Society for the Study of Economic Inequality

Working Paper Series

Measuring rank mobility with variable population size

Walter Bossert
Conchita D'Ambrosio

ECINEQ WP 2014 - 350

ECINEQ 2014 - 350
December 2014

www.ecineq.org

Measuring rank mobility with variable population size*

Walter Bossert†
University of Montreal, Canada

Conchita D'Ambrosio
INSIDE, University of Luxembourg, Luxembourg

Abstract

We provide characterizations of a class of rank-mobility measures and of a specific member of this class. These measures are based on the Kemeny distance for orderings. We use the well-known replication-invariance property to ensure that our measures are applicable in variable-population settings. The rank-based approach to mobility has a natural connection with the study of social status. Rank-based measures are widely applied in empirical research but their theoretical foundation is still in need of further investigation, and we consider our approach to be a contribution towards this objective.

Keywords: Rank mobility, Kemeny distance, variable population.

JEL Classification: D63.

*Financial support from the Fonds National de la Recherche Luxembourg, the Fonds de Recherche sur la Société et la Culture of Québec and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

†**Contacts details**: (W. Bossert) Department of Economics and CIREQ, University of Montreal, Canada. E-mail: walter.bossert@umontreal.ca; (C. D'Ambrosio) INSIDE, University of Luxembourg, Luxembourg. E-mail: conchita.dambrosio@uni.lu.

1 Introduction

The measurement of mobility is an increasingly important area within the analysis of social index numbers. The fundamental issue to be addressed is the design of measures that reflect the extent to which members of a society move across social or economic boundaries from one period to the next. A crucial aspect that distinguishes mobility from most other criteria that are used to assess the performance of a society (such as income inequality or poverty) is that mobility is difficult—if not impossible—to define without any reference to intertemporal considerations. Of course, intertemporal approaches to the measurement of inequality, poverty and other social phenomena have been explored but they can also be defined without any difficulties in a single-period setting; in contrast, there is no mobility without movement. As a consequence, the arguments of a mobility measure are pairs of indicators of economic or social status—one indicator for each of the time periods under consideration.

Another characteristic of the concept of mobility is that it is multifaceted. As Fields (2008) summarizes, six mobility concepts can be found in the economics literature: time independence, positional movement, share movement, non-directional income movement, directional income movement, and equalizer of longer-term incomes. Excellent surveys and guides to the literature are also provided by Maasoumi (1998), Fields and Ok (1999) and Jäntti and Jenkins (2014).

In this paper we contribute to the measurement of positional movement or, more specifically, to the measurement of the movement across the ranks held by individuals in a society. The rank-based approach to mobility has a natural connection with the study of social status. Rank-based measures are widely applied in empirical research (see, for example, Dickens, 1999) but, to the best of our knowledge, only few contributions such as D'Agostino and Dardanoni (2009) and Cowell and Flachaire (forthcoming) investigate them from a theoretical perspective. We employ the basic setup of these studies but use different methods and arrive at alternative classes of measures.

D'Agostino and Dardanoni (2009) phrase the problem in terms of (partial) permutation matrices and use a subgroup-consistency property to obtain an additive structure of their criteria. Much of their analysis is devoted to dominance criteria in a fixed-population setting, although they discuss variable-population issues as well without providing formal characterizations. In contrast, we explicitly deal with variable-population considerations by imposing replication invariance and characterizing the resulting class of rank-mobility measures. An additive structure results in our setting from a well-established strengthening of the triangle inequality as employed by Kemeny and Snell (1962) and Can and Storcken (2013).

Cowell and Flachaire (forthcoming) propose classes of indices involving various status concepts in a fixed-population setting. Their approach is very flexible and is based on a general measure of distance between individual statuses. The latter may or may not be directly (that is, independently of the position of others) observable. As such, rank mobility is not a central issue in their framework.

In order to perform rank-mobility comparisons across societies with different population sizes, we employ the standard replication-invariance axiom. In our setting, replication invariance demands that if a pair of population rankings is replicated, rank mobility remains unchanged. Thus, unlike much of the existing literature on the subject, our measures are applicable in a variable-population framework.

Two dominant measures of non-parametric rank correlation have been established in the literature—namely, Spearman's (1904) ρ index and Kendall's (1938) τ index. D'Agostino and Dardanoni (2009) characterize rank-mobility preorders that are linked to Spearman's ρ index, whereas we focus on measures based on Kendall's τ index. We note that Kendall's τ index and related measures are not included in D'Agostino and Dardanoni's (2009) class. As is the case for the index proposed by D'Agostino and Dardanoni (2009, p. 1796), our measure assumes values between zero and one. Clearly, this does not apply to the corresponding measures of rank correlation; their values are between minus one and one.

