The Measurement of Rank Mobility

Marcello D’Agostino
Dipartimento di Scienze Umane
Università di Ferrara
dgm@unife.it

Valentino Dardanoni
Dipartimento SEAF
Università di Palermo
vdardano@unipa.it

KEYWORDS: Rank mobility, Mobility measurement, Concordance, Partial matrices, Spearman’s index.

JEL numbers: D31, D63.

This is the post-print authors’ version of the paper: M. D’Agostino, V. Dardanoni. The measurement of rank mobility. Journal of Economic Theory 144(2009):1783–1803.

*We would like to thank an anonymous referee for extremely useful suggestions which have significantly contributed to improve the paper. Address for correspondence: Valentino Dardanoni, Dipartimento SEAF, Università di Palermo, Viale delle Scienze, 90143 Palermo, Italy, tel. 39-091-6626221, fax 39-091-489346.
Abstract

In this paper we investigate the problem of measuring social mobility when the social status of individuals is given by their rank. In order to sensibly represent the rank mobility of subgroups within a given society, we address the problem in terms of partial permutation matrices which include standard (“global”) matrices as a special case. We first provide a characterization of a partial ordering on partial matrices which, in the standard case of global matrices, coincides with the well-known “concordance” ordering. We then provide a characterization of an index of rank mobility based on partial matrices and show that, in the standard case of comparing global matrices, it is equivalent to Spearman’s $\rho$ index.
1 Introduction

When discussing social mobility issues, a basic distinction is usually made between *intergenerational mobility* (how the distribution of some relevant measure of individual status changes between different generations in a given society) and *intragenerational mobility* (how the distribution of individual status changes among a group of individuals over a given period of their lifetime). As a vehicle of discussion, we shall concentrate on intergenerational mobility, but all our considerations and results could be easily transposed to the intragenerational case. All the information about a social mobility context is then contained in a bivariate cumulative distribution function, which describes the distribution of two random variables capturing fathers' and sons' socio-economic status. It is widely believed that socioeconomic mobility is somewhat an elusive concept, difficult to define, let alone to measure: as remarked by Fields and Ok [11] in a recent survey “...the mobility literature does not provide a unified discourse of analysis. ...a considerable rate of confusion confronts a newcomer in the field.”¹ This may be contrasted with the literature on income inequality, where a consensus has emerged on what concepts of inequality mean, on the correct theoretical procedures to measure it, and on how to go from theory to empirical applications.

One of the main challenges in mobility measurement is the precise definition of individual socio-economic status, and its practical evaluation using available data. Typically, mobility studies use data such as current or permanent income, consumption, occupational prestige, education etc. For ease of presentation, assume that income data is used for this purpose. In this paper we explore the possibility that the socio-economic status of each individual is given by his *rank*, i.e. by its relative position in the generation to which he belongs. This way of defining individual status seems quite natural and intuitively satisfying. Indeed, the Encyclopedia Britannica defines social status as “the relative rank that an individual holds” and some analysts follow this common-sense interpretation by *equating* the concept of social status with that of rank.

It is generally agreed that the intuitive notion of intergenerational mobility tries to capture an idea of “dissimilarity” between two vectors, namely those representing, respectively, the social status of fathers and sons. If one holds the view that the status of an individual is well captured by his rank, then rank mobility is a natural concept worth of investigation. The rank-based approach has a further advantage: disregarding all information on fathers’ and sons’ incomes, besides that reflected in their rank, the researcher is allowed to focus on comparisons which do not depend on the marginal distributions, since these are both reduced to a common uniform one. To take a simple artificial example, consider three “societies” A, B and C, each consisting of five

¹See also Maasoumi [21] for a survey on mobility measurement.
families, such that the incomes of fathers and sons are associated as in Table 1. A quick
inspection shows that these three societies display an equal amount of rank mobility, since the coupling of fathers’ and sons’ ranks is exactly the same in all of them. In fact, in any bivariate joint distribution of fathers and sons’ incomes, marginal distributions contain information of a static nature, while mobility is a dynamic concept which captures how the two distributions are coupled together. When comparing bivariate distributions which differ widely in their marginals (such as A, B and C in Table 1), it may be quite difficult to capture the dependence of the incomes of sons upon those of fathers, which is commonly deemed as a normatively unappealing feature for a society.
Rank mobility is one way of capturing the extent of such origin dependence and, in this sense, comparing rank mobility can be seen as conceptually similar to the common practice of using quantile matrices for comparing Markov mobility chains.²

Table 1: Three societies displaying the same rank mobility.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fathers</td>
<td>100 150 200 250 300</td>
<td>220 500 700 820 1000</td>
<td>190 380 610 800 960</td>
</tr>
<tr>
<td>Sons</td>
<td>150 200 100 250 350</td>
<td>520 640 400 800 1220</td>
<td>280 320 190 420 600</td>
</tr>
</tbody>
</table>

²However, since using income ranks and disregarding all other information certainly involves a good deal of simplification, the notion of rank mobility is not without its critics; for example, one may legitimately claim that the information about income differences between adjacent ranks should be relevant to the appraisal of social mobility.
³This case typically arises when \((x_1, y_1), \ldots, (x_n, y_n)\) is actually a random sample generated from a continuous distribution (so that ties occur with probability zero).
sons’ income are axiomatized, among others, by [4, 5, 10, 12, 16, 20, 22, 24, 26] and
capture different aspects of social mobility than the present paper. These contributions
may be considered more complementary than alternative to our approach. Fields ([9],
chapter 6) compares some theoretical properties of various indices of income mobility
including some indices of rank mobility such as rank correlation, and Buchinsky et al.
[3] compare their empirical properties in an application to French income mobility.4

Given our assumption of no ties in the marginal distributions, all the information
concerning the rank mobility of a society is contained in a permutation matrix $P$,
with typical element $P(i, j) = 1$ if there is a family in this society whose father
has rank $i$ and son has rank $j$, and 0 otherwise. The problem is how to turn this
information into a quantitative measure.5 In order to achieve a faithful and consistent
representation of the rank mobility of subgroups within a given society, we address the
problem in terms of partial permutation matrices (defined in Section 2) which include
standard (“global”) matrices as a special case, and argue that a representation of the
rank mobility of a given subgroup of the population in terms of global matrices would be
paradoxical. After observing that a standard decomposability property — which is key
in the characterization of additively separable indices — cannot be sensibly assumed
in the context of measuring rank mobility, we take advantage of our representation
in terms of partial matrices to define a weaker form of decomposability which can
be safely assumed. As an intermediate step, in section 3 we provide (Theorem 1) a
characterization of a partial ordering on partial matrices which in the case of global
matrices coincides with the well-known “concordance” ordering.

We then provide (in Theorems 2 and 3) a rather natural and simple characteriza-
tion (up to a monotonic transformation) of an index of rank mobility based on partial
matrices and show that, in the standard case of comparing the mobility of two global
matrices of equal size, this is equivalent to Spearman’s $\rho$ index. We also discuss (in
Proposition 2 and 1) another interesting incomplete preorder which lies between con-
cordance and the class of its completions (over partial matrices of equal size) singled
out by Theorem 2. We finally show, in the last section, how the characterized index can
be naturally extended to one which provides a complete preorder of partial matrices of
different size. Our results seem to provide reasonable grounds for adopting this kind of
index (rather than other alternative ones, such as Kendall’s $\tau$ or Spearman’s footrule)
in the measurement of rank mobility.

4The only characterization of a mobility index which takes explicit account of ranks that we are
aware of is [20]. King’s index uses information not only on the ranks of fathers’ and sons’ incomes,
but also on their numerical values, and thus his approach in not directly comparable with ours, but
can be considered complementary.

