER OF UNITS CHANGING
SEPH EMPLOYMENT SIZE GROUP

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS CHANGING SEPH EMPLOYMENT SIZE GROUP	
	NO CHANGE IN SEPH SIZE GROUP	CHANGE IN SEPH SIZE GROUP			STANDARD ERROR	CV
AGRICULTURE	62198 93.50	4323 6.50	66521	89.61	629.4	0.1456
FORESTRY	5686 93.12	420 6.88	6106	79.99	111.5	0.2654
FISHING AND TRAPPING	3856 92.98	291 7.02	4147	82.14	83.4	0.2872
MINING	3532 89.87	398 10.13	3930	78.87	96.7	0.2429
MANUFACTURING	35016 85.08	6139 14.92	41155	88.95	631.3	0.1028
CONSTRUCTION	77867 93.20	5683 6.80	83550	76.42	803.5	0.1414
TRANSPORTATION	24590 88.73	3124 11.27	27714	83.19	430.4	0.1378
COMMUNICATIONS AND OTHER UTILITIES	2038 86.03	331 13.97	2369	85.06	71.3	0.2157
WHOLESALE TRADE	37935 90.08	4179 9.92	42114	80.57	557.9	0.1335
RETAIL TRADE	110053 93.41	7758 6.59	117811	74.47	1025.8	0.1322
FINANCE INSURANCE AND REAL ESTATE	46254 92.10	3965 7.90	50219	81.83	568.5	0.1434
EDUCATION AND HEALTH SERVICES	34080 89.63	3941 10.37	38021	90.76	507.1	0.1287
OTHER SERVICES	141700 89.92	15890 10.08	157590	77.73	1543.5	0.0971
TOTAL	584805 91.20	56442 8.80	641247	80.15	2446.6	0.0433

