Institute of Economic Development and Research SCHOOL OF ECONOMICS University of the Philippines System

Discussion Paper NO. 75-15

September 1975

THE EFFECTS OF AGE MISSTATEMENT ON ESTIMATES OF NET MIGRATION

by

Bryan L. Boulier

NOTE: IEDR Discussion Papers are prelimary versions circulated privately to elicit critical comment. References in publications to Discussion Papers should be cleared with the authors.

THE EFFECTS OF AGE MISSTATEMENT ON ESTIMATES OF NET MIGRATION

by

Bryan L. Boulier

There are many procedures for estimating net Among the most widely used are variants of migration. the census survival method. The procedure for estimating net migration to an area using the forward census survival method is to project the initial population of the area to the date of the second census and to subtract the projected population, from the enumerated population, the difference being estimated net migration. The survival rates for the projection are calculated from the enumerated national populations of the appropriate age groups at the two census dates (e.g., the ratio of the number age 15-19 in 1/970 to the number age 5-9 in 1960.) It has been demonstrated that census survival methods implicitly take into account the relative coverage of the two censuses as well as age misstatement, so that it is very often the case that these methods are preferred to other residual estimation techniques such as vital statistics or life table survival procedures (Hamilton and Anderson, 1944; Zachariah, 1962; and Hamilton, 1966.) There has been much analysis of the biases in the estimates of the volume of net migration rates when the two censuses have differing

degrees of enumeration error (Price, 1955; Zachariah, 1962; Hamilton, 1966; and Stone, 1967.) It has been assumed that the consequences of age misstatement are identical to the consequences of enumeration error since the transfer of persons out of an age-group by age misstatement is analogous to an undercount of the population in the age group by a similar percent. This note presents for the first time an explicit derivation of the effects of age misstatement on estimates of the volume of net migration and net migration rates and demonstrates that results which have been obtained in the analysis of enumeration error are not always applicable to age misstatement. The estimates from the census survival rate (CSR) method are compared with those obtained from the use of life table survival rates (LTSR).

Assumptions and Definitions.

We shall analyze a simple case of age misstatement where a certain fraction $\underline{\alpha}$ of an age group is reported as belonging to an older age group. Such age transfers may arise, for instance, from deliberate misstatement of age, rounding of ages, or inaccuracy of estimation by enumerators who have to guess the ages of respondents not knowing their ages. It will be assumed that there are no

* .

有大大 医电影通讯 医克雷斯 医克雷斯氏 医克斯克氏病 医电影 医电影克克氏病 医电影 ាក់ស ព្រះក្រុមស្រុសស្រុស និក្សា នេះ និស្សាស្សាក្តី ក្រុមស្រុសស្រុស និស្សាស្សាក្សា ស្រុសស្រុសស្រុសស្រុសស្រុសស្ San file langer i solfforgum och bli och ann er engo enn er i jedliften br The standard of the control of the standard of · Marin Astronomist (1997年) はい Begras に Marina に Astronomist (1997年) Astronomist (19 "我们们,我们<mark>是我们的</mark>我们的是一个,我们的,我们的人们也不是一个人,我们的人们的,我们就是我们的 TO MORE FOR ONLY THE CONTRACT CONTRACT OF THE CONTRACT C अपने कल रेक्ट्रांस्ट्रिक्ट (१८००) १८ देखा । १९५० च्या १००० १००० १००० १००० १०० १००० १**०० १००**० व्यक्ति हो हो हो। is about one of a service of the ser oration for the second state of the second s "自然来来的","我们是这种的,你们的**我生,我的**我们们,我们们也没有更多。""我们们,你们们们也没有是 The state of the s * * 20 * 2 * 25 * Litter (14 * 5) 1 * 1 * 2 .

Martin Adams to the contract

(1) 日本の書から、 (1) 日本の書から、 (2) 日本の書から、 (2) 日本の書から、 (2) 日本の書から、 (3) 日本の書から、 (4) 日本の書から、 (4

that there is a complete count of the population in both censuses, and that all migration occurs at the end of the intercensal period. Zachariah (1962) examines the consequences of differential mortality, previously cited authors have considered cases of incomplete enumeration, and the last assumption ensures that in the absence of age misstatement the forward estimate of migration correctly estimates the volume of net migration. Siegel and Hamilton (1952) analyze other patterns of the timing of migration and the results obtained here are easily generalized to those patterns.

We shall consider four age groups in a population when the length of the age interval equals the length of the intercensal period (e.g., five year age groups and five year census intervals.) Let us define the following terms:

- P(x,t) = true national population in age group x at time t (x = 1, 2, 3, 4; t = 1, 2),
- P*(x,t) = reported national population age group x time t,

 s_x = true survival rate from age group x to x + 1,

 M_{x} = true net migration age x,

 \widehat{H}_{x} = estimated net migration age x from LTSR method.