5Note that if $(x_1, y_1), \ldots, (x_n, y_n)$ is a random sample and we multiply by $\frac{1}{n}$ the permutation
matrix $P$ we obtain the so called empirical joint rank distribution function (see Block et al. [2]).
2 Subgroup mobility and partial permutation matrices

Let $F$ denote the set of all families who live in a given society and consider a subset $A \subseteq F$; examples of interesting subsets are the families which live in a given geographical location, or which belong to a given ethnic group, or whose fathers have a given education level etc. Sometimes we may be interested in exploring how the status of individuals change from one generation to the next for members of this particular subset. We could call this kind of information the rank mobility of $A$ with respect to $F$. Observe that this is not the same as considering the rank mobility of $A$ w.r.t. $A$, because individuals’ rank is calculated with respect to the whole of $F$. A simple example may help to clarify. Consider a society $F$ consisting of six families in which the distributions of fathers’ and sons’ incomes is summarized in Table 2 below:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fathers</td>
<td>100</td>
<td>150</td>
<td>200</td>
<td>250</td>
<td>300</td>
<td>350</td>
</tr>
<tr>
<td>Sons</td>
<td>150</td>
<td>200</td>
<td>100</td>
<td>250</td>
<td>350</td>
<td>300</td>
</tr>
</tbody>
</table>

Table 2: Incomes of fathers and sons in $F$.

Now, consider the subset $A$ of $F$ consisting only of the third, fourth and fifth families. If we consider only $A$ and calculate the rank of individuals with respect to this specific subset, then there is no rank mobility from one generation to the next. Viceversa, there is clearly a sense in which the families in $A$ exhibit some status mobility, which is made apparent when the status is calculated with respect to the whole of $F$: the son of the third family has lost two positions with respect to his generation, while the son of the fifth family has gained one position.

This kind of “partial” mobility information, i.e. restricted to a subset of a whole set $F$ of families, will then be described by an $n \times n$ matrix which differs from a permutation matrix because it can have rows and columns with zeros only. Such matrices are called partial permutation matrices.\footnote{See e.g. Horn and Johnson [17] for definitions and some properties of partial permutation matrices.} When necessary for clarity, we shall call ordinary permutation matrices global. More formally, the set $\mathcal{P}_n$ of $n \times n$ partial permutation matrices is defined as follows: a matrix $P$ belongs to $\mathcal{P}_n$ if and only if, for all $i = 1, \ldots, n$ and $j = 1, \ldots, n$, we have: (i) $P(i, j) \in \{0, 1\}$; (ii) $\sum_i P(i, j) \leq 1$; (iii) $\sum_j P(i, j) \leq 1$. Notice that, under this definition, global matrices are nothing but a special case of partial matrices.

Now, suppose $A$ and $B$ are disjoint subsets of a set $F$ of $n$ families. Clearly the partial permutation matrices that describe the rank mobility of $A$ and $B$ with respect
to $F$, call them $P$ and $Q$, will belong to $\mathcal{P}_n$ and will be disjoint in a related sense that is expressed by the following definition:

**Definition 1.** $P, Q \in \mathcal{P}_n$ are disjoint if

$$P(i, j) = 1 \text{ and } Q(m, k) = 1 \Rightarrow i \neq m \text{ and } j \neq k.$$  

**Corollary 1.** $P, Q \in \mathcal{P}_n$ are disjoint if and only if $P + Q \in \mathcal{P}_n$, where $+$ is the usual sum of matrices.

Therefore, the rank mobility with respect to $F$ of disjoint subsets of $F$ is represented by disjoint partial matrices. Notice that if we partition $F$ into $m$ mutually exclusive and exhaustive subsets $A_1, \ldots, A_m$, the rank mobility of these subsets with respect to $F$ will be described by mutually disjoint partial matrices $P_1, \ldots, P_m$ such that $P = P_1 + \cdots + P_m$. Let us say that a partial matrix $P$ is atomic if there exists exactly one $i$ and one $j$ such that $P(i, j) = 1$, that is, if we are considering a subset containing exactly one family. We shall use the lower case letter $p$ (possibly with subscripts) to denote atomic matrices. Clearly, any partial permutation matrix in $\mathcal{P}_n$ will be equal to a sum $p_1 + \cdots + p_k$ of $k \leq n$ atomic matrices, where $k = n$ only for global matrices.

Observe that any $n \times n$ partial matrix $P$ can be regarded as representing the rank mobility of some set of families $A$ with respect to some “society” $F$ of size $n$ that includes it. Indeed, it is always possible to find an $F$ such that the rank mobility determined by the marginal distributions of fathers’ and sons’ income is represented by a global matrix that includes $P$. So, the above corollary implies that the sum of two disjoint partial matrices can always be regarded as representing the rank mobility of some suitable subgroup $A$ of a possible “society” $F$. Therefore, in this abstract setting, we can forget about “real” families, groups and societies and concentrate only on the partial matrices that represent their rank mobility. However, to avoid long-winded sentences, we shall often abuse of the more concrete terminology and speak, for instance, of “a society (group) $P$” to mean “a society (group) whose rank mobility is represented by the (partial) matrix $P$”, or of “the family $(i, j)$ (in a matrix $P$)” to mean “the family in which the father’s rank is $i$ and the son’s rank is $j$ (in a society whose rank mobility is represented by $P$)”.

We are interested in investigating the properties of some suitable ordering $\preceq$ over the class $\mathcal{P} = \bigcup_{i=1}^{\infty} \mathcal{P}_i$ such that $P \preceq Q$ can be taken as meaning that the matrix $Q$ exhibits at least the same degree of social mobility as the matrix $P$.

## 3 Axiomatizing rank mobility orderings

In this section we shall start investigating the ordering relation $\preceq$, by restricting our attention to comparisons between matrices of equal size. In other words, the charac-
terized orderings will be such that \( P \preceq Q \) holds true only if both \( P \) and \( Q \) belong to the same subclass \( \mathcal{P}_n \) for some fixed \( n \). We shall sometimes use the notation \( \preceq_n \) for the restriction of \( \preceq \) to \( \mathcal{P}_n \), but will omit the subscript \( n \) whenever it is clear from the context that we are comparing matrices of the same size. Then, in Section 4 we shall discuss how \( \preceq \) can be extended in a natural way into a rank mobility index which applies to all partial matrices (of arbitrary size), so inducing a complete preorder over the whole class \( \mathcal{P} \).

Given two matrices \( P \) and \( Q \) in \( \mathcal{P}_n \) (for some given \( n \)), when can we say that \( Q \) displays at least the same rank mobility as \( P \)? We first introduce and discuss some plausible axioms to impose on the ordering \( \preceq \) and then derive characterization theorems following an incremental approach. As a first step, in Section 3.1, we shall only assume that \( \preceq \) is a preorder, that is a reflexive and transitive binary relation. Then, we shall derive, in Theorem 1, a characterization of what we propose as the basic rank mobility ordering from two axioms. Such ordering allows only for comparisons between partial matrices of equal size which satisfy the further condition of being “similar” (in the sense explained in Definition 2) and is not complete on any of the classes \( \mathcal{P}_n \). Next, in Section 3.2, we shall investigate possible extensions of this basic preorder. Assuming that, for all \( n \), \( \preceq_n \) is a complete preorder of \( \mathcal{P}_n \), we shall introduce further axioms which will allow us to obtain sharper characterizations in Theorems 2 and 3.