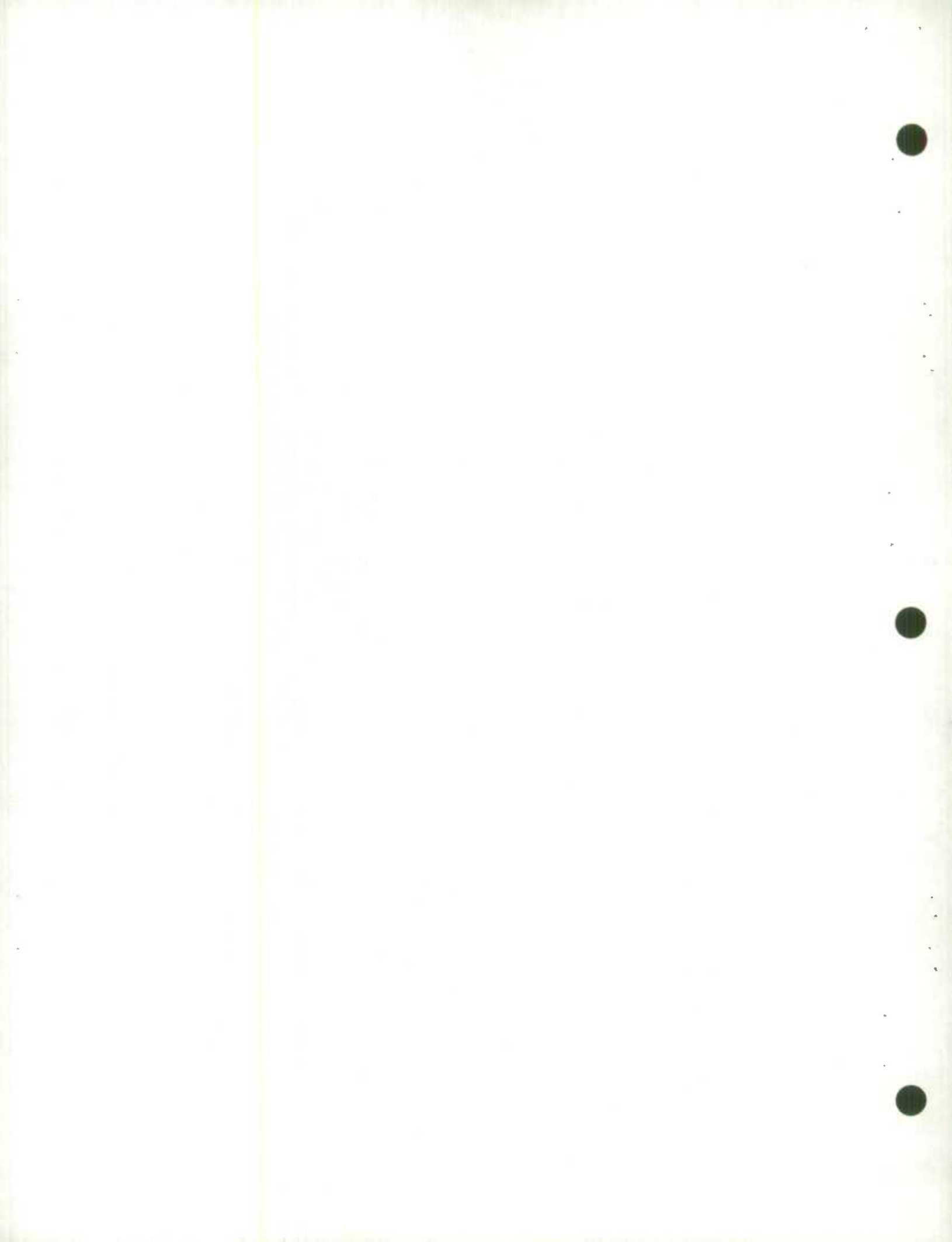


Table B.9

ESTIMATION OF THE TOTAL NUMBER OF UNITS REQUIRING
A NON-LEGAL UPDATE TO THE LABEL NAMETHE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO LEGAL,
FINANCIAL, OR CORPORATE CHANGES

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS REQUIRING A NON-LEGAL UPDATE TO THE LABEL NAME	
	NO CHANGE IN LABEL NAME	CHANGE IN LABEL NAME			STANDARD ERROR	CV
AGRICULTURE	62217 96.07	2543 3.93	64760	87.24	487.8	0.1918
FORESTRY	5752 96.75	193 3.25	5945	77.89	76.6	0.3970
FISHING AND TRAPPING	3963 98.05	79 1.95	4042	80.06	44.3	0.5623
MINING	3670 95.92	156 4.08	3828	76.78	61.6	0.3943
MANUFACTURING	37806 94.50	2190 5.50	39796	86.01	394.5	0.1801
CONSTRUCTION	77233 95.57	3576 4.43	80809	73.91	639.4	0.1788
TRANSPORTATION	25679 96.54	919 3.46	26598	79.84	240.8	0.2617
COMMUNICATIONS AND OTHER UTILITIES	2186 96.77	73 3.23	2259	81.11	35.1	0.4811
WHOLESALE TRADE	37217 93.63	2530 6.37	39747	76.04	440.7	0.1742
RETAIL TRADE	103423 92.13	8832 7.87	112255	70.96	1086.0	0.1230
FINANCE INSURANCE AND REAL ESTATE	46457 95.16	2362 4.84	48819	79.55	443.4	0.1877
EDUCATION AND HEALTH SERVICES	35259 94.17	2181 5.83	37440	89.37	385.2	0.1766
OTHER SERVICES	136858 90.08	15063 9.92	151921	74.93	1497.8	0.0994
TOTAL	577521 93.42	40697 6.58	618218	77.27	2198.9	0.0540



Table B.10

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE
IN LOCATION POSTAL CODE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES
AND FOR WHICH THE LABEL ADDRESS IS AT LEAST LOCATION

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN IN LOCATION POSTAL CODE	
	NO CHANGE IN LOCATION POSTAL CODE	CHANGE IN LOCATION POSTAL CODE			STANDARD ERROR	CV
AGRICULTURE	58793 98.04	1174 1.96	59967	80.78	334.6	0.2850
FORESTRY	5592 99.43	32 0.57	5624	73.68	31.6	0.9830
FISHING AND TRAPPING	3281 97.65	79 2.35	3360	66.55	44.3	0.5623
MINING	3357 100.00	0 0.00	3357	67.37	0.0	.
MANUFACTURING	35794 98.34	604 1.66	36398	78.66	210.9	0.3491
CONSTRUCTION	69248 97.98	1430 2.02	70678	64.65	408.5	0.2856
TRANSPORTATION	22394 94.72	1248 5.28	23642	70.97	278.8	0.2234
COMMUNICATIONS AND OTHER UTILITIES	1968 97.28	55 2.72	2023	72.64	30.5	0.5574
WHOLESALE TRADE	35013 97.06	1061 2.94	36074	69.01	289.6	0.2729
RETAIL TRADE	96870 96.59	3419 3.41	100289	63.39	688.0	0.2012
FINANCE INSURANCE AND REAL ESTATE	40683 96.88	1312 3.12	41995	68.43	333.4	0.2541
EDUCATION AND HEALTH SERVICES	31551 96.44	1163 3.56	32714	78.09	284.9	0.2449
OTHER SERVICES	129247 97.79	2915 2.21	132162	65.19	680.2	0.2333
TOTAL	533791 97.36	14492 2.64	548283	68.53	1271.7	0.0877

