2017-1

Generalization of the Census Survival Ratio Method and Decomposability of Net Migration

Takashi Inoue

January 2017

Working Paper Series, Institute of Economic Research
Aoyama Gakuin University

Generalization of the Census Survival Ratio Method and Decomposability of Net Migration

Takashi Inoue

Department of Public and Regional Economics
College of Economics
Aoyama Gakuin University
4-4-25 Shibuya, Shibuya-ku
Tokyo, 150-8366 JAPAN

Abstract

The census survival ratio method is valuable for estimating the volume of migration from only static statistics, that is, census data. Particularly, in the case of areas where or in the case of periods when we cannot obtain migration data, this method has a high utility value. However, since estimators calculated using the method are net migration and cannot be decomposed into in- and out-migrations nor into long- and short-distance migrations, the results of its application often disappoint us. Although the method has such a disadvantage, it has scarcely been modified or generalized since its frame was established in 1940s.

This paper generalizes the method by reconsidering its preconditions, establishes the concept of decomposability of net migration, and decomposes the CoCSIR that is a demographic measure developed by the author. As a result of that, we understand that internal net migration of a municipality can be clearly decomposed into long- and short-distance migrations that are different in direction from one another, and that the decomposing method has a crucial advantage.

1. Introduction

The census survival ratio method (hereafter, referred to as CSR method) is valuable for estimating the volume of migration from only static statistics, that is, census data. Particularly, in the case of areas where or in the case of periods when we cannot obtain migration data, this method has a high utility value. However, since estimators calculated using the method are net migration and cannot be decomposed into in- and out-migrations nor into long- and short-distance migrations, the results of its application often disappoint us. For example, if the volume of in-migration into an area is the same as that of out-migration from the area, no matter how large it is, it is not observable at all because the volume of net migration in the area becomes a zero value. Although the method has such a disadvantage, it has scarcely been modified or generalized since its frame was established in 1940s.

Thus, this paper attempts to generalize CSR method by reconsidering its preconditions, and to enable us to decompose net migration calculated using the generalized method. After formulating CSR method in Chapter 2 and generalizing the method in Chapter 3, this paper establishes the concept of decomposability of net migration in Chapter 4 through the discussion in Chapters 2 and 3, and attempts to decompose the CoCSIR (Cohort Cumulative Social Increase Ratio) that is a demographic measure developed by the author[1], by applying the concept to it.

[1] The author developed the CoCSIR in 2002, and attempted to formulate it in 2014. Shimizu (2006, 2009) examined the efficacy of the measure by using actual data.

2. Formulation of the Census Survival Ratio Method

Although CSR method are separated into the forward, reverse, and average method according to how to calculate survival ratios, there is no essential difference among these three methods. Accordingly, this paper has all discussions on the basis of the forward method, which is the most standard of them. As is well known, the following two preconditions are required for using CSR method: the population of a study area is closed, i.e., there is no migration between a study area and its outside; death rates show a uniform distribution in a study area. This method is how to estimate the volume of net migration in a subarea of the study area under these two preconditions. To formulate the frame of the method, the study area and subareas are considered to be one nation and municipalities constituting the nation, respectively, and some variables are defined for a certain cohort as below:

$p_i(t)$: the population of municipality *i* at time *t*;

$p_i(t+1)$: the population of municipality *i* at time *t*;

$d_i(t,t+1)$: the number of deaths in municipality *i* in the period from time *t* to time *t*+1;

$m_i(t,t+1)$: the volume of net migration in municipality *i* in the period from time *t* to time *t*+1;

$m_i^{\mathrm{INTL\_I}}(t,t+1)$: the volume of immigration into municipality *i* from the outside of a nation in the period from time *t* to time *t*+1;

$m_i^{\mathrm{INTL\_E}}(t,t+1)$: the volume of emigration from municipality *i* to the outside of a nation in the period from time *t* to time *t*+1;

$m_i^{\mathrm{INTL}}(t,t+1)$: the volume of international net migration between municipality *i* and the outside of a nation in the period from time *t* to time *t*+1;

$m_i^{\mathrm{INT}}(t,t+1)$: the volume of internal net migration between municipality *i* and all the other municipalities in the period from time *t* to time *t*+1;

N: set of area numbers of all municipalities constituting a nation.