### 3.1 The concordance ordering

While it is intuitively clear that it is meaningful to compare two standard (i.e. global) permutation matrices representing the rank mobility of two societies \( F \) and \( F' \), it is not quite as clear whether it is equally meaningful to compare partial matrices — representing, say, the rank mobility of some subset \( A \in F \) w.r.t. \( F \) and the rank mobility of some subset \( B \in F' \) w.r.t. \( F' \) — when they have different marginal distributions.\(^7\) We shall therefore start by restricting the comparison to a clear-cut case.

**Definition 2.** We say that:

1. Two matrices \( P, Q \in \mathcal{P}_n \) are similar if
   
   \[
   \{ i | P(i,j) = 1 \text{ for some } j \} = \{ i | Q(i,j) = 1 \text{ for some } j \}
   \]

   \[
   \{ j | P(i,j) = 1 \text{ for some } i \} = \{ j | Q(i,j) = 1 \text{ for some } i \}.
   \]

2. A matrix \( P \in \mathcal{P}_n \) is monotone if, for all \( i, j, m, k \) such that \( P(i,j) = 1 \) and \( P(m,k) = 1 \), we have \( (i - m)(j - k) > 0 \).

\(^7\)Loosely speaking, by “marginal distributions” of a partial matrix \( P \) we mean the following: the marginal distribution of the fathers is the set of all \( i \) such that \( P(i,j) = 1 \) for some \( j \), and the marginal distribution of the sons is the set of all \( j \) such that \( P(i,j) = 1 \) for some \( i \).
So, two matrices are similar when they have equal marginal distributions. Note that the definition of similarity induces an equivalence relation on $\mathcal{P}_n$. Moreover, observe that within each similarity set there is a unique monotone matrix which can be considered as displaying the least amount of mobility:

**Axiom 1** (Monotonicity). For all $n$, and for all distinct $P, Q \in \mathcal{P}_n$ such that $P$ is monotone and similar to $Q$ we have $P \prec Q$.

Notice also that a matrix is monotone if, and only if, there is a strictly increasing function from fathers’ rank to sons’ rank.$^8$

The second axiom requires that the sum of disjoint partial matrices is a monotonic operation:

**Axiom 2** (Subgroup Consistency). For all $n$ and for every $P_1, P_2, P_3, P_4 \in \mathcal{P}_n$ such that $P_1$ is disjoint with $P_2$ and $P_3$ is disjoint with $P_4$

$$P_1 \preceq P_3 \text{ and } P_2 \preceq P_4 \Rightarrow P_1 + P_2 \preceq P_3 + P_4.$$ 

Similar axioms are commonly used in the literature on income inequality [28], poverty [15] and mobility measurement [12], where they usually imply a fundamental, and practically useful, decomposability property: given an arbitrary partition of a population into $k$ subgroups, the problem of measuring a certain feature in the overall population can be reduced to the $k$ separate problems of measuring that feature in each of the $k$ subgroups.$^9$ It must be stressed that the above axiom cannot be interpreted as asserting a similar decomposability of the rank mobility of a society $F$ into the rank mobility of their subgroups. In the terminology used in the introduction, this would amount to asserting that, given a partition of $F$ into $A_1, \ldots, A_k$, the rank mobility of $F$ w.r.t. $F$, can be decomposed into the rank mobility of $A_1$ w.r.t. $A_1$, $A_2$ w.r.t. $A_2$, etc., where the rank of each individual is evaluated with reference to the subgroup to which it belongs. However, this would clearly be paradoxical in the context of measuring rank mobility.

To see why, recall the simple example given in Section 2 (see Table 2), considering a society made of six families. Let $A_1$ be the subgroup consisting of the third, fourth and

---

$^8$The reader may find it helpful to compare our concept of similar matrices with the well-known Frechet class of distributions with fixed marginals, and our monotonicity axiom with the lower bound in the Frechet class (see e.g. Nelsen [25]).

$^9$Though such axioms are widely accepted in these contexts, for a critical discussion see Foster and Sen [14].

$^{10}$For this interpretation of the decomposability property in the context of social mobility see, for instance, Fields and Ok [10]. The term decomposability has different interpretations in other theoretical and applied contexts.
fifth families, and $A_2$ the subgroup consisting of the first, second and sixth. Now, it is clear that the rank mobility of the whole $F$ w.r.t. $F$ is greater than zero, while the rank mobility of $A_1$ w.r.t. $A_1$ and the rank mobility of $A_2$ w.r.t. $A_2$ are both, intuitively, equal to zero. Hence, we cannot hope that rank mobility enjoys such a strong decomposability property.\textsuperscript{11} However, our axiom states a \textit{weaker} decomposability property: given a partition of $F$ into $A_1, \ldots, A_n$, the rank mobility of $F$ w.r.t. $F$, can be decomposed into the rank mobility of $A_1$ w.r.t. $F$, $A_2$ w.r.t. $F$, etc., where the rank of each individual is evaluated with reference to the whole society $F$. Accordingly, the axiom is expressed in terms of \textit{partial} permutation matrices for a society $F$, which are obtained from the global permutation matrix for $F$ by omitting information concerning some of the families, and are intended to represent the rank mobility of given subgroups with respect to the whole of $F$. Our subgroup consistency axiom, therefore, cannot be interpreted as allowing us to measure the rank mobility of a population in terms of \textit{independent} measurements of the rank mobility of its subgroups. Indeed, our axiom is better understood as a monotonicity requirement on the sum of disjoint partial permutation matrices, and, from this point of view, it expresses a basic logical property that seems quite uncontroversial.

Suppose, now, that we have a matrix $P$ such that for $k < m$ and $l < n$ we have $P(k, l) = P(m, n) = 1$, and consider another matrix $Q$ such that $P(i, j) = P(i, j)$ for all $i \neq k, m$ and $j \neq l, n$, and $Q(k, n) = Q(m, l) = 1$. In words, $Q$ differs from $P$ because there has been an inversion of social status between two families, such that before the inversion the higher status father had the higher status son, while after the inversion the lower status father has the higher status son. Intuitively, such an inversion should be mobility-increasing.\textsuperscript{12} Under these circumstances we say that $Q$ has been obtained from $P$ by \textit{inverting} $(k, l)$ and $(m, n)$. We write $P \prec Q$ whenever $Q$ can be obtained from $P$ by means of such an inversion.

Suppose, a given matrix $Q$ can be obtained from $P$ by a \textit{sequence} of inversions. We can intuitively conclude that $Q$ displays more social mobility than $P$, and write $P \preceq_C Q$. Formally:

\textbf{Definition 3.} $P \preceq_C Q$ if and only if there is a finite sequence of matrices $P_0, \ldots, P_k$, with $k \geq 0$, such that (i) $P_0 = P$, (ii) $P_k = Q$ and (iii) if $k > 0$, $P_{i-1} \prec P_i$ for all $i = 1, \ldots, k$.

It can be easily checked that $\preceq_C$ is a partial ordering\textsuperscript{13} defined on each set of similar

\textsuperscript{11}That a mechanical application of standard decomposability properties is clearly nonsensical in this context is remarked, for example, by Cowell \cite{Cowell5} page 144, who explicitly states that distance measures based on ranks are not decomposable.

\textsuperscript{12}Such swaps are well-known in the mathematical statistics \cite{29} and economics \cite{8, 1, 6} literature, where it is often assumed that they are always mobility-increasing.

\textsuperscript{13}That is, besides being reflexive and transitive, it is also \textit{antysimmetric}, namely: $P \preceq_C Q$ and
matrices. The reason for the choice of the superscript "C" is that, when the similarity class consists of the global matrices in \( P_n \), \( \preceq^C \) is called the concordance ordering in the mathematical statistics literature, see e.g. Tchen [29] and Kimeldorf and Sampson [19].