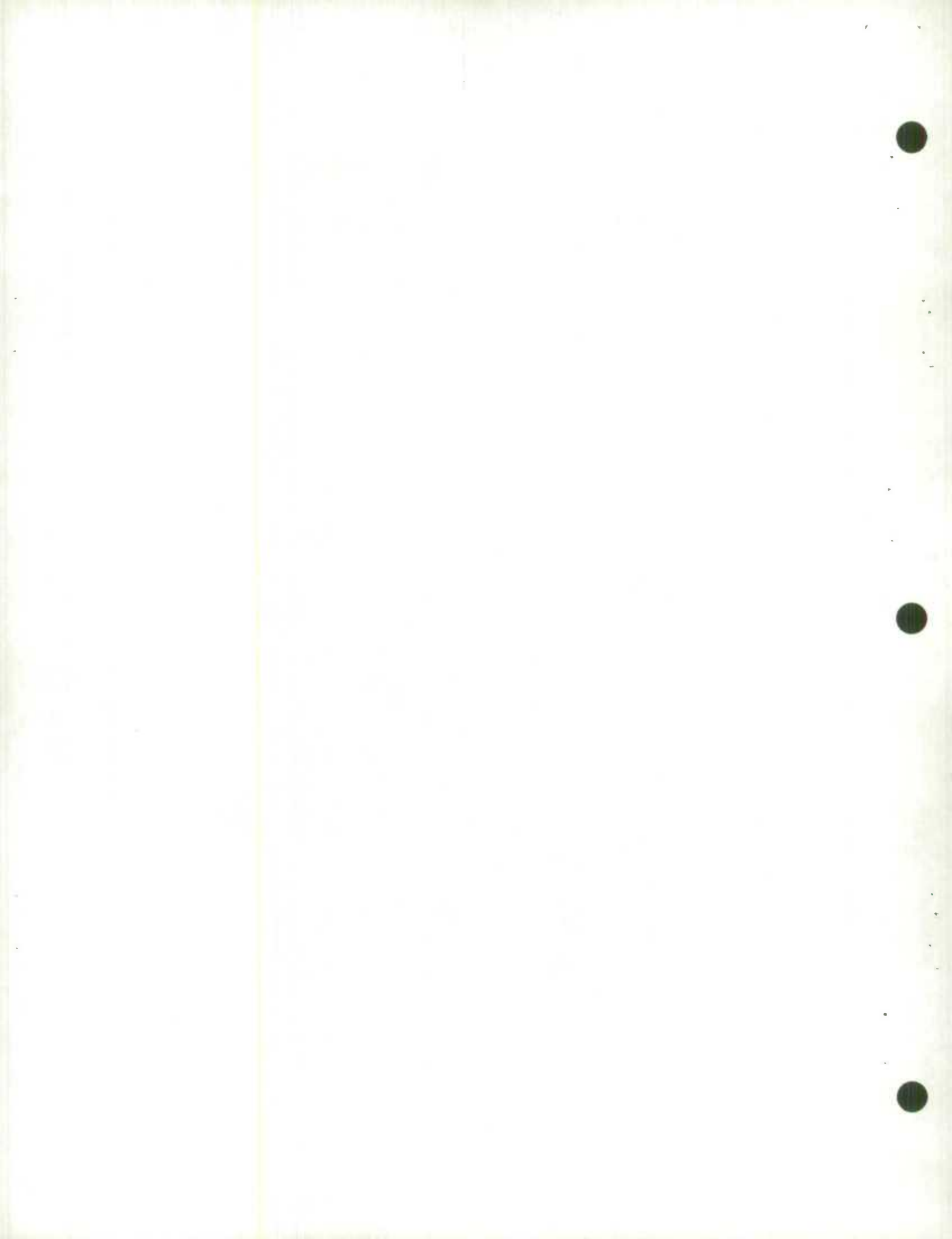


Table B.11

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE IN SGC

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES
AND FOR WHICH THE LABEL ADDRESS IS AT LEAST LOCATION

ESTIMATED NUMBER OF UNITS
PERCENTAGE OF DOMAIN

INDUSTRY GROUP	NO CHANGE IN SGC	CHANGE AT FIRST LEVEL (FIRST 2 DIGITS)	CHANGE AT SECOND LEVEL (NEXT 2 DIGITS)	CHANGE AT THIRD LEVEL (LAST 3 DIGITS)	ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY
AGRICULTURE	58989 98.37	0 0.00	880 1.47	98 0.16	59967	80.78
FORESTRY	5624 100.00	0 0.00	0 0.00	0 0.00	5624	73.68
FISHING AND TRAPPING	3308 98.45	0 0.00	52 1.55	0 0.00	3360	66.55
MINING	3357 100.00	0 0.00	0 0.00	0 0.00	3357	67.37
MANUFACTURING	35945 98.76	0 0.00	302 0.83	151 0.41	36398	78.66
CONSTRUCTION	69129 97.81	119 0.17	834 1.18	596 0.84	70678	64.65
TRANSPORTATION	22920 96.95	0 0.00	525 2.22	197 0.83	23642	70.97
COMMUNICATIONS AND OTHER UTILITIES	1987 98.22	0 0.00	36 1.78	0 0.00	2023	72.64
WHOLESALE TRADE	35258 97.74	0 0.00	571 1.58	245 0.68	36074	69.01
RETAIL TRADE	98722 98.44	0 0.00	1282 1.28	285 0.28	100289	63.39
FINANCE INSURANCE AND REAL ESTATE	41208 98.13	0 0.00	612 1.46	175 0.42	41995	68.43
EDUCATION AND HEALTH SERVICES	31842 97.33	0 0.00	581 1.78	291 0.89	32714	78.09
OTHER SERVICES	129571 98.04	0 0.00	2105 1.59	486 0.37	132162	65.19
TOTAL	537860 98.10	119 0.02	7780 1.42	2524 0.46	548283	68.53