The Kendall τ index is at the core of the Kemeny distance (also referred to as the swap distance), which is one of the most prominent distance measures for orderings; see, for instance, Kemeny (1959) and Kemeny and Snell (1962). By its nature, the rank-mobility setting seems ideally suited for employing the literature on measuring the distance between orderings—in this specific case, the rankings of the individuals in a society before and after a move from one period to the next.

The Kemeny distance is characterized in Kemeny and Snell (1962) but, as pointed out in a remarkable contribution by Can and Storcken (2013), one of the axioms employed in the original characterization is redundant. As a consequence, Can and Storcken (2013) succeed in obtaining a considerable strengthening of the result due to Kemeny and Snell (1962). The axiom in question is a reducibility condition—the only property used by Kemeny and Snell (1962) that (at least implicitly) links distances between orderings involving different numbers of objects to be ranked. Can and Storcken's (2013) observation that the Kemeny distance can be characterized without this axiom represents, in our opinion, a very fundamental and important contribution to this literature.

We make use of the results established by Can and Storcken (2013) to obtain a characterization of a rank-mobility measure that is a variable-population variant of the Kemeny distance. The specific (population-size-dependent) multiplicative factor is determined by the replication-invariance property familiar from the theory of economic index numbers and a normalization axiom. To the best of our knowledge, replication invariance has not appeared in the literature on measuring the distance between orderings. This is likely the case because it is a natural property in the context of measurement issues involving the ranking of individuals but is not of immediate appeal in the more abstract setting of measuring the distance between orderings.

2 The (generalized) Kemeny distance

We begin with a brief review of the Kemeny distance and its recent characterization by Can and Storcken (2013), formulated for the case in which the alternatives under consideration can be identified with positive integers; this is done to facilitate the application of the requisite results to our setting that involves the measurement of mobility. Clearly, this does not involve any loss of generality. Furthermore, although one of the fundamental contributions of the paper by Can and Storcken (2013) is that the Kemeny distance is characterized in a fixed-size setting, we use a variable-population framework so as to directly apply their observations to our rank-mobility framework.

Let $N = \mathbb{N} \setminus \{1\}$. For $n \in N$, we consider a set of alternatives $\{1, \ldots, n\}$. The set of all orderings (that is, all reflexive, complete and transitive binary relations) on $\{1, \ldots, n\}$ is denoted by $\mathcal{R}^n$.

A distance function for orderings is a function $d\colon \cup_{n \in N} \mathcal{R}^n \times \mathcal{R}^n \to \mathbb{R}_+$. An ordering $R^1 \in \mathcal{R}^n$ is between $R^0 \in \mathcal{R}^n$ and $R^2 \in \mathcal{R}^n$ if

$$R^0 \cap R^2 \subseteq R^1 \subseteq R^0 \cup R^2.$$

Let $n \in N$ and let $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a bijective function. For $R \in \mathcal{R}^n$, we define the relation R_π by letting, for all $a, b \in \{1, \ldots, n\}$,

$$(\pi(a), \pi(b)) \in R_\pi \ \Leftrightarrow \ (a, b) \in R.$$

One of the most prominent distance functions is what is usually referred to as the Kemeny distance d_K; see Kemeny (1959), Kemeny and Snell (1962) and Can and Storcken (2013), for instance. It is defined by letting, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,

$$d_K(R^0, R^1) = |R^0 \setminus R^1| + |R^1 \setminus R^0|.$$

The following axioms are familiar from the literature on the measurement of distance between orderings.

Zero at identity only. For all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \ \mathcal{R}^n$,

$$d(R^0, R^1) = 0 \ \Leftrightarrow \ R^0 = R^1.$$

Symmetry. For all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n_+ \times \mathcal{R}^n$,

$$d(R^1, R^0) = d(R^0, R^1).$$

Strong triangle inequality. For all $n \in N$ and for all $R^0, R^1, R^2 \in \mathcal{R}^n$,

$$d(R^0, R^2) \leq d(R^0, R^1) + d(R^1, R^2),$$

and the inequality is satisfied with an equality if and only if R^1 is between R^0 and R^2.

Neutrality. For all $n \in N$, for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$ and for all bijective functions $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$,

$$d(R^0_\pi, R^1_\pi) = d(R^0, R^1).$$

Normalization. For all $n \in N$,

$$\min \{d(R^0, R^1) \mid (R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n \text{ and } R^0 \neq R^1\} = 1.$$

These properties are of immediate appeal not only for general distance functions but also in the specific context of measuring rank mobility.