Obviously from the above definitions,

$$m_i(t,t+1) = m_i^{\mathrm{INTL}}(t,t+1) + m_i^{\mathrm{INT}}(t,t+1) \ , \qquad (1)$$

$$m_i^{\mathrm{INTL}}(t,t+1) = m_i^{\mathrm{INT\_I}}(t,t+1) - m_i^{\mathrm{INTL\_E}}(t,t+1) \ . \qquad (2)$$

The two preconditions that the population of a nation is closed and death rates show a uniform distribution in a nation are expressed by the following two formulas: for a certain cohort and all *i*,

$$m_i^{\mathrm{INT\_I}}(t,t+1) = m_i^{\mathrm{INTL\_E}}(t,t+1) = 0 \ , \qquad (3)$$

$$\frac{d_i(t,t+1)}{p_i(t)} = \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} p_i(t)} \ . \qquad (4)$$

Accordingly, CSR method is formulated as follows: if equations (3) and (4) hold for a certain cohort and all i, then equation (5) is established, that is,

$$m_i(t,t+1) = p_i(t+1) - p_i(t) \cdot \frac{\sum_{i \in N} p_i(t+1)}{\sum_{i \in N} p_i(t)} \ . \tag{5}$$

Obviously from equations (2) and (3),

$$m_i^{\mathrm{INTL}}(t,t+1) = 0 \ . \tag{6}$$

Hence, strictly speaking, equation (5) is rewritten in

$$m_i^{\mathrm{INT}}(t,t+1) = p_i(t+1) - p_i(t) \cdot \frac{\sum_{i \in N} p_i(t+1)}{\sum_{i \in N} p_i(t)} \ . \tag{7}$$

In the above formulation, the left-hand and right-hand sides of equation (4) denote death rates in municipality i and in the whole nation, respectively. Furthermore, $\sum_{i \in N} p_i(t+1) / \sum_{i \in N} p_i(t)$ in equation (7) means, for a certain cohort, a nationwide cohort change ratio as well as a nationwide survival ratio, because it is assumed that there is no migration between the nation and its outside.

Let us derive equation (7) from two preconditions, i.e., equations (3) and (4). First, to prove this, a minor proposition is necessary for proof. Summing up both sides of equation (1) for all i $(i \in N)$,

$$\sum_{i \in N} m_i(t,t+1) = \sum_{i \in N} m_i^{\mathrm{INT}}(t,t+1) + \sum_{i \in N} m_i^{\mathrm{INTL}}(t,t+1) \ . \tag{8}$$

Since regarding internal migration, the number of all origins is quite the same as that of all destinations,

$$\sum_{i \in N} m_i^{\mathrm{INT}}(t,t+1) = 0 \ . \tag{9}$$

Equations (8) and (9) lead to

$$\sum_{i \in N} m_i(t,t+1) = \sum_{i \in N} m_i^{\mathrm{INTL}}(t,t+1) \ . \tag{10}$$

From equation (6), we obtain

$$\sum_{i \in N} m_i(t,t+1) = 0 \ . \tag{11}$$

Second, we can obtain the following formula from the demographic equation for a certain cohort and all i:

$$p_i(t+1) = p_i(t) - d_i(t,t+1) + m_i(t,t+1) \ . \tag{12}$$

By using equation (1), we can transform equation (12) into equation (13), which as mentioned below, has an important role in generalization of CSR method.

$$m_i^{\mathrm{INT}}(t,t+1) = p_i(t+1) - p_i(t)\left\{1 - \frac{d_i(t,t+1)}{p_i(t)} + \frac{m_i^{\mathrm{INTL}}(t,t+1)}{p_i(t)}\right\} \ . \tag{13}$$

Substituting equations (4) and (6) into equation (13),

$$m_i^{\mathrm{INT}}(t,t+1) = p_i(t+1) - p_i(t)\left\{1 - \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} p_i(t)}\right\} \ . \tag{14}$$

Meanwhile, summing up both sides of equation (12) for all i $(i \in N)$,

$$\sum_{i\in N} p_i(t+1) = \sum_{i\in N} p_i(t) - \sum_{i\in N} d_i(t,t+1) + \sum_{i\in N} m_i(t,t+1) \ . \tag{15}$$

After substituting equation (11) into equation (15), and then, dividing both sides of it by $\sum_{i\in N} p_i(t)$, we obtain

$$\frac{\sum_{i\in N} p_i(t+1)}{\sum_{i\in N} p_i\,(t)} = 1 - \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} P_i(t)} \ . \tag{16}$$

Finally, substituting equation (16) into equation (14), we can prove equation (7) to be established.