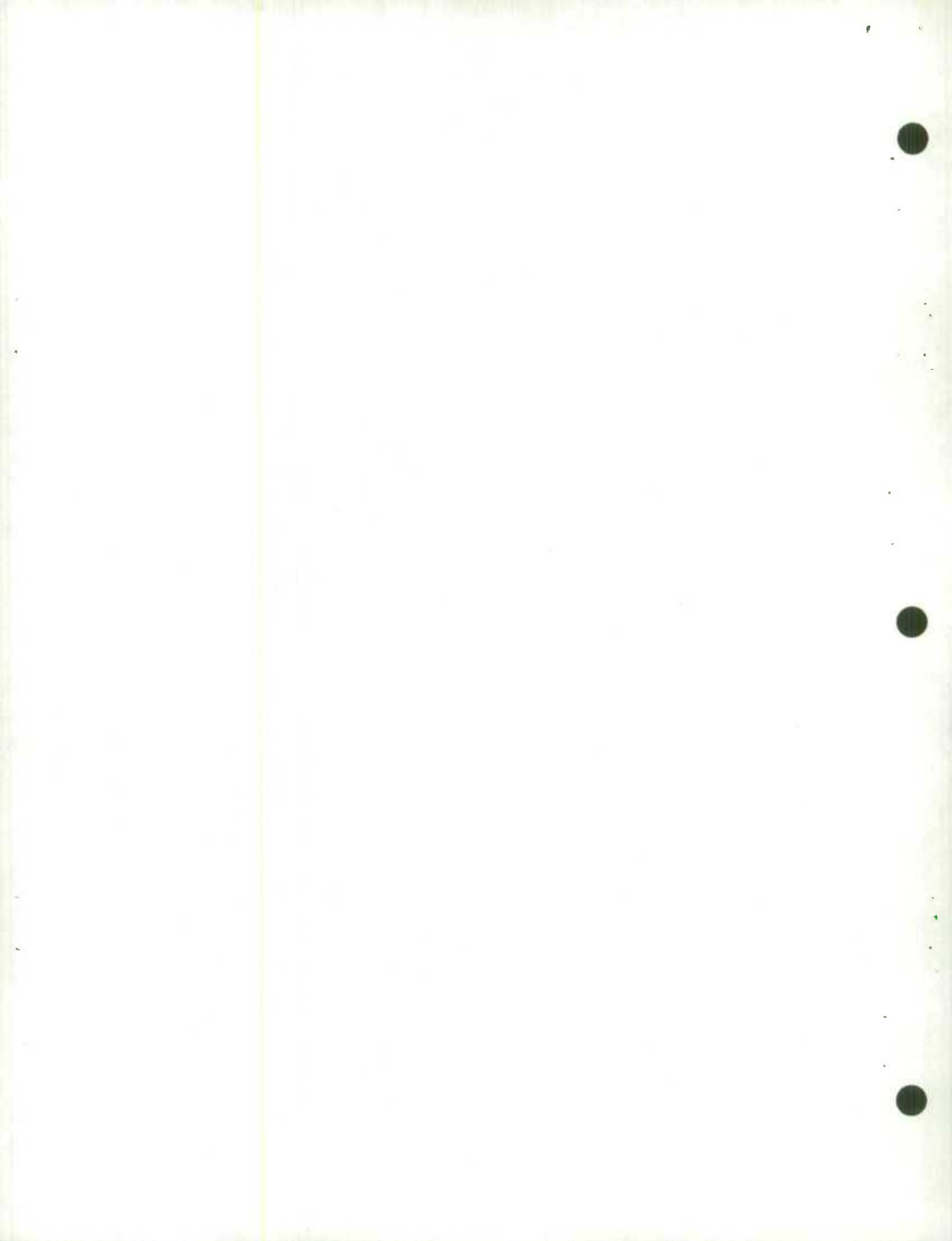


Table B.12

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE IN SGC

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES
AND FOR WHICH THE LABEL ADDRESS IS AT LEAST LOCATION

INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN SGC	
	NO CHANGE IN SGC	CHANGE IN SGC			STANDARD ERROR	CV
AGRICULTURE	58989 98.37	978 1.63	59967	80.78	305.8	0.3126
FORESTRY	5624 100.00	0 0.00	5624	73.68	0.0	.
FISHING AND TRAPPING	3308 98.45	52 1.55	3360	66.55	36.2	0.6905
MINING	3357 100.00	0 0.00	3357	67.37	0.0	.
MANUFACTURING	35945 98.76	453 1.24	36398	78.66	183.0	0.4038
CONSTRUCTION	69129 97.81	1549 2.19	70678	64.65	425.0	0.2743
TRANSPORTATION	22920 96.95	722 3.05	23642	70.97	213.9	0.2961
COMMUNICATIONS AND OTHER UTILITIES	1987 98.22	36 1.78	2023	72.64	25.0	0.6850
WHOLESALE TRADE	35258 97.74	816 2.26	36074	69.01	254.6	0.3120
RETAIL TRADE	98722 98.44	1567 1.56	100289	63.39	468.6	0.2990
FINANCE INSURANCE AND REAL ESTATE	41208 98.13	787 1.87	41995	68.43	259.4	0.3294
EDUCATION AND HEALTH SERVICES	31842 97.33	872 2.67	32714	76.09	247.6	0.2838
OTHER SERVICES	129571 98.04	2591 1.96	132162	65.19	641.8	0.2477
TOTAL	537860 98.10	10423 1.90	548283	68.53	1090.5	0.1046