**Theorem 1.** Within each set of similar matrices, \( \preceq^C \) is the smallest\(^{14}\) preorder which satisfies Axiom 1 and Axiom 2.

A proof of this theorem is given in Appendix A. The concordance ordering \( \preceq^C \) is a very well established and much studied ordering of bivariate distributions. It is a partial ordering which, in the space of global permutation matrices, is a subrelation of many important complete orders, for example, those induced by the popular nonparametric indices of concordance such as Kendall’s \( \tau \) and Spearman’s \( \rho \), see e.g. Schweizer and Wolff [27]. The theorem then says that all reflexive and transitive relations \( \preceq \) which satisfy Axioms 1 and 2 must have a common area of agreement equal to \( \preceq^C \). Atkinson [1] first applies the concordance ordering to mobility measurement; Dardanoni [6] applies it to a Markov chain model of social mobility, and shows the equivalence of a version of this ordering to some very intuitive concepts of greater social mobility. In particular, appropriately defining father’s and sons’ status as monotonic functions of their rank, Theorem 4 of Dardanoni [6] can be adapted to this context to show the equivalence of \( \preceq^C \) with useful partial orderings for making rank mobility comparisons in ways that parallel the classical Lorenz ordering.

On the other hand, \( \preceq^C \) allows for comparisons between similar matrices only and, while this restriction is immaterial when comparing global matrices, it makes the comparison of partial matrices impossible except for the artificial special case in which the matrices have exactly the same marginal distributions. Moreover, being a partial ordering, \( \preceq^C \) does not even allow for comparisons of all matrices in a given similarity set. Thus, in order to be able to compare all mobility contexts in \( P_n \) (for any \( n \)), we must focus on preorders \( \preceq \) such that every restriction \( \preceq_n \) is a complete preorder of \( P_n \).\(^{15}\) Clearly, even assuming that each \( \preceq_n \) is complete, Axioms 1 and 2 are not sufficient to provide a unique characterization (since these axioms are satisfied by several distinct preorders which are complete on every \( P_n \), e.g. the above mentioned \( \rho \) and \( \tau \).) From this point of view, Theorem 1 only implies that every preorder which is complete on each \( P_n \) and satisfies the axioms must include \( \preceq^C \), so that the properties expressed by the axioms can be considered as minimal requirements on any suitable mobility ordering. Thus, in the sequel, we shall take our mobility ordering \( \preceq \) to be a complete preorder.

\[^{14}\text{In terms of set-inclusion.}\]

\[^{15}\text{Recall that } \preceq_n \text{ is a complete preorder of } P_n \text{ when for all } P, Q \in P_n, \text{ either } P \preceq_n Q \text{ or } Q \preceq_n P.\]
preorder of each \( P_n \) and seek for extra axiomatic properties that allow us to uniquely characterize it.

### 3.2 Completing the concordance ordering

In this section we investigate the possible completions over each \( P_n \) of the basic concordance ordering characterized in Theorem 1. We shall therefore assume that each restriction \( \preceq_n \) of our mobility ordering \( \preceq \) is a complete preorder of \( P_n \) and consider the class of such preorders satisfying Axioms 1 and 2. (As implied by Theorem 1, they must all include the concordance ordering.) Our aim, now, is to investigate how these axioms can be expanded to single out a suitable mobility ordering from this class.

There are two distinct intuitive aspects of the notion of “greater mobility” which emerge from its conceptual analysis. One aspect, which is apparent in the standard definitions of some well-known orderings — such as the concordance ordering and the complete preorder based on Kendall’s function \( \tau \) — stems from the idea that there is an increase in mobility when two families interchange their relative position. On the other hand, from a different angle, mobility is related to the distance between father’s and sons’s status within each family, and overall mobility of a group of families may be construed as the aggregation of the degrees of mobility exhibited by all the families in that group.\(^{16}\)

Now, for a single family in a society \( F \), such that father’s rank is \( i \) and son’s rank is \( j \), we can take \( |i − j| \) as measuring the social distance between father’s and son’s social status. This basic intuition is captured by the following:

**Axiom 3 (Atomic Monotonicity).** For all \( n \) and for any two atomic matrices \( p, q \in P_n \) such that \( p(i, j) = q(i', j') = 1 \),

\[
p \preceq q \iff |i − j| \leq |i' − j'|.
\]

Notice that, although this axiom forces a unique complete preorder of atomic matrices, it is not sufficient to uniquely characterize \( \preceq_n \) in the whole domain of \( P_n \). It should also be noted\(^ {17} \) that this axiom forces one to care about rank alterations only as far as the extent of these alterations is concerned, disregarding where these alterations occur. For example, this axiom forces one to see an equal amount of mobility in an atomic permutation matrix in which father’s rank is the worst and son’s rank is one above the worst, and in an atomic permutation matrix in which father’s rank is the second best

\(^{16}\)Clearly these two concepts of mobility (one which considers the interplay of families and the other which considers families in isolation) are interrelated, since single families cannot change relative positions without affecting other families.

\(^{17}\)We thank a referee for pointing this out to us.
and son’s rank is the best. One may find this objectionable from a normative point of view, for lower income dynamics are often deemed more important from this angle. On the other hand, it can be observed that such concerns appear to be more significant when considering the difference between father’s and son’s actual income, rather than between their rank. Clearly, the same amount of income movement may appear to be more significant as the level of income decreases, but it is not obvious that the same should hold true for movement of ranks, since in this case all the information concerning the actual income levels is discarded. It may well be that the actual movement of income observed within a low rank family is much smaller than the one observed within a high rank family. Since any purely ordinal approach simply discards this kind of information, it is not obvious how to assign higher weight to lower income dynamics.

Let’s now introduce some notation which will simplify considerably the following discussion. If “$P$” denotes a matrix in a given space $\mathcal{P}_n$, then “$P^m$”, with $m \geq n$, will denote the matrix in $\mathcal{P}_m$ which coincides with $P$ wherever $P$ is defined and contains only 0’s everywhere else, i.e. the matrix defined as follows:

$$P^m(i, j) = P(i, j) \text{ for all } i, j \leq n \text{ and }$$

$$P^m(i, j) = 0 \text{ otherwise.}$$

On the other hand, if “$P$” denotes a matrix in $\mathcal{P}_m$, we shall attach no meaning to the notation “$P^k$” with $k < m$. Observe that, by definition,

1. $(P^k)^m = P^m$ for every $m \geq k$

2. $(P + Q)^m = P^m + Q^m$ for every disjoint $P, Q \in \mathcal{P}_k$ with $k \leq m$.

Let us say that a matrix $P$ is null if $i = j$ for all $(i, j)$ such that $P(i, j) = 1$. Intuitively, a null-matrix says that the subgroup for which it is defined displays no mobility at all. The following axiom is an adaptation of the well-known Archimedean Property to our setting:

**Axiom 4 (Archimedean Property).** For every $m$ and all $P, Q \in \mathcal{P}_m$, the strict inequality $P \prec Q$ holds if and only if there is an $n \geq m$ and a non-null $R \in \mathcal{P}_n$, disjoint with $P^n$, such that

$$P^n + R \sim Q^n.$$  

The meaning of this axiom is that there is no gap between the mobility displayed by two partial matrices (groups of families) which cannot be bridged if one takes into account the possibility that the societies in which these are embedded are expanded. That is, if the total number of families is increased and the group displaying less mobility is expanded with new families which display positive mobility, it is always possible to equalize the mobility of the other group.
It must be stressed that such an axiom is acceptable only if one is concerned with the absolute amount of mobility displayed by the partial matrices, disregarding any consideration of the size of the groups whose mobility they represent.\footnote{On the other hand, a “relative” notion of rank mobility, which takes into account the size of the groups, can be soundly based on the absolute notion outlined in this section. On this point see Section 4 below.} To intuitively grasp this axiom one can usefully compare it with a similar property which holds in the field of real numbers, where \( p < q \) holds true if and only if there is a positive \( r \) such that \( p + r = q \). In fact, this axiom could be taken as a definition of the relation \( P \prec Q \), once it has been made clear that this relation concerns absolute rank mobility comparisons.