3. Generalization of the Census Survival Ratio Method

In fact, the two preconditions mentioned above are too limited to lead to equation (7). In other words, the necessary and sufficient condition that equation (7) holds for a certain cohort and all *i* is less limited than the two preconditions expressed by equations (3) and (4). Thus, this chapter attempts to derive the condition from equation (7) and other equations that always hold.

Because both equations (10) and (15) always hold for a certain cohort and all *i*, by substituting equation (10) into (15) and dividing both sides of it by $\sum_{i\in N} p_i(t)$, we get the following important equation that forms one of the cores of discussion on the above necessary and sufficient condition:

$$\frac{\sum_{i\in N} p_i(t+1)}{\sum_{i\in N} p_i\,(t)} = 1 - \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} P_i(t)} + \frac{\sum_{i\in N} m_i^{\mathrm{INTL}}(t,t+1)}{\sum_{i\in N} p_i(t)} \quad . \qquad (17)$$

Substituting equation (17) into (7), we get equation (18).

$$m_i^{\mathrm{INT}}(t,t+1) = p_i(t+1) - p_i(t)\left\{1 - \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} p_i(t)} + \frac{\sum_{i\in N} m_i^{\mathrm{INTL}}(t,t+1)}{\sum_{i\in N} p_i(t)}\right\} \quad . \qquad (18)$$

By subtracting each side of equation (18) from each side of equation (13) that always holds and forms another one of the cores of discussion on the condition, respectively,

$$\frac{d_i(t,t+1)}{p_i(t)} - \frac{m_i^{\mathrm{INTL}}(t,t+1)}{p_i(t)} = \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} P_i(t)} - \frac{\sum_{i\in N} m_i^{\mathrm{INTL}}(t,t+1)}{\sum_{i\in N} p_i(t)} \quad . \qquad (19)$$

This result means that we can deduce equation (19) from (7). On the contrary, if assuming that equation (19) holds for a certain cohort and all *i*, we can derive equation (18) from equations (13) and (19), and then, can obtain equation (7) from equations (17) and (18). As a result, it is proved that equation (19) is the necessary and sufficient condition that equation (7) holds for a certain cohort and all *i*. However, equation (19) is much complicated, and therefore it is not practical to consider equation (19) to be a precondition of CSR method for generalization.

Thus, the author proposes more practical and slightly more limited preconditions that are expressed as below:

$$\frac{m_i^{\mathrm{INTL}}(t,t+1)}{p_i(t)} = \frac{\sum_{i\in N} m_i^{\mathrm{INTL}}(t,t+1)}{\sum_{i\in N} p_i(t)} \quad , \qquad (20)$$

$$\frac{d_i(t,t+1)}{p_i(t)} = \frac{\sum_{i\in N} d_i(t,t+1)}{\sum_{i\in N} P_i(t)} \quad . \qquad (21)$$

If both of the above equations hold for a certain cohort and all *i*, equation (19) is satisfied. Consequently, equations (20) and (21) are appropriate for the new preconditions of CSR method. Let us compare them with the existing preconditions, i.e., equations (3) and (4). Because equation (21) is quite the same as preceding equation (4), it is assumed that death rates show a uniform distribution in a nation. Meanwhile, equation (20) is much less limited than equation (3), and means that international net migration rates have the same value at all municipalities. Since the assumption based on equation (20) does not require the nation for a closed population that is hypothetical, it is much more available than that based on equation (3), and can also be interpreted as follows: long-distance net migration rates have the same value at all municipalities. Generally, the volume of long-distance migration is less than that of short-distance, even if the range of migration is confined to the interior of a nation. This suggests that the proportion of the volume of long-distance net migration to that of internal net migration is small, and that therefore, the assumption based on equation (20) is acceptable not only international net migration but also for internal net migration.