As a remark aside, note that some authors (including Can and Storcken, 2013) formulate the second part of the strong triangle inequality as a separate axiom but, considering the close connection to the standard triangle inequality, we follow Kemeny (1959) and Kemeny and Snell (1962) in combining the two requirements.

Can and Storcken (2013) prove, among other results, the following two theorems.

Theorem 1. *A distance function d satisfies zero at identity only, symmetry, the strong triangle inequality and neutrality if and only if there exists a set $\{c_n \in \mathbb{R}_{++} \mid n \in N\}$ such that, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,*

$$d(R^0, R^1) = c_n d_K(R^0, R^1).$$

Theorem 2. *A distance function d satisfies zero at identity only, symmetry, the strong triangle inequality, neutrality and normalization if and only if $d = d_K$.*

Theorem 2 represents a substantial improvement of the corresponding result reported in Kemeny (1959) and Kemeny and Snell (1962). These earlier authors employ, in addition to the axioms in the above theorem, a reducibility property that applies across different sizes of the sets of alternatives over which the orderings are defined. Loosely speaking, reducibility requires that the distance between two orderings is unchanged if alternatives are deleted that are ranked identically at the top or at the bottom of the two orderings; see Can and Storcken (2013) for details. Can and Storcken (2013) show that reducibility is redundant in the Kemeny and Snell (1962) characterization of the Kemeny distance and, as a consequence, come up with a characterization that is not only formally but also conceptually much stronger than the original axiomatization. In addition to establishing that the axioms employed in the original characterization are not independent, they manage to get by without the only axiom that (at least implicitly) imposes restrictions on distance measurements across orderings that rank the elements of sets involving different numbers of alternatives. Although we define the notion of a distance function and the axioms in a variable-size setting, it is clear that fixed-size versions are readily available because each property only applies to orderings involving the same number of alternatives to be ranked. In contrast. Kemeny and Snell's (1962) reducibility axiom cannot even be defined in a fixed-size framework and, thus, they did not succeed in providing a characterization on the domain for which Can and Storcken's (2013) results are valid.

3 Measures of rank mobility

We now identify a rank-mobility measure with a distance function and, thus, use the terms 'distance function' and 'rank-mobility measure' interchangeably. Two more properties of such a measure are employed in this section. For $k \in N$, $\mathbf{1}_k$ is the k-dimensional vector in $\mathbb{R}^k$ that consists of k ones. For $n, k \in N$ and $R \in \mathcal{R}^n$, the k-fold replication $R\mathbf{1}_k \in \mathcal{R}^{nk}$ of

R is defined as follows. For each $a \in \{1, \ldots, n\}$, denote the k replicas of a by $a_1, \ldots, a_k$. For all $a, b \in \{1, \ldots, n\}$ and for all $i, j \in \{1, \ldots, k\}$, let

$$(a_i, b_j) \in R\mathbf{1}_k \;\Leftrightarrow\; (a, b) \in R.$$

Replication invariance is an axiom that is familiar from various subfields of the theory of economic index numbers. It requires that if we replicate the orderings in both periods, the value of the rank mobility measure remains unchanged.

Replication invariance. For all $n, k \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,

$$d(R^0\mathbf{1}_k, R^1\mathbf{1}_k) = d(R^0, R^1).$$

Our final property is a normalization analogous to that employed in the usual characterizations of the Kemeny distance. We normalize our index by requiring that its possible values have one as the least upper bound. Clearly, the ranking of pairs of orderings according to our measure is not affected by this normalization; as can be seen from the results reported below, if this normalization property is not imposed, we obtain a one-parameter class of (ordinally equivalent) measures. As is the case for replication invariance, this is a requirement that involves variable-population considerations.

Supremum of one. $\sup \{d(R^0, R^1) \mid (R^0, R^1) \in \cup_{n \in N} \mathcal{R}^n \times \mathcal{R}^n\} = 1.$

The remainder of this section is devoted to our main characterization theorem, along with a preliminary result that is of interest in its own right. First, we add replication invariance to the list of axioms used in Theorem 1 and prove a result that narrows down the class of rank-mobility measures accordingly. Then, we add supremum of one to obtain our main characterization result.