Observe now that, within a given class \( \mathcal{P}_n \), a partial permutation matrix \( P \) is uniquely determined by the set \( S(P) = \{(i, j) | P(i, j) = 1\} \). We call \( S(P) \) the characteristic set of \( P \). We can prove the following:

**Theorem 2.** Let \( \preceq \) be a preorder over \( \mathcal{P} \) such that \( \preceq \) is complete over \( \mathcal{P}_n \) for every \( n \in \mathbb{N} \). Then \( \preceq \) satisfies Axioms 1–4 if and only if there is a strictly increasing and strictly convex function \( f : \mathbb{N} \to \mathbb{N} \) such that, for all \( n \) and for all \( P, Q \in \mathcal{P}_n \),

\[
P \preceq Q \iff \sum_{(i,j) \in S(P)} f(|i - j|) \leq \sum_{(i,j) \in S(Q)} f(|i - j|).
\]

A proof of this theorem is given in Appendix B. Theorem 2 shows that Axioms 1–4 characterize (up to a monotonic transformation) a class of additive mobility indices which depend on the choice of an appropriate weighting function \( f \). It is interesting to notice that, within the space of global matrices, two important indices of ordinal association which would seem appropriate to (im)mobility measurement, namely Kendall’s \( \tau \) and Spearman’s footrule (see e.g. Kendall and Gibbons \cite{18} for definitions and a discussion of their properties) do not belong to the class defined in Theorem 2.

Consider for example the global permutation matrices \( P, P' \) and \( P'' \) in \( \mathcal{P}_4 \) with the following characteristic sets:

\[
S(P) = \{(1, 1), (2, 4), (3, 3), (4, 2)\},
S(P') = \{(1, 3), (2, 1), (3, 4), (4, 2)\},
S(P'') = \{(1, 1), (2, 3), (3, 4), (4, 2)\}.
\]

Using any of the mobility indices, say \( M \), in the class characterized by Theorem 2, the mobility of \( P, P' \) and \( P'' \) will be equal to

\[
M(P) = H(f(0) + f(2) + f(0) + f(2)),
M(P') = H(f(2) + f(1) + f(1) + f(2)),
M(P'') = H(f(0) + f(1) + f(1) + f(2)).
\]
for some strictly increasing and strictly convex $f$ and strictly increasing $H$. Now, if we adopted Spearman’s footrule as a mobility measure, which corresponds to letting $f$ be the identity function, $P$ and $P''$ would display the same amount of mobility, since $0 + 2 + 0 + 2 = 0 + 1 + 1 + 2$. However, $P$ can be derived from $P''$ by an inversion of the families $(2,3)$ and $(3,4)$. Thus $P'' \prec C P$, and so Spearman’s footrule is inconsistent with $\preceq_C$. This failure of Spearman’s footrule to satisfy the basic ordering $\preceq_C$ makes it unsuitable for measuring rank mobility.

On the other hand, it is easy to show that, in the class of global matrices, Kendall’s $\tau$ does indeed agree with $\preceq_C$ (see e.g. Schweizer and Wolff [27]). Nevertheless, it cannot satisfy all our axioms, as can be seen by observing that $P$ and $P'$ have the same value of Kendall’s $\tau$, while any of the indices of Theorem 2 would deliver different values, since in $P'$ there are two families with social distance equal to 2 (as in $P$), but, in addition, there are also two families with positive social distance (since $f$ is strictly increasing).

Now, Theorem 2 characterizes a class of complete preorders of each $P_n$ which, while being small enough to exclude some important mobility indices, is still too wide and its practical application is dependent on the choice of an appropriate function $f$. At this stage a natural option, suggested by a referee, is to look for a preorder which seeks agreement in equivalence (1) for all strictly increasing and strictly convex functions $f : \mathbb{N} \to \mathbb{N}$. Let $d_P$ denote the vector of rank differences $|i - j|$’s for the families in $S(P)$, with dimension equal to the cardinality of $S(P)$, and let $\tilde{d}_P$ denote the decreasing rearrangement of $d_P$. Standard results in the stochastic dominance literature imply the following:

**Proposition 1.** For all $n$ and for all $P, Q \in P_n$ such that $|S(P)| = |S(Q)| = k \leq n$, the following conditions are equivalent:

1. $\sum_{(i,j) \in S(P)} f(|i - j|) \leq \sum_{(i,j) \in S(Q)} f(|i - j|)$ for all strictly increasing and strictly convex function $f : \mathbb{N} \to \mathbb{N}$;

2. $\sum_{i=1}^{j} q_i^P \leq \sum_{i=1}^{j} q_i^Q$, $j = 1, \ldots, k$.

Proposition 1 gives an empirically verifiable condition which is equivalent to the agreement in equivalence (refsumoff) for all strictly increasing and strictly convex functions $f$. Let us $\preceq_M$ be the preorder defined by Condition 2 of Proposition 1 above,

---

19 The axiomatic characterization of a preorder which seeks agreement in equivalence (1) for all strictly increasing and strictly convex functions $f$ is an interesting topic for future research. Some inspiration may come from the work of Dubra, Macleroni and Ok [7], who characterize an expected utility representation for a potentially incomplete preference relation over lotteries by means of a set of von Neumann-Morgenstern utility functions.

which is defined on any \( P, Q \in \mathcal{P}_n \) (for some fixed \( n \)) with the same number of non-zero entries. The following proposition shows that, within each class of similar matrices, \( \preceq^M \) is in fact an extension of \( \preceq^C \):

**Proposition 2.** For all \( n \) and for all similar \( P, Q \in \mathcal{P}_n \), \( P \preceq^C Q \) implies \( P \preceq^M Q \) but the converse does not apply.

A proof that, within each class of similar matrices, \( \preceq^C \) implies \( \preceq^M \) is in Appendix C. That the converse is not true can be established by the following example: let \( P \) be the global matrix in \( \mathcal{P}_3 \) with characteristic set \( S(P) = \{(1,2), (2,1), (3,3)\} \), and let \( Q \) be the global matrix in \( \mathcal{P}_3 \) with characteristic set \( S(Q) = \{(1,1), (2,3), (3,2)\} \). Then \( d^P = (1,1,0) \) and \( d^Q = (0,1,1) \), so that \( \tilde{d}^P = \tilde{d}^Q \). Therefore, Condition 2 of Proposition 1 applies and \( P \sim^M Q \), but it can be easily checked that \( P \) and \( Q \) are not comparable by \( \preceq^C \).

While the preorder \( \preceq^M \) can be easily applied and allows the researcher to order a greater number of partial matrices than the concordance ordering, it is not a complete preorder of each class \( \mathcal{P}_n \). To obtain a completion for every \( \mathcal{P}_n \), the most natural route consists in imposing further conditions which allow us to pinpoint one element in the class of preorders characterized in Theorem 2, which are complete for every \( \mathcal{P}_n \) (that is, to appropriately choose a given strictly increasing and strictly convex \( f \)). We now show that the choice of a specific function \( f \) can be guided by the following argument: consider an inversion of two families \( (i,j) \) and \( (i+1,j+1) \), so that 1 is the distance between the fathers and 1 is the distance between the sons before the inversion. We shall write \( P \prec Q \) when \( Q \) is obtained from \( P \) by means of such an inversion.