It is assumed that migrants move in age group x but are age x + 1 at the date of the second census. Finally, it is assumed that a fraction α of persons age x = 2 are erroneously reported as being age x = 3.

Census Survival Rate Estimates of Net Migration.

Table 1 shows the true and reported national and regional populations at both census dates. The CSR estimates of net migration are:

(1)
$$M_X^* = p^*(x + 1, t + 1) - s_X^*(x,t) p^*(x,t)$$
.

The National CSR are:

(2)
$$s_1^* = \frac{P*(2,2)}{P*(1,1)} = s_1 (1 - \alpha),$$

(3)
$$s_2^* = \frac{P*(3,2)}{P*(2,1)} = \frac{s_2}{1-\alpha} + \frac{\alpha s_1}{1-\alpha} \left[\frac{P(1,1)}{P(2,1)} \right]$$
, and

(4)
$$s_3^* = \frac{p*(4,2)}{p*(3,1)} = \frac{s_3}{1+a \frac{p(3,1)}{p(2,1)}}$$

The CSR are rather complex, s_1^* is the product of the true—survival rate s_1 and one minus the fraction transferred to the next older age group. s_2^* is a combination of the true survival rate, the age misstatement factor, and the national age distribution. s_3^* is also a function of the national age distribution.

Table 2 shows the estimated volume of net migration to the region calculated from equation (1) after some simplification of the resulting expressions. There are several interesting results which may be derived from the table. First, consider M_1^* . The estimated volume of migration is a fraction of true migration depending upon the proportion of persons whose ages are misstated. The smaller is , the closer is estimated to true migration. A migration rate which relates migration to the end of the period population $(1-M_1)$ relates migration to the end of the period population $(1-M_1)$ is unbiased; a migration rate which relates estimated migration to the beginning population $(1-M_1)$ and which would be used for population projection is biased towards zero. (See Hamilton (1965) for a discussion of alternative measures of migration rates.)

The results for M_1^* are similar to those obtained in the analysis of enumeration error. If the population is underenumerated by a fraction α for both censuses, the volume of net migration is underestimated by the same proportion but migration rates c culated from the end of period or beginning of period populations will be unbiased. For age misstatement, as was noted, the latter rate is biased.

Now we turn to migration at other ages. M_2^* equals the true volume of net migration plus terms which include the survival rate s_1 , the regional and national age distributions, migration into the preceding age bracket, and the extent of age misstatement. If the proportionate age distributions of the regional and national population in the relevant age groups are identical at the time of the first census, the term in square brackets equals zero, and $M_2^* = M_2 + \alpha M_1$, so that M_2^* is very close to M_2 if α is small, M_1 is small, or both. If however, the age distribution are identical the bracketed term is non-zero and is weighted by α $s_1p(2,1)$. Even if α is small, the value of the term could be large since p(2,1) is ordinarily large. There is no presumption as to the direction of bias in the estimator M_2^* , and rates calculated from M_2^* will generally

be biased. The relationship between M_3^* and M_3 depends upon the age distributions of the regional and national populations at the time of the first census, the survival rate s_3 , and α . If the proportionate age distributions of the national and regional populations in the relevant age intervals are equal at the time of the first census, the term in the square brackets equals zero, and $M_3^* = M_3$. If the term is non-zero, then the smaller is α the closer is M_3^* to M_3 . But even if α is small, the error may be large since α times the bracketed term is multiplied by $s_3p(3,1)/(1 * \frac{p(2,1)}{p(3,1)})$. As in the case of M_2^* , there is no presumption as to the direction of the bias in the estimator M_3^* and the rates calculated from it will generally be biased.

Life Table Survival Rate Estimates of Net Migration.

Despite the biases note above, it should not be forgotten, however, that the CSR method is likely to provide estimates superior to those calculated from LTSR methods. Suppose we assume that the national population has the same age distribution as the stationary population associated with the life table so that the s_{χ} may be identified with life table survival rates. The forward

estimates of migration using LTSR and the enumerated regional population are also shown in Table 2. The estimates include terms of the form α s_x p(x,1) which may well be large even if α is small. For instance, consider $\hat{\mathbb{M}}_1 = (1-\alpha)$ $\mathbb{M}_1 - \alpha$ s₁p(1,1). Persons who are transferred to an older age group are counted as outmigrants. The CSR method implicitly adjusts for this age transfer. $\hat{\mathbb{M}}_2$ counts as in-migrants persons who transfer from age x = 2 at the time of the second census (α s₁p(1,1) + α M₁) and the survivors of the persons who misstated their age at the time of the first census (α s₂p(2,1)). $\hat{\mathbb{M}}_3$ counts as outmigrants the persons who were erroneously enumerated as part of age group x = 3 at the time of the first census (α p(2,1)) timess₃, the survival rate appropriate for survival from $x_3^2 = 3$ to x = 4.