Theorem 3. *A rank-mobility measure d satisfies zero at identity only, symmetry, the strong triangle inequality, neutrality and replication invariance if and only if there exists $c \in \mathbb{R}_{++}$ such that, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,*

$$d(R^0, R^1) = \frac{c}{n^2} d_K(R^0, R^1).$$

Proof. The 'if' part of the theorem statement is straightforward to verify. To prove the 'only-if' part, suppose that the axioms of the theorem statement are satisfied. By Theorem 1, there exists a set $\{c_n \in \mathbb{R}_{++} \mid n \in N\}$ such that, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,

$$d(R^0, R^1) = c_n d_K(R^0, R^1). \tag{1}$$

Let $n, k \in N$ and let $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$ be such that $R^0 \neq R^1$. By definition of the k-fold replication of an ordering, we have, for all $a, b \in \{1, \ldots, n\}$ and for all $i, j \in \{1, \ldots, k\}$,

$$(a_i, b_j) \in R^0\mathbf{1}_k \;\Leftrightarrow\; (a, b) \in R^0$$

and

$$(a_i, b_j) \in R^1\mathbf{1}_k \;\Leftrightarrow\; (a,b) \in R^1.$$

Consider any $(a,b) \in R^0 \setminus R^1$. By definition,

$$\begin{array}{c}(a_1,b_1),\ldots,(a_1,b_k) \in R^0\mathbf{1}_k \setminus R^1\mathbf{1}_k,\\ \vdots \\ (a_k,b_1),\ldots,(a_k,b_k) \in R^0\mathbf{1}_k \setminus R^1\mathbf{1}_k.\end{array}$$

Thus, for each pair in $R^0 \setminus R^1$, there are k^2 pairs in $R^0\mathbf{1}_k \setminus R^1\mathbf{1}_k$ and, analogously, for each pair in $R^1 \setminus R^0$, there are k^2 pairs in $R^1\mathbf{1}_k \setminus R^0\mathbf{1}_k$. Hence,

$$|R^0\mathbf{1}_k \setminus R^1\mathbf{1}_k| + |R^1\mathbf{1}_k \setminus R^0\mathbf{1}_k| = k^2(|R^0 \setminus R^1| + |R^1 \setminus R^0|) = k^2 d_K(R^0,R^1).$$

Using (1), we obtain

$$\begin{aligned} d(R^0\mathbf{1}_k, R^1\mathbf{1}_k) &= c_{nk}(|R^0\mathbf{1}_k \setminus R^1\mathbf{1}_k| + |R^1\mathbf{1}_k \setminus R^0\mathbf{1}_k|)\\ &= c_{nk}k^2 d_K(R^0,R^1)\end{aligned}$$

and

$$d(R^0,R^1) = c_n d_K(R^0,R^1).$$

Therefore, replication invariance requires that, for all $n,k \in N$,

$$c_{nk}k^2 d_K(R^0,R^1) = c_n d_K(R^0,R^1). \tag{2}$$

Because $R^0 \neq R^1$, $d_K(R^0,R^1) > 0$ and, thus, (2) demands that

$$c_{nk}k^2 = c_n \quad \text{for all } n,k \in N. \tag{3}$$

Setting $n = 2$ in (3), we obtain

$$c_{2k} = \frac{c_2}{k^2} \quad \text{for all } k \in N \tag{4}$$

and, letting $m = 2k$ and defining $c = 4c_2 \in \mathbb{R}_{++}$, (4) can be rewritten as

$$c_m = \frac{c_2}{(m/2)^2} = \frac{4c_2}{m^2} = \frac{c}{m^2} \quad \text{for all even } m \in N. \tag{5}$$

Now let $n \in N$ be odd. Let $k = 2$ in (3), which implies that $nk = 2n$. Substituting into (4), it follows that

$$4c_{2n} = c_n \tag{6}$$

and, because $2n$ is even, (5) implies

$$c_{2n} = \frac{c}{4n^2}. \tag{7}$$

Combining (6) and (7), it follows that

$$c_n = \frac{c}{n^2} \quad \text{for all odd } n \in N$$

and, together with (5),

$$c_n = \frac{c}{n^2} \quad \text{for all } n \in N.$$

Substituting back into (1) yields the desired result. ■

Theorem 4. *A rank-mobility measure d satisfies zero at identity only, symmetry, the strong triangle inequality, neutrality, replication invariance and supremum of one if and only if, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,*

$$d(R^0, R^1) = \frac{1}{n^2} d_K(R^0, R^1).$$

Proof. The 'if' part of the theorem statement is straightforward to verify. To prove the 'only-if' part, suppose that the axioms of the theorem statement are satisfied. By Theorem 3, there exists $c \in \mathbb{R}_{++}$ such that, for all $n \in N$ and for all $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$,

$$d(R^0, R^1) = \frac{c}{n^2} d_K(R^0, R^1). \tag{8}$$

Clearly, for any fixed $n \in N$, the maximal value of $d(R^0, R^1)$ as defined in (8) subject to the constraint that $(R^0, R^1) \in \mathcal{R}^n \times \mathcal{R}^n$ is attained for the case in which R^0 is antisymmetric and R^1 is the inverse of R^0. Without loss of generality, suppose that the elements of $\{1, \ldots, n\}$ are labeled so that

$$(1,2), \ldots, (1,n) \in R^0 \setminus R^1,$$
$$\vdots$$
$$(n-1, n) \in R^0 \setminus R^1$$

and

$$(n, n-1), \ldots, (n, 1) \in R^1 \setminus R^0,$$
$$\vdots$$
$$(2,1) \in R^1 \setminus R^0.$$