We can generalize CSR method in the basis of the above discussion. A region is considered to be located in a nation, and some new variables are introduced. These variables are defined for a certain cohort as below:

$m_i^{\mathrm{INREG}}(t,t+1)$: the volume of interregional net migration between municipality i and all municipalities outside a region in the period from time t to time t+1;

$m_i^{\mathrm{INTRA}}(t,t+1)$: the volume of intraregional net migration between municipality i and all the other municipalities inside a region in the period from time t to time t+1;

R: set of area numbers of all municipalities constituting a region.

Obviously from the above definitions,

$$m_i^{\mathrm{INT}}(t,t+1) = m_i^{\mathrm{INREG}}(t,t+1) + m_i^{\mathrm{INTRA}}(t,t+1) \ . \qquad (22)$$

Furthermore, we can express a generalized CSR method as follows: if equations (23) and (24) hold for a certain cohort and all i $(i \in R)$, then equation (25) is established, that is,

$$\frac{m_i^{\mathrm{INREG}}(t,t+1)}{p_i(t)} = \frac{\sum_{i\in R} m_i^{\mathrm{INREG}}(t,t+1)}{\sum_{i\in R} p_i(t)} \ , \qquad (23)$$

$$\frac{d_i(t,t+1)}{p_i(t)} = \frac{\sum_{i\in R} d_i(t,t+1)}{\sum_{i\in R} P_i(t)} \ , \qquad (24)$$

$$m_i^{\mathrm{INTRA}}(t,t+1) = p_i(t+1) - p_i(t)\cdot\frac{\sum_{i\in R} p_i(t+1)}{\sum_{i\in R} p_i(t)} \ . \qquad (25)$$

We can decompose internal net migration into long- and short-distance migrations, that is, into interregional and intraregional migrations, using equations (7), (22), and (25), as mentioned in the next chapter.

4. Decomposability of Net Migration

4-1. Theoretical Explanation

As mentioned in the preceding chapter, we can decompose internal net migration into interregional and intraregional migrations, using equations (7), (22), and (25). By substituting equations (7) and (25) into equation (22), for a certain cohort and all i $(i \in R)$, we get

$$m_i^{\mathrm{INREG}}(t, t+1) = m_i^{\mathrm{INT}}(t, t+1) - m_i^{\mathrm{INTRA}}(t, t+1)$$

$$= \left\{ p_i(t+1) - p_i(t) \cdot \frac{\sum_{i \in N} p_i(t+1)}{\sum_{i \in N} p_i(t)} \right\} - \left\{ p_i(t+1) - p_i(t) \cdot \frac{\sum_{i \in R} p_i(t+1)}{\sum_{i \in R} p_i(t)} \right\}$$

$$= P_i(t) \left\{ \frac{\sum_{i \in R} p_i(t+1)}{\sum_{i \in R} p_i(t)} - \frac{\sum_{i \in N} p_i(t+1)}{\sum_{i \in N} p_i(t)} \right\} . \qquad (26)$$

Therefore, internal net migration can be decomposed into two components as follows:

$$m_i^{\mathrm{INT}}(t, t+1) = P_i(t) \left\{ \frac{\sum_{i \in R} p_i(t+1)}{\sum_{i \in R} p_i(t)} - \frac{\sum_{i \in N} p_i(t+1)}{\sum_{i \in N} p_i(t)} \right\} + \left\{ p_i(t+1) - p_i(t) \cdot \frac{\sum_{i \in R} p_i(t+1)}{\sum_{i \in R} p_i(t)} \right\} . \qquad (27)$$