Example.

Table 3 presents estimates of net migration using hypothetical data. The national population is assumed to be stationary with $s_1 = .99$, $s_2 = .98$, and $s_3 = .97$, and it is assumed that five percent of the age group x = 2 is erroneously reported as belonging to age group x = 3. The CSR's calculated from the enumerated national population are $s_1^* = .9405$, $s_2^* = 1.0842$, and $s_3^* = .9229$. The

regional population is assumed to have the same true sursame pattern of age misreporting as vival rates and the the national population. At time t = 1, the regional population is constructed by taking 10 percent of the national population age x = 1, 12 percent of the number x = 2, 11 percent of the number x = 3, and 10 percent of the number x = 4. Part A of the table gives the basic data, part B compares true net migration with migration estimated by the CSR wethod and by the LTSR method, and part C compares migration rates. It is readily seen from parts B and C of the table that the CSR estimates are superior to the LTSR estimates. $\widehat{\mathbb{N}}_1$ is less than half of M_1 , \widehat{M}_2 is more than three times M_2 , and \widehat{M}_3 , is negative while M_3 is Marge and positive. In contrast the CSR estimates are all within ten percent of the correct values. Nonetheless, it is important to note that in the presence of modest age misstatement and age distributions of the regional and national populations which are not very dissimilar, the errors in the CSR estimates of net migration and net migration rates are large enough to be important for some analyses.

Summary.

In this paper, we have shown the consequences

of age misstatement for census survival rate and life table survival rate estimates of net migration. In the presence of age misstatement, CSR estimates of migration and migration rates are generally biased although the direction of bias is not always clear. Moreover, estimation errors due to age misstatement are not necessarily similar to those which result from underenumeration of a population. Although CSR estimates of net migration are usually superior to those obtained by LTSR methods, the above analysis suggests that in some cases it is probably worthwhile to attempt to adjust the data for age misreporting and then to apply the LTSR method for additional estimates to be compared with those obtained by the CSR procedure.

REFERENCES

- Hamilton, C. Horace. 1965. "Practical and Mathematical Considerations in the Formulation and Selection of Migration Rates," <u>Demography</u>, II, 429-463.
- Hamilton, C. Horace. 1966. "Effect of Census Errors on the Measurement of Net Migration," <u>Demography</u>, III. 393-415.
- 3. Hamilton, C. Horace and F.M. Anderson, 1944. "Use of Survival Rate Method in Measuring Net Migration,"

 Journal of the American Statistical Association,

 XXXIX, 197-206.
- 4. Price, Daniel O. 1955. "Examination of Two Sources of Error in the Estimation of Net Internal Migration,"

 Journal of the American Statistical Association, L, 689-700.
- 5. Siegel, Jacobs and C. Horace Hamilton. 1952. "Some Considerations in the Use of the Residual Method of Estimating Net Migration," <u>Journal of the American Statistical Association</u>, XLVII, 475-500.
- Stone, Leroy O. 1967. "Evaluating the Relative Accuracy and Significance of Net Migration Estimates," <u>Demography</u>, IV, 310-330.
- 7. Zachariah, K.C. 1962. "A Note on the Census Survival Ratio Method of Estimating Net Migration," <u>Journal of the American Statistical Association</u>, LVII, 175-183.

TABLE	TABLE 1. True and Reported:			
		National and Regions		
True	Formal Population	oronal Population	SI	
t = 1 P(1,1)	P(2,1)			
t = 2 P(1,2)	*s1P(1,1)	The state of the s		
Reported		P(3,1) syp(2,1)	P(4,1)	
$t : 1 P^*(1,1) = P(1,1)$	P*(2 1)*((163)	P(4,2)*	
t = 2 P*(1,2)=P(1,2)	$P^*(2,1)^*(1-\alpha)P(2,1)$	P*(3,1)=p(3,1)	3863	
	=(1-a)s1 P(1,1)	P*(3,2)=P(3,2)+ aP(2,1)		
True	Regional Population	*2P(2,1)+ %S1P(1,1)		
t = 1 p(1,1)	p(2,1)	•		
t = 2 p(1,2)	p(2,2)=s ₁ p(1,1)+ M.	p(3,1)		
Reported	;·1	p(3,2)=8,p(2,1)+y	p(4,1)	
$c : 1 p^*(1,1) = p(1,1)$	p*(2,1)=(1,2)	Z ₁₁₁ , / - 6	P (4,2)=8,2p.	
t : 2 p*(1,2)=p(1,2)	$p^*(2,2)=(1-a)p(2,1)$ = $(1-a)p(2,2)$	p*(3,1) + p(2,1)		