Thus,

$$|R^0 \setminus R^1| + |R^1 \setminus R^0| = 2 \sum_{i=1}^{n-1} i = n(n-1)$$

and the maximal value of $d(R^0, R^1)$ for fixed $n \in N$ is

$$\frac{c}{n^2} n(n-1) = \frac{c(n-1)}{n}.$$

This expression is increasing in n and, thus,

$$\sup \{d(R^0, R^1) \mid (R^0, R^1) \in \cup_{n \in N} \mathcal{R}^n \times \mathcal{R}^n\} = \lim_{n \to \infty} \frac{c(n-1)}{n} = c.$$

Supremum of one implies that $c = 1$ and the proof of the theorem is complete. ■

4 Concluding remarks

The contribution of this paper consists primarily of an alternative approach to the measurement of rank mobility as initiated by D'Agostino and Dardanoni (2009). While the notion of rank mobility seems to rest on a solid conceptual foundation, it is clear that there are shortcomings as well (as is the case for all areas of economic measurement in which no consensus has been reached yet as far as the existence of a single superior index—or class of indexes—is concerned). We propose a class of replication-invariant mobility indices that generalize one of the two most fundamental measures of non-parametric rank correlation—namely, Kendall's (1938) τ index. The rank-mobility preorders D'Agostino and Dardanoni (2009) characterize are linked to the other common measure of correlation, which is Spearman's (1904) ρ index. Thus, our analysis is by no means intended to diminish the importance of alternative suggestions that have appeared in the literature. When added to D'Agostino's and Dardanoni's (2009) fundamental contribution, the work reported here may be viewed as providing an additional argument in favor of the further exploration of rank-based mobility measurement.

Without going into any technical details, we note that the rank-based setting can be applied to the measurement of income mobility in a straightforward manner. All that is required is the addition of an axiom that ensures that only ranks matter and, furthermore, slight modifications of the zero-at-identity axiom and the second part of the strong triangle inequality so as to take into account this ranks-only property.

References

Can, B. and T. Storcken (2013), A re-characterization of the Kemeny distance, Maastricht University School of Business and Economics, RM/13/009.

Cowell, F.A. and E. Flachaire (forthcoming), Measuring mobility, in: R. Scott and S. Kosslyn (eds.), *Emerging Trends in the Social and Behavioral Sciences*, Sage Publishing, Thousand Oaks.

D'Agostino, M. and V. Dardanoni (2009), The measurement of rank mobility, *Journal of Economic Theory* **144**, 1783–1803.

Dickens, R. (1999), Caught in a trap? Wage mobility in Great Britain: 1975-1994, *Economica* **67**, 477–497.

Fields, G.S. (2008), Income mobility, in: L. Blume and S. Durlauf (eds.), *The new Palgrave Dictionary of Economics*, Palgrave Macmillan, New York.

Fields, G.S. and E.A. Ok (1999), The measurement of income mobility: an introduction to the literature, in: J. Silber (ed.), *Handbook of Income Inequality Measurement,* Kluwer Academic Publishers, Norwell, pp. 557–596.

Jäntti, M. and S.R. Jenkins (2014), Income mobility, in: A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution, Vol. 2,* Elsevier, Amsterdam.

Kemeny, J.G. (1959), Mathematics without numbers, *Daedalus* **88**, 577–591.

Kemeny, J.G. and J.L. Snell (1962), Preference rankings: an axiomatic approach, in: J.G. Kemeny and J.L. Snell (eds.), *Mathematical Models in the Social Sciences,* Blaisdell Publishing Company, New York, pp. 9–23.

Kendall, M.G. (1938), A new measure of rank correlation, *Biometrika* **30**, 81–93.

Maasoumi, E. (1998), On mobility, in: D. Giles and A. Ullah (eds.), *Handbook of Applied Economic Statistics,* Marcel Dekker, New York, pp. 119–175.

Spearman, C. (1904), The proof and measurement of association between two things, *American Journal of Psychology* **15**, 72–101.