It may be reasonable to assume that all the inversions of this type are minimal inversions and that they all generate an equivalent mobility increase.\(^{21}\) This assumption is expressed by the following axiom:

**Axiom 5** (Minimal Inversion). For all \( n \) and for all matrices \( P, Q, R \in \mathcal{P}_n \),

\[
P \prec Q \text{ and } P \prec R \implies Q \sim R.
\]

Then we can prove the following:

**Theorem 3.** Let \( \preceq \) be a preorder over \( \mathcal{P} \) such that \( \preceq_n \) is complete over \( \mathcal{P}_n \) for every \( n \in \mathbb{N} \). Then \( \preceq \) satisfies Axioms 1–5 if and only if for all \( n \) and for all \( P, Q \in \mathcal{P}_n \),

\[
P \preceq Q \iff \sum_{(i,j) \in S(P)} (i - j)^2 \leq \sum_{(i',j') \in S(Q)} (i' - j')^2.
\]

\(^{21}\)By contrast, we might assume that the increase of mobility generated by an inversion should depend, somehow, also on the values of the father’s rank \( i \) and the son’s rank \( j \) in the inverted families. In this case, of course, not all minimal inversions would be equivalent.
A proof is given in Appendix D. It can be easily verified that, within the set of global matrices, Theorem 3 characterizes (up to a monotonic transformation) the well-known Spearman index of ordinal association, since the latter (which is better described as an immobility index) can be written as

\[ \rho(P) = 1 - \frac{6 \sum_{(i,j) \in S(P)} (i - j)^2}{n^3 - n} \]

(see e.g. Kendall and Gibbons [18], page 8).

On the other hand, the ordering characterized in Theorem 3 is not restricted to populations’ comparisons. For partial permutation matrices, the theorem provides a means for comparing the status mobility of different subgroups when the concept of social status we are interested in refers to the rank of individuals in the whole society. As an example, recall again the society \( F \) considered in in Section 2 (see Table 2), and assume that the third, fourth and fifth family belong to a first group, while the first, second and sixth belong to a second group. It is then easily calculated that families in the first group exhibit a greater level of rank mobility than those in the second; since, applying Theorem 3, we have \( 4 + 0 + 1 > 1 + 1 + 1 \).

4 A rank mobility index

Theorem 3 provides a characterization of a preorder \( \preceq \), which is complete over each class \( \mathcal{P}_n \) consisting of all partial matrices of a given fixed size \( n \). However, in many situations it would be useful to go beyond the preorder \( \preceq \), since:

- when comparing two matrices, say \( P \) and \( Q \), we may want to conclude not only that \( P \) displays more rank mobility than \( Q \), but also to be able to claim in a meaningful way that, say, \( P \) displays 20% more mobility than \( Q \);
- the characterized ordering allows us to compare matrices of the same size in a meaningful way, but it is of no use when it comes to comparing matrices of different size. Thus, it would be desirable to be able to compare partial matrices over the set \( \mathcal{P} = \bigcup_{i=1}^{\infty} \mathcal{P}_n \) of partial matrices of any size;
- \( \preceq \) allows us to compare all partial matrices of the same size \( n \), and so to compare the overall rank mobility of arbitrary subgroups, \textit{regardless of the number of families \((i, j)\) composing the subgroup}. Although this may be intuitively sound when one is interested in measuring the absolute amount of rank mobility exhibited by a subgroup, alternatively one may be interested in a relative notion of mobility which takes into account the number of families which compose the subgroups which are being compared.
Let then $I_n : \mathcal{P}_n \mapsto \mathbb{R}$ be a rank mobility index implicitly characterized by Theorem 3:

$$I_n(P) = F_n\left( \sum_{(i,j) \in S(P)} (i-j)^2 \right)$$

where $F_n$ is some strictly increasing function. A first reasonable property to impose on $I_n(P)$ is that $F_n$ is in fact a similarity transformation, that is,

$$I_n(P) = c_n\left( \sum_{(i,j) \in S(P)} (i-j)^2 \right)$$

for some positive constant $c_n$. If we now let $I : \mathcal{P} \mapsto \mathbb{R}$ denote a rank mobility index which allows comparisons of all matrices in $\mathcal{P}$, it is then natural to assume that, whenever $P \in \mathcal{P}_n$, $I(P) = I_n(P)$.

A further property that one may wish to impose on such a rank mobility index $I$ is that it achieves a fixed maximum, which can be conventionally chosen to be 1, and this maximum is reached by any “antimonotone” partial matrix. In particular, let $A_{k,n} \in \mathcal{P}_n$ denote the partial matrix such that (i) the cardinality of $S(A_{k,n})$ is $k$ (that is, $A_{k,n}$ contains $k$ families), and (ii) families’ ranks are as follows: $(1,n)$, $(2,n-1)$, ..., $(k,n-k+1)$. It can then be argued that, among all partial matrices $P \in \mathcal{P}_n$ with $S(P)$ having cardinality $k$, $A_{k,n}$ displays the maximum amount of mobility, that is, for any $k$ and $n$, $I(A_{k,n}) = 1$. It then immediately follows that the rank mobility index we seek has the following form:

$$I(P) = \frac{1}{(k^3 - k)/3 + k(n-k)^2} \sum_{(i,j) \in S(P)} (i-j)^2$$

for any partial matrix $P \in \mathcal{P}_n$ with cardinality of $S(P)$ equal to $k$. Notice that to derive the index above we have used the fact that, for any $A_{k,n}$, $\sum_{(i,j) \in S(A_{k,n})}(i-j)^2 = (k^3 - k)/3 + k(n-k)^2$.

Two properties of our rank mobility index $I$ may be worth noticing. First, the index actually measures the percentage of rank mobility (measured according to the formula characterized in Theorem 3) of any given partial matrix $P$ with respect to the theoretically maximum achievable rank mobility for a subgroup with the same number of families in a society of the same size. Secondly, when $P$ is a global matrix (that is, $k = n$), it is immediately checked that $I$ coincides with Spearman’s index $\rho$, after allowing for the fact that Spearman’s index range is $[-1, 1]$ while $I(P) \in [0, 1]$ and, as remarked above, $\rho$ measures immobility rather than mobility.
Proof. We first show that the $\preceq^C$ ordering (which, we recall, is defined on each set of similar partial matrices) satisfies the axioms. It is obvious that it satisfies Axiom 1. As for Axiom 2, suppose $P_1, P_2$ and $P_3, P_4$ are mutually disjoint, and $P_1 \preceq^C P_3$ and $P_2 \preceq^C P_4$. Then there exists a sequence $Q_0, \ldots, Q_k$ of partial matrices in $\mathcal{P}_n$ such that (i) $Q_0 = P_1$, (ii) $Q_k = P_3$ and, whenever $k > 0$, (iii) $Q_i \preceq Q_{i+1}$ for $i = 0, \ldots, k - 1$. Similarly, there exists a sequence $R_0, \ldots, R_{k'}$ of partial matrices in $\mathcal{P}_n$ such that (i) $R_0 = P_2$, (ii) $R_{k'} = P_4$ and, whenever $k' > 0$, (iii) $R_i \preceq R_{i+1}$ for $i = 0, \ldots, k' - 1$. Suppose $k' > k$. Then, it is easy to see that, since $P_1, P_2$, and $P_3, P_4$ are mutually disjoint, the sequence

$$Q_0 + R_0, \ldots, Q_k + R_k, Q_k + R_{k+1}, \ldots, Q_k + R_{k'}$$

is such that (i) $Q_0 + R_0 = P_1 + P_2$, (ii) $Q_k + R_{k'} = P_3 + P_4$, (iii) $Q_i \preceq Q_{i+1}$ and $R_i \preceq R_{i+1}$ for $i = 0, \ldots, k - 1$, and (iv) $Q_k + R_j \preceq Q_k + R_{j+1}$ for $j = k, \ldots, k' - 1$. Hence, by definition of $\preceq^C$, we have that $P_1 + P_2 \preceq^C P_3 + P_4$. The argument is similar when $k > k'$.