	TABLE 1	l. True and Reported: Na	TABLE 1. True and Reported: National and Regional Populations	S
True		National Population		
t = 1	P(1,1)	P(2,1)	P(3,1)	P(4,1)
t = 2	P(1,2)	P(2,2)=s ₁ P(1,1), P(1,1), P(1,1)	P(3,1)=s ₂ P(2,1)	P(4,2)=s ₃ P(3,1)
Reported	rol .			
н н	P*(1,1)=P(1,1)	$P^*(2,1)^{\pi}(1-\alpha)P(2,1)$	$P*(3,1)=P(3,1)+ \alpha P(2,1)$	P*(4,1)=P(4,1)
ւ ն	P*(1,2)=P(1,2)	$P*(2,1)=(1-\alpha)P(2,2)$	$P*(3,2)=P(3,2)+\alpha P(2,2)$	$P*(4,2)=s_3P(3,1)$

#82P(2	$=(1-\alpha)s_1 P(1,1)$		
P*(3,2)=P(3,2	$P^*(2,1)=(1-a)P(2,2)$	P*(1,2)=P(1,2)	t 1
P*(3,1)=P(3,1	$P*(2,1)=(1-\alpha)P(2,1)$	P*(1,1)=P(1,1)	н п

	National Population		
	P(2,1) P(2,2)=s ₁ P(1,1)	P(3,1) P(3,1)*s ₂ P(2,1)	P(4,1) P(4,2)=s ₃ P(3,
(1,1)	$P*(2,1)=(1-\alpha)P(2,1)$	P*(3,1)=P(3,1)+ aP(2,1)	P*(4,1)=P(4,
(1,2)	$P*(2,1)=(1-\alpha)P(2,2)$ = (1-\alpha)S1 P(1.1)	$P*(3,2)=P(3,2)+\alpha P(2,2)$ = $SOP(2,1)+$	P*(4,2)=s ₃ P(

щ	щ	
$P*(3,1)=P(3,1)+ \alpha P(2,1)$	P*(3,2)=P(3,2)+ aP(2,2) =s2P(2,1)+ as1P(1,1)	
$P*(2,1)*(1-\alpha)P(2,1)$	$P*(2,1)=(1-\alpha)P(2,2)$ =(1-\alpha)s ₁ P(1,1)	
P*(1,1)=P(1,1)	P*(1,2)=P(1,2)	
1 11	ti 13	

 $p(4,2)=s_3p(3,1)+M_3$

p(4,1)

 $p^*(4,2)$ =s3p(3,1)+M3

p*(3,2) = p(3,2) + p(2,2)= $s_2p(2,1) +$ $a | s_1p(1,1) + M_1 + M_2$

 $p^*(2,2)=(1-\alpha)p(2,2)$ = $(1-\alpha) s_1p(1,1)+M_1$

 $p^*(2,1)=(1-a)p(2,1)$

 $p^{*}(1,1)=p(1,1)$

رد ::

Reported

p(1,2)

t = 2

p(1,1)

ر <u>.</u> ...

True

 $p^*(1,2)=p(1,2)$

 $p^*(4,1)=p(4,1)$

 $p^*(3,1) + p(2,1)$

A. Census Survival Method

$$M_{1}^{*} = (1 - \alpha)$$

$$M_{2}^{*} = M_{2} + \alpha M_{1} + \alpha s_{1} p(2,1) \cdot \left[\frac{p(1,1)}{p(2,1)} - \frac{P(1,1)}{P(2,1)} \right]$$

$$M_{3}^{*} = M_{3} + \frac{\alpha s_{3} p(3,1)}{1 + \alpha \frac{P(2,1)}{P(3,1)}} \cdot \left[\frac{P(2,1)}{P(3,1)} - \frac{p(2,1)}{p(3,1)} \right]$$

B. Life Table Survival Method

$$M_1 = (1 - \alpha) M_1 - \alpha s_1 p (1,1)$$
 $M_2 = M_2 + \alpha M_1 + \alpha s_1 p (1,1) + \alpha s_2 p (2,1)$
 $M_3 = M_3 - \alpha s_3 p (2,1)$