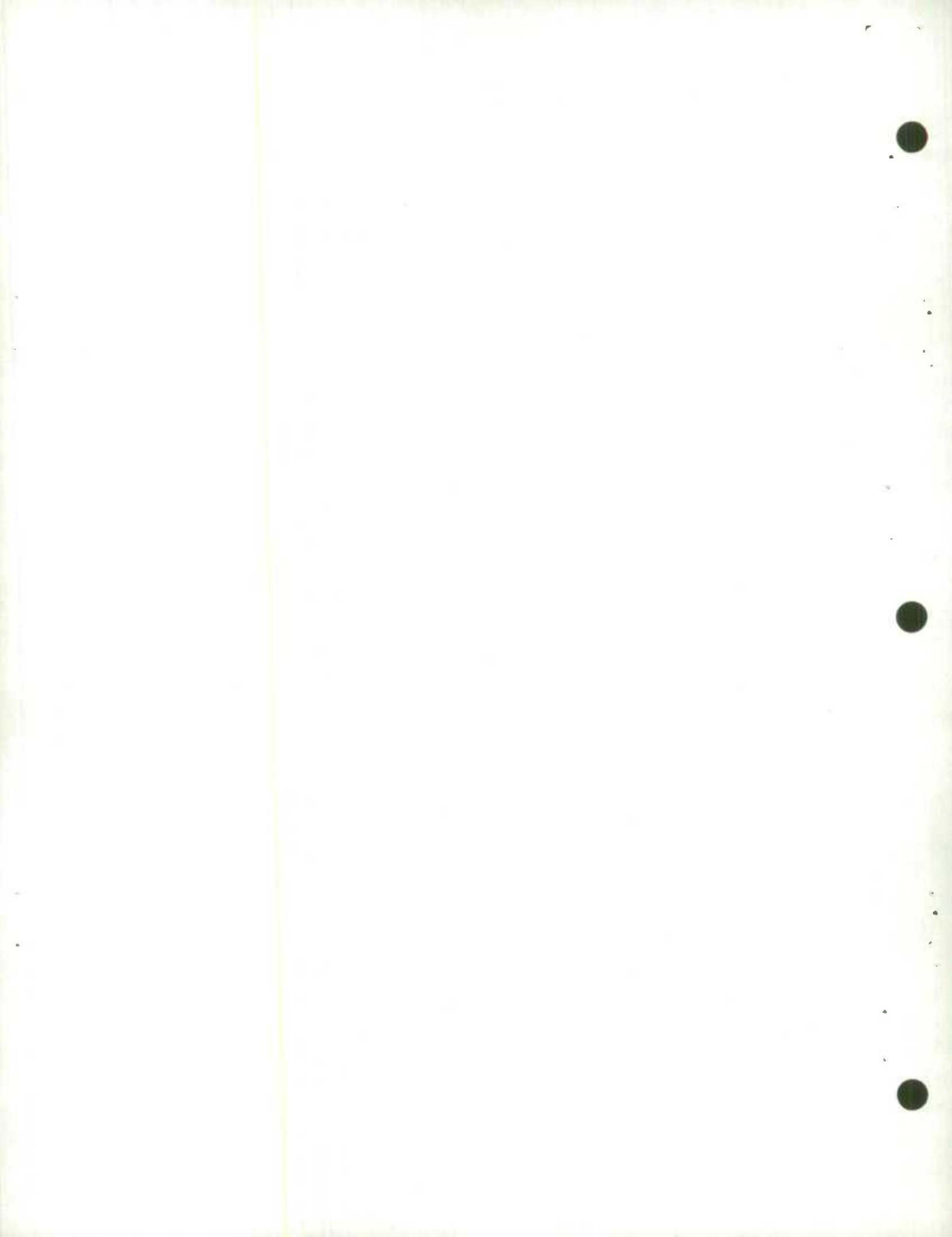