Next, we show that if a preorder $\preceq$ satisfies the axioms, then it must include the concordance ordering. This is sufficient to conclude that $\preceq^C$ is the smallest preorder satisfying the axioms.

Suppose $\preceq$ satisfies the axioms. Let $P$ and $Q$ be two matrices such that $P \preceq^C Q$. By definition, this means that $P$ and $Q$ are similar and there is a sequence of matrices $P_0, \ldots, P_k$, with $k \geq 0$, such that $P_0 = P$, $P_k = Q$ and, if $k > 0$, $P_i \preceq P_{i+1}$ for all $i = 1, \ldots, k - 1$. Now consider the $i$-th inversion step, and suppose it is such that, for some $j, m, l, n$ with $j < m$ and $l < n$, $P_{i-1}(j, l) = P_{i-1}(m, n) = 1$ and $P_i(j, n) = P_i(m, l) = 1$. Consider the matrix $P^*_{i-1}$ such that only $P^*_{i-1}(j, l) = P^*_{i-1}(m, n) = 1$, while all the other entries are 0 (that is, its characteristic set $S(P^*_{i-1})$ is equal to $\{(j, l), (m, n)\}$). Let also $P^*_{i}$ be the similar matrix such that only $P^*_{i}(j, n) = P^*_{i}(m, l) = 1$, while all the other entries are 0 (that is, its characteristic set $S(P^*_{i})$ is equal to $\{(j, n), (m, l)\}$).

Clearly

$$P_{i-1} = (P_{i-1} - P^*_{i-1}) + P^*_{i-1} \quad \text{and} \quad P_i = (P_i - P^*_{i}) + P^*_{i}.$$ 

Moreover, $P_{i-1} - P^*_{i-1} = P_i - P^*_{i}$ and, since $j < m$ and $l < n$, the matrix $P^*_{i-1}$ is a monotone matrix, so that $P^*_{i-1} \preceq P^*_i$ (by Axiom 1). Therefore,

$$P_{i-1} = (P_{i-1} - P^*_{i-1}) + P^*_{i-1} \preceq (P_i - P^*_{i}) + P^*_i = P_i$$

by Axiom 2. Hence, $P_{i-1} \preceq P_i$. The same argument holds for all $i$, therefore $P \preceq Q$. This shows that $P \preceq^C Q$ implies $P \preceq Q$ for all $P, Q$, i.e. $\preceq^C$ is included in $\preceq$. Since
$\preceq$ was an arbitrary preorder satisfying the axioms, $\preceq^C$ is included in all the preorders satisfying the axioms.

\section*{B Proof of Theorem 2}

\subsection*{B.1 Preliminary lemmas}

The following lemmas turn out to be useful to make the proof of Theorem 2 more readable.

\begin{lemma}
Let $\preceq$ be a preorder over $P$ such that $\preceq_n$ is complete over $P_n$ for every $n \in \mathbb{N}$. If $\preceq$ satisfies Axiom 4, then

$$ P \preceq Q \iff P^m \preceq Q^m, $$

for all $n$, all $P, Q \in P_n$ and all $m \geq n$.

\end{lemma}

\begin{proof}
Suppose $P \preceq Q$ and $P^m \not\preceq Q^m$, that is $P^m \succ Q^m$, for some $m \geq n$. Then, by Axiom 4, there is an $m' \geq m$ and a non-null $R \in P_{m'}$ such that $R$ is disjoint with $Q^{m'}$ and

$$ Q^{m'} + R \sim P^{m'}.$$

Since $m' \geq n$, it follows again from Axiom 4, that $Q \prec P$, against the hypothesis. Hence:

$$ P \preceq Q \implies P^m \preceq Q^m. \quad (2) $$

For the converse, suppose that for some $m > n$, $P^m \preceq Q^m$ and $P \succ Q$. Then, by Axiom 4,

$$ P^k \sim Q^k + R, \quad (3) $$

for some $k \geq n$ and some non-null $R \in P_k$. Now, if $k \geq m$, this implies, again by Axiom 4, that $Q^m \prec P^m$ against the hypothesis. If $k < m$, it follows from (3), by (2), that $P^m \sim (Q^k + R)^m \sim Q^m + R^m$. Since $R^m$ is non-null, it follows, again by Axiom 4, that $Q^m \prec P^m$, against the hypothesis. \hfill \Box

\end{proof}

\begin{lemma}
Let $\preceq$ be a preorder over $P$ such that $\preceq_n$ is complete over $P_n$ for every $n \in \mathbb{N}$. If $\preceq$ satisfies Axiom 4, then

$$ P \preceq P + R, $$

for all $n$, and all $P, R \in P_n$ such that $P + R \in P_n$.

\end{lemma}
Proof. Suppose \( P \triangleright P + R \). Then, by Axiom 4,
\[
P^m \sim (P + R)^m + S \sim P^m + R^m + S
\]
for some \( m \geq n \) and some non-null \( S \in \mathcal{P}_m \). Since \( R^m + S \) is non-null, this would imply, again by Axiom 4, that \( P \triangleright P \), which is impossible. \( \square \)

Let us denote by \( \perp_n \) the unique matrix \( P \in \mathcal{P}_n \), that we call the \textit{empty matrix}, such that \( P(i, j) = 0 \) for all \( i, j \), that is the \( n \times n \) matrix which is everywhere undefined. By definition, (i) \( \perp_n \) is a null matrix (ii) for every \( P \in \mathcal{P}_n \), \( \perp_n \) is disjoint with \( P \), and (iii) \( \perp_n + P = P \).

\textbf{Lemma 3}. Let \( \preceq \) be a preorder over \( \mathcal{P} \) such that \( \preceq_n \) is complete over \( \mathcal{P}_n \) for every \( n \in \mathbb{N} \). If \( \preceq \) satisfies Axioms 2, 3 and 4, then
\[
\perp_n \sim P \prec Q
\]
for every null matrix \( P \in \mathcal{P}_n \) and every non-null matrix \( Q \) in \( \mathcal{P}_n \).

\textit{Proof}. First, recall that (by Axiom 4) \( \perp_n \prec Q \), for every \( n \) and every non-null \( Q \in \mathcal{P}_n \), since \( \perp_n + Q = Q \sim Q \). Hence, we only have to show that \( P \sim \perp_n \) for every null \( P \in \mathcal{P}_n \). If \( P = \perp_n \), then it is trivially true that \( P \sim \perp_n \). Consider, then, the case that \( P \neq \perp_n \). Let us first show that \( p \sim \perp_n \) for every null atomic matrix \( p \) in \( \mathcal{P}_n \). Suppose, that \( p \triangleright \perp_n \). Then, by Axiom 4, \( p^m \sim \perp_m + R = R \) for some \( m \geq n \) and some non-null \( R \in \mathcal{P}_m \). Since \( R \) is non-null, \( R = r + T \) for some non-null atomic matrix \( r \) and some, possibly null, matrix \( T \) in \( \mathcal{P}_m \). By Axiom 3, \( p^m \prec r \) and, by Lemma 2, \( p^m \prec r + T = R \). This is a contradiction, since we had before concluded that \( p^m \sim R \). Suppose, then, that \( \perp_n \triangleright p \). By Axiom 4, \( p^m + R \sim \perp_m \) for some \( m \geq n \) and some non-null \( R \in \mathcal{P}_m \). Now, since \( \perp_m \prec Q \) for every non-null \( Q \in \mathcal{P}_m \), it follows that \( \perp_m \prec R \). Then, by Lemma 2, \( \perp_m \prec p^m + R \) against the previous conclusion that \( \perp_m \sim p^m + R \). Hence, since \( \preceq \) is a complete preorder of \( \mathcal{P}_m \), \( \perp_m \sim p \).