The first and second members in the right hand side of equation (27) denote interregional and intraregional net migrations, respectively. In other words, these two components denote long- and short-distance net migrations, respectively. Let $p_{\mathrm{NAT}}(t)$ be the population of a nation at time *t*, and let $p_{\mathrm{REG}}(t)$ be the population of a region at time *t*, $\sum_{i \in N} p_i(t) = p_{\mathrm{NAT}}(t)$ and $\sum_{i \in R} p_i(t) = p_{\mathrm{REG}}(t)$ are given, and thereby, equation (27) is rewritten in

$$m_i^{\mathrm{INT}}(t, t+1) = P_i(t) \left\{ \frac{p_{\mathrm{REG}}(t+1)}{p_{\mathrm{REG}}(t)} - \frac{p_{\mathrm{NAT}}(t+1)}{p_{\mathrm{NAT}}(t)} \right\} + \left\{ p_i(t+1) - p_i(t) \cdot \frac{p_{\mathrm{REG}}(t+1)}{p_{\mathrm{REG}}(t)} \right\} . \qquad (28)$$

Equation (28) forms the core of the concept "decomposability of net migration." Since all variables in the right hand side of equation (28) are obtained from census data, equation (28) means that we can decompose internal net migration into long- and short-distance components using only census data, i.e., static statistics. This is the largest advantage of generalization that this paper suggests.

Two fractions $p_{\mathrm{REG}}(t+1)/p_{\mathrm{REG}}(t)$ and $p_{\mathrm{NAT}}(t+1)/p_{\mathrm{NAT}}(t)$ in equation (28) show cohort change ratios from time *t* to time *t*+1 for a certain cohort regarding a region and a nation, respectively. However, either of them is not a survival ratio because they do not vary by only death factors. In addition, by dividing both sides of equation (28) by $p_i(t)$, we get the following simple formula:

$$\frac{m_i^{\mathrm{INT}}(t,t+1)}{p_i(t)} = \left\{\frac{p_{\mathrm{REG}}(t+1)}{p_{\mathrm{REG}}(t)} - \frac{p_{\mathrm{NAT}}(t+1)}{p_{\mathrm{NAT}}(t)}\right\} + \left\{\frac{p_i(t+1)}{p_i(t)} - \frac{p_{\mathrm{REG}}(t+1)}{p_{\mathrm{REG}}(t)}\right\} . \qquad (29)$$

In equation (29), the left hand side shows the net migration rate in municipality *i* from time *t* to time *t*+1, and the right hand side shows the sum of two differences in cohort change ratio from time *t* to time *t*+1, i.e., the difference between a region and a nation; and the difference between municipality *i* and a region.

4-2. Application to Actual Data

This section attempts to decompose CoCSIRs developed by the author, by applying the concept decomposability of net migration, that is, equation (28) to them. The CoCSIR is a measure to estimate the long-term influence of social increase, i.e., net migration on an area, and defined for a certain cohort as a ratio of the volume of net migration accumulated after the cohort reaches age 10-14 to the population size when the cohort reaches age 10-14. Because the numerator of the CoCSIR means net migration, we can easily decompose the CoCSIR by decomposing its numerator.

We set Akita City, Akita Prefecture, and Japan for a municipality, a region, and a nation, respectively. Akita Prefecture, which is located in the northeast part of Japan, is one of the most depopulated area in Japan. Akita City is the largest city in Akita Prefecture and its capital as well. Data used for calculation are 1950-2010 census populations and 2015-40 population projections by 5-year age class.

Figures 1, 2, and 3 show results of application, and illustrate the change of Akita City’s CoCSIRs regarding internal, interregional, and intraregional migrations, respectively. According to these figures, the change of CoCSIRs based on internal migration is relatively flat and near to zero for any cohort. In contrast, the change of CoCSIRs based on interregional and intraregional migration is concave or convex at age 20-24. This suggests that Akita City pushes the population to the outside of Akita Prefecture and pulls the population from the inside of Akita Prefecture, and means that internal net migration of Akita City can be clearly decomposed into long- and short-distance migrations that are different in direction from one another. The above discussions enable us to understand that the decomposing method suggested by this paper has a crucial advantage. However, against our expectations, the volume of long-distance net migration is not much less than that of short-distance in the case of Akita City.