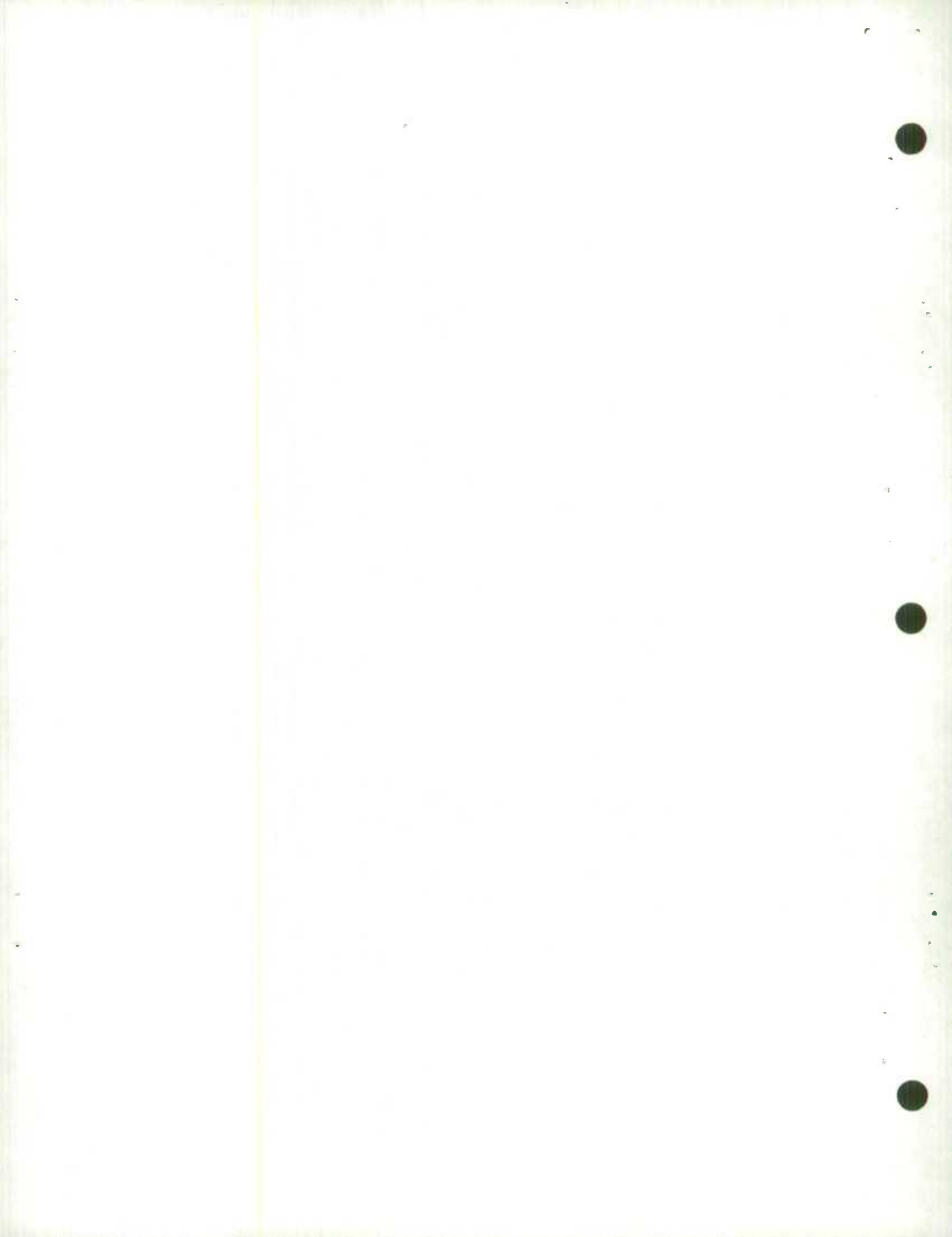