TABLE 3. An Illustrative Example of Age Misstatement and Estimation of Net Migration by Census Survival Rate and Life Table Survival Rate Methods

A. <u>Data</u>

	Nat	ional E	opulati	Lon .		R	egional	Populati	lon
		True	Rep	ported			ue		rted
Age	/ <u>Time</u>	(t=1,2	(t	=1,2)		t=1	t=2	t=1	t=2
	- 1	100,00		,000		10,000	10,000	10,000	10,00
	• 2	99,00		,050		11,880	•	11,286	10,35
	= 3	97,02		.,970		10,672	12.142	11,266	12,68
x	= 4	94,10	9 94	,109		9,411	10,852	9,411	10,85
В.	Number	of Net	Migran	its					
			3						
		3	······································	,		x=1	x =2	X≅	3
	True (Mx)	я •			1,000	500	50	0
	•	Surviv	1 Rate	!		•		-	
		hod (Mx	*			950	451	45	5
		able Şu		Rate		, , , ,	-732	43	_
		hod (Mg				455	1,627	- 7	4
	MC C	noa (nga	Ø .			433	1,027	-,	O
		33 \$4. ★	4,						
C.	Net Mi	gration	Rates (1)					
	•	True		Census	Survi	<u>lva1</u>	Life	Table	
	M ₁	M ₂	Мз	M ₁ *	M ₂ *	M ₃ *	м ₁	M ₂	М3
Begi	inning							•	
	.100	.042	.047	.095	.040	.040	.046	.144	007
End	of Per	iod							
	.092		.046	.096	.036	.042	.044	.128	007
				-			- · ·		

⁽¹⁾ Beginning of period rates are calculated from $M_{\rm X}/p({\rm x},1)$ and end of the period rates are calculated from $M_{\rm X}/p({\rm x}+1,2)$. The true rates use true migration and the true regional population; the CSR and LTSR estimates of migration rates use the reported regional population as denominators.

Common Co	
A Transis (A SARAMAN) - 10 Sac a SARAMAN) Sac at a SARAMAN	
en och i videlligen viden i videlligen program videlligent	
polytic management specific and general specific constitutions.	
Company - Marketine Company - Company Company - Company Company - Company Company	
Vol. o 1 Ko rdino Con 11 - Alberton Sci. o 1 Alberton	
Conserve - Re disses equals consideration energy - Applic	
ng med i ma dalapi a agantu na <mark>hasalan</mark> Cilah mangal apia	
econ commente concern vendamente concern a quanto	
engles str atoure: pendru i ny stain puest e nyst <mark>ain</mark>	
edalia: 32 delagram militro confessionia: montali respektario:	
crea - creatine. neturi-ciptule. cotto-ciptule.	
Chair or an elocation Table 2 - Alexandre A NOTA - Chair Alexandre A NOTA - Chair Alexandre	
Second Indiament Second a present agency a present agency Arabbas	
cons. consense: trois consense: drocs resident: drocs resident:	
SERVICE CONTRACTOR AND CONTRA	
Septim sidentification Description of the description State Septimization	
projektivis delegen engles projektivis projektivis managanis	
Colores of Williams Colores of Marines Colores of Marines	
Chestics consistence contract in deligation contract in the sequence contract in the sequence	
erke son de de arabitet in La de de La de arabitet in de arabitet	
Control Colored (Action Colored Colore	
Contract of Manager Contract	
Comment of	
Charac Assembler The Control of Control The Control of Control The	
control of the property of the control of the contr	
Company of the Compan	
Coperation of Conference Coperation of Coperation Coperation of Cope	
per type per medigalisation in e to the term to management the period of the publication	
Experience (experience) material consideration material (experience)	
condition to confidence in the confidence of the	
Executive Sendings Record Colleges Sendings Sendings	
Construct of Century,	
Make and companies and a second secon	
- And Andrews -	
COMPANY CONTRACTOR CON	
Andrew - Addition Only Continue All Cont	
The state of millions of the state of the st	
e de analogo de Ca ragoline Constante de la caragoline Constante de analogo de Caragoline	
Hardware (1986) (1986) Objection (1986) (1986) William (1986)	
The Control of Control	
production of the designation of the control of the	
- construction of the cons	
Sealer (Sealethine) Sealer (Sealethine) Sealer (Sealethine)	
1996 - Washing 1996 - Washing 1996 - Washing 1996 - Washing	
The second of th	
repeate a 12 repeateur - 12 repeateu	
100m	
ero de arto, e en en contrato lemento de la contrato de arto	
- Marian - Americanismo Spiriture - Marianismo AMERIA - Marianismo	
of mount or or (Constitution)	
General Control Applica	