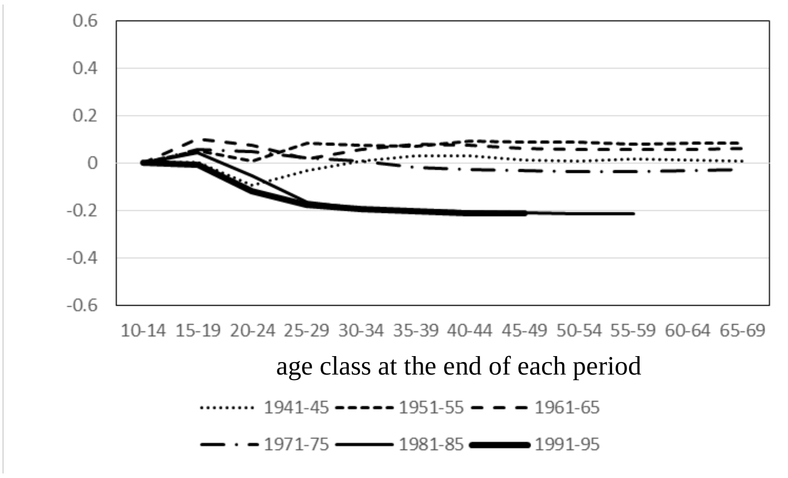

Fig. 1 The Change of Akita City's CoCSIRs Based on Internal Migration

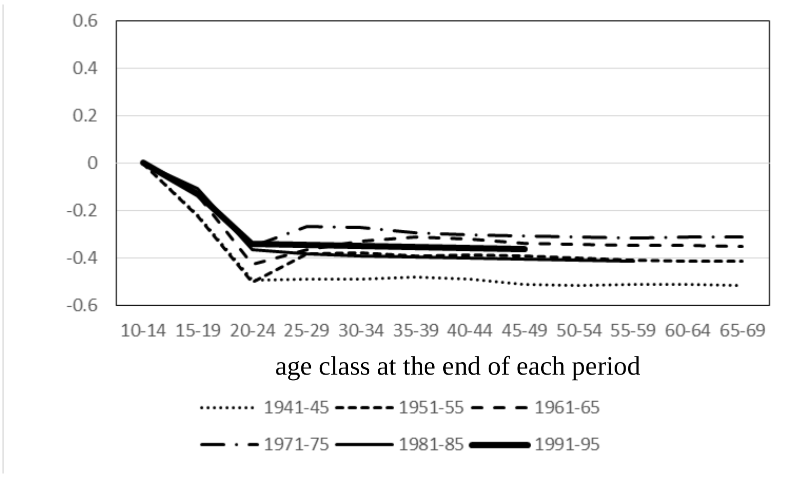

Fig. 2 The Change of Akita City's CoCSIRs Based on Interprefectural Migration

Fig. 3 The Change of Akita City's CoCSIRs Based on Intraprefectural Migration

5. Conclusion

This paper generalizes CSR method by reconsidering its preconditions, establishes the concept of decomposability of net migration, and decomposes the CoCSIR by using the concept. As a result of that, we understand that internal net migration of Akita City can be clearly decomposed into long- and short-distance migrations that are different in direction from one another, and that the decomposing method has a crucial advantage. However, against our expectations, the volume of long-distance net migration is not much less than that of short-distance in the case of Akita City. Some assumptions including this point need to be examined in further studies by using not only census data but also migration statistics.

Acknowledgements

This paper is a result of a study that received financial support from the Japan Society for the Promotion of Science by grant-in-aid (grant number: 16H03525).

References

Inoue, T., 2002, “The Migration Turnaround of Japan from the Demographic Viewpoint.” in Arai, Y., and Kawaguchi, T. and Inoue, T. eds, *Migration in Japan: Life Course and Regionality*, Kokon Shoin, Tokyo.

Inoue, T., 2014, “On the Mathematical Formulation of the Cohort Cumulative Social Increase Ratio.” *Working Paper Series, Institute of Economic Research, Aoyama Gakuin University*, 2014-4, pp.1-18.

Shimizu, M., 2006, “On the Quantum and Temp of Cumulative Net Migration.” *Journal of Population Problems*, Vol.62, No.4, 41-60.

Shimizu, M., 2009, “The Cohort Cumulative Social Increase Ratio by Municipality.” *Jinkogaku Kenkyu* (*The Journal of Population Studies*), No.44, 33-42.