Table B.13

ESTIMATION OF THE TOTAL NUMBER OF UNITS WITH A CHANGE IN PROVINCE CODE

THE DOMAIN IS THE SET OF ALL ACTIVE UNITS WITH NO MAJOR LEGAL CHANGES
AND FOR WHICH THE LABEL ADDRESS IS AT LEAST LOCATION

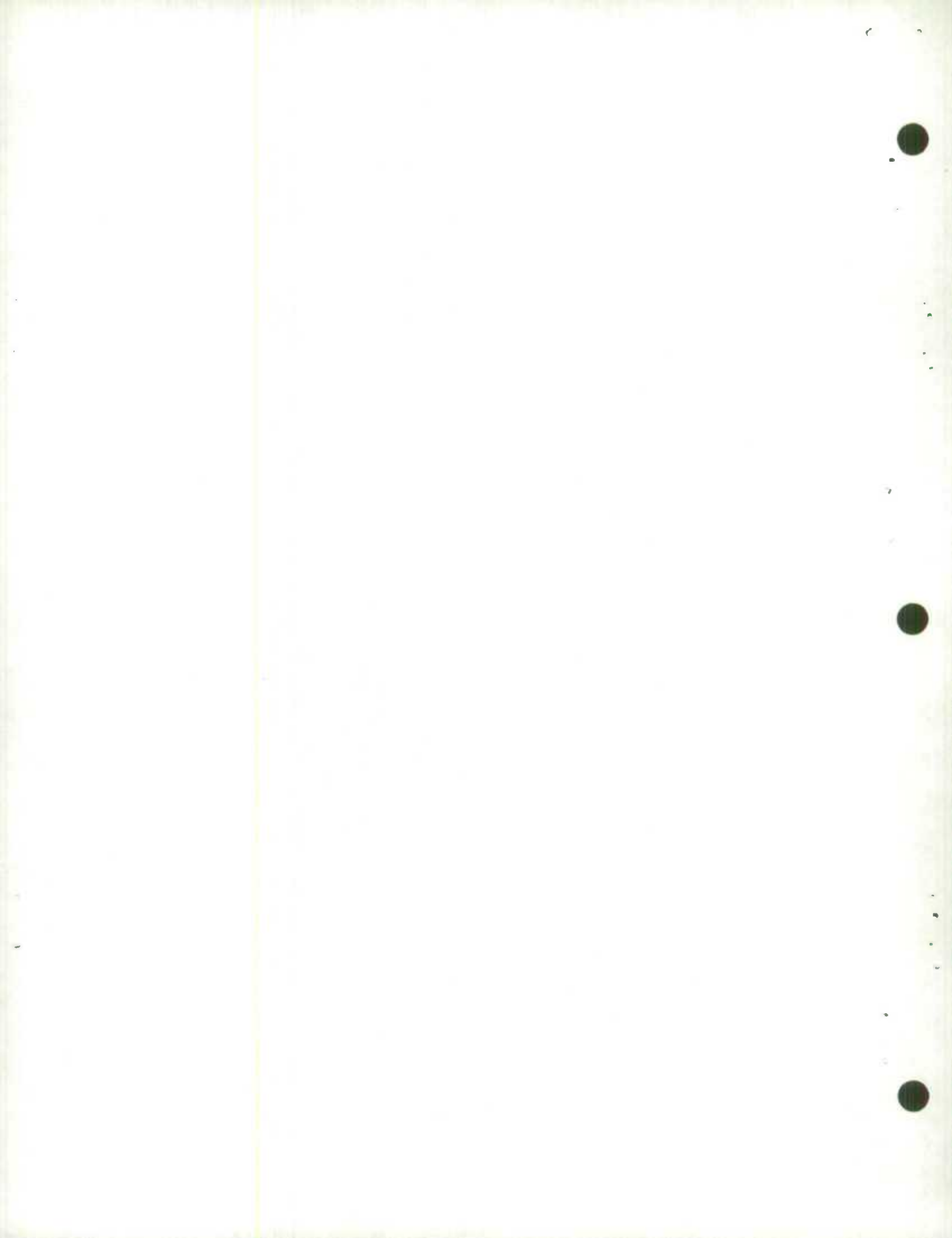
INDUSTRY GROUP	ESTIMATED NUMBER OF UNITS PERCENTAGE OF DOMAIN		ESTIMATED DOMAIN TOTAL	DOMAIN AS ESTIMATED PERCENTAGE OF INDUSTRY	ESTIMATES OF RELIABILITY FOR THE ESTIMATED NUMBER OF UNITS WITH A CHANGE IN PROVINCE CODE	
	NO CHANGE IN PROVINCE CODE	CHANGE IN PROVINCE CODE			STANDARD ERROR	CV
AGRICULTURE	59967 100.00	0 0.00	59967	80.78	0.0	.
FORESTRY	5624 100.00	0 0.00	5624	73.68	0.0	.
FISHING AND TRAPPING	3360 100.00	0 0.00	3360	66.55	0.0	.
MINING	3357 100.00	0 0.00	3357	67.37	0.0	.
MANUFACTURING	36398 100.00	0 0.00	36398	78.66	0.0	.
CONSTRUCTION	70559 99.83	119 0.17	70678	64.65	118.5	0.9955
TRANSPORTATION	23576 99.72	66 0.28	23642	70.97	65.1	0.9919
COMMUNICATIONS AND OTHER UTILITIES	2023 100.00	0 0.00	2023	72.64	0.0	.
WHOLESALE TRADE	36074 100.00	0 0.00	36074	69.01	0.0	.
RETAIL TRADE	100289 100.00	0 0.00	100289	63.39	0.0	.
FINANCE INSURANCE AND REAL ESTATE	41733 99.38	262 0.62	41995	68.43	150.4	0.5731
EDUCATION AND HEALTH SERVICES	32714 100.00	0 0.00	32714	78.09	0.0	.
OTHER SERVICES	132162 100.00	0 0.00	132162	65.19	0.0	.
TOTAL	547836 99.92	447 0.08	548283	68.53	202.3	0.4525



Appendix C

Estimate of Variance for the Result in

Section 5.2



Estimate

$$\hat{y} = \left(\frac{x_{11}}{x_{11} + x_{12}} \right) \times \left(\frac{x_{11} + x_{12} + x_{21} + x_{22}}{N} \right) \times N$$

Approximate Estimate of Variance of \hat{y}

$$\hat{Var}(\hat{y}) \cong \left(\frac{N}{n} \right)^2 \left[1 + \frac{x_{12} \times (x_{21} + x_{22})}{(x_{11} + x_{12})^2} \right]^2 \times \hat{Var}(x_{11})$$

$$+ \left(\frac{N \times x_{11}}{n} \right)^2 \left[\frac{x_{21} + x_{22}}{(x_{11} + x_{12})^2} \right]^2 \times \hat{Var}(x_{12})$$

$$+ \left(\frac{N \times x_{11}}{n} \right)^2 \left[\frac{1}{x_{11} + x_{12}} \right]^2 \times \hat{Var}(x_{21} + x_{22})$$

$$- 2 \times \left[\frac{N}{n} \times \left\{ 1 + \frac{x_{12} \times (x_{21} + x_{22})}{(x_{11} + x_{12})^2} \right\} \right] \times \left[\frac{N \times x_{11}}{n} \times \frac{(x_{21} + x_{22})}{(x_{11} + x_{12})^2} \right] \times \hat{Cov}(x_{11}, x_{12})$$

$$+ 2 \times \left[\frac{N}{n} \times \left\{ 1 + \frac{x_{12} \times (x_{21} + x_{22})}{(x_{11} + x_{12})^2} \right\} \right] \times \left[\frac{N \times x_{11}}{n} \times \frac{1}{(x_{11} + x_{12})} \right] \times \hat{Cov}(x_{11}, x_{21} + x_{22})$$

$$- 2 \times \left[\frac{N \times x_{11}}{n} \times \frac{(x_{21} + x_{22})}{(x_{11} + x_{12})^2} \right] \times \left[\frac{N \times x_{11}}{n} \times \frac{1}{(x_{11} + x_{12})} \right] \times \hat{Cov}(x_{12}, x_{21} + x_{22})$$

$$\text{where } \hat{Var}(x_{11}) = \left(\frac{N-n}{N} \right) \times \left(\frac{N-1}{N} \right) \times \frac{n^2}{(n-1)} \times \left(\frac{x_{11}}{N} \right) \times \left(1 - \frac{x_{11}}{N} \right)$$

etc.

$$\text{and } \hat{Cov}(x_{11}, x_{12}) = - \left(\frac{N-n}{N} \right) \times \left(\frac{N-1}{N} \right) \times \frac{n^2}{(n-1)} \times \left(\frac{x_{11}}{N} \right) \times \left(\frac{x_{12}}{N} \right)$$

etc.

005

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA



1010148862