If \( P \) is not an atomic matrix and \( P \neq \perp_n \), then \( P = p_1 + \cdots + p_k \) for some \( k \) such that \( 1 < k \leq n \), with each \( p_i \) \((1 \leq i \leq k)\) being a null atomic matrix. As we have just established, \( p_i \sim \perp_n \) for all \( i = 1, \ldots, k \). Hence, by Axiom 2 (and recalling that \( \perp_n + \perp_n = \perp_n \)), \( P \sim \perp_n \). \( \square \)

Say that two matrices \( P \) and \( Q \) are \textit{atomically equivalent} if, for every \( k \geq 0 \), they contain the same number of non-zero entries \((i, j)\) with \(|i - j| = k\). Clearly, if \( P \) and \( Q \) are atomically equivalent, there are atomic matrices \( p_1, \ldots, p_m \) and \( q_1, \ldots, q_m \), with \( m \leq n \), such that \( P = p_1 + \cdots + p_m, Q = q_1 + \cdots + q_m \) and, by Axiom 3, \( p_i \sim q_i \) for \( i = 1, \ldots, m \). Hence, by Axiom 2, if \( P \) and \( Q \) are atomically equivalent, then \( P \sim Q \).
Remark 1. Given any two matrices $P, Q \in \mathcal{P}_n$, one can always find a sufficiently large $m$ and a matrix $R$ in $\mathcal{P}_m$ such that $R$ is atomically equivalent to $Q^m$ and disjoint with $P^m$. For this purpose, it is sufficient to take $m = 2n$ and $R$ equal to the matrix such that (i) $R(i,j) = 0$ for all $i, j \leq n$ and (ii) $R(n+i, n+j) = 1$ if and only if $Q(i,j) = 1$. Using this method, if $P_1, \ldots, P_k$ are matrices in $\mathcal{P}_n$, one can always find suitable matrices $P'_1, \ldots, P'_k \in \mathcal{P}_{kn}$ such that (i) $P'^{kn}_i \sim P'_i$ for $i \leq k$ and (ii) all the $P'_i$ are mutually disjoint.

Lemma 4. Let $\preceq$ be a preorder over $\mathcal{P}$ such that $\preceq_n$ is complete over $\mathcal{P}_n$ for every $n \in \mathbb{N}$. If $\preceq$ satisfies Axioms 2 and 4, then for all $n$, and all $P, Q, R, S \in \mathcal{P}_n$ such that $P$ is disjoint with $R$ and $Q$ is disjoint with $S$,

$$P \sim Q \text{ and } P + R \sim Q + S \implies R \sim S.$$  

Proof. Let us assume that $P \sim Q$ and $P + R \sim Q + S$. Suppose, ex absurdo, that $R \not\sim S$.

Case 1: $R > S$. Then, it follows from Axiom 4, that $R^m \sim S^m + T$ for some $m \geq n$ and some non-null $T \in \mathcal{P}_m$. By Remark 1, there are $m' \geq m$ and $U \in \mathcal{P}_{m'}$ such that $U \sim P^{m'}$ and $U$ is disjoint with $S^m + T$. Hence, by Lemma 1 and Axiom 2,

$$P^{m'} + R^{m'} \sim U + S^{m'} + T \sim Q^{m'} + S^{m'}.$$  

So, by Axiom 4, $U + S^{m'} \prec Q^{m'} + S^{m'}$ (since $T$ is non-null). However, by Axiom 2, $U + S^{m'} \sim Q^{m'} + S^{m'}$ (since $U \sim P^{m'} \sim Q^{m'}$ by hypothesis and Lemma 1), which is a contradiction.

Case 2: $R < S$. This case is similar to Case 1 and is left to the reader. \qed

Now, consider the set $\mathcal{P} = \bigcup_{i=1}^\infty \mathcal{P}_i$ of all partial matrices. We define the subset $\Delta_k$, $k \geq 0$, of $\mathcal{P}$ as the set of all $P \in \mathcal{P}$ such that for all $(i, j) \in S(P)$, $|i - j| \leq k$. Notice that $\Delta_k \subseteq \Delta_n$ whenever $k \leq m$. The matrices in $\Delta_0$ are the null matrices. We shall also write $\Delta^*_k$ for $\Delta_k \cap \mathcal{P}_n$. Moreover, given two matrices $P, Q \in \mathcal{P}_n$, let us say that $Q$ is contained in $P$ if $P(i, j) = 1$ for all $i, j \in \{1, \ldots, n\}$ such that $Q(i, j) = 1$. Recall that every partial matrix can be uniquely expressed as a sum of atomic matrices.

Lemma 5. Let $\preceq$ be a preorder over $\mathcal{P}$ such that $\preceq_n$ is complete over $\mathcal{P}_n$ for every $n \in \mathbb{N}$. If $\preceq$ satisfies Axioms 2, 3 and 4, then for all $n$ and all $P, Q \in \mathcal{P}_n$, $P \preceq Q$ if and only if the number of non-null atomic matrices contained in $P$ is less than or equal to the number of non-null atomic matrices contained in $Q$.

Proof. Suppose first that the number of non-null atomic matrices contained in $P$ is less than or equal to the number of non-null atomic matrices contained in $Q$. Let $p_1, \ldots, p_j$
be the non-null atomic matrices in $P$ and $q_1, \ldots, q_k$ the non-null atomic matrices in $Q$, with $j \leq k \leq n$. Then $P = p_1 + \cdots + p_j + R$ for some null $R \in P_n$ and $Q = q_1 + \cdots + q_j + S$ for some possibly non-null $S \in P_n$. By Axiom 3 all the non-null atomic matrices in $\Delta_k^n$ are equivalent to each other and therefore, by Axiom 2, $p_1 + \cdots + p_j \sim q_1 + \cdots + q_j$. Moreover, by Lemma 3, $R \preceq S$. Hence, again by Axiom 2, $P \preceq Q$. Suppose now that the number of non-null atomic matrices in $P$ is strictly greater than the number of non-null atomic matrices in $Q$. Let $p_1, \ldots, p_k$ the non-null atomic matrices in $P$ and $q_1, \ldots, q_j$ the non-null atomic matrices in $Q$, with $j \leq k \leq n$. So, $P = p_1 + \cdots + p_j + R$ for some non-null $R \in P_n$ and $Q = q_1 + \cdots + q_j + S$ for some non-null $S \in P_n$. Moreover, by Lemma 3, $S \prec R$ and, as argued above, $p_1 + \cdots + p_j \sim q_1 + \cdots + q_j$. So, by Axiom 2, $Q = q_1 + \cdots + q_j + S \preceq p_1 + \cdots + p_j + R = P$. By Lemma 4, $Q \sim P$ would imply that $S \sim R$ which, given that $S$ is null and $R$ is non-null, is ruled out by Lemma 3. Therefore, we can conclude that $Q \prec P$. \hfill $\Box$

**Lemma 6.** Let $\preceq$ be a preorder over $P$ such that $\preceq_n$ is complete over $P_n$ for every $n \in \mathbb{N}$. If $\preceq$ satisfies Axioms 2–4, then for every $n$ and every atomic matrix $p \in P_n$, there is an $m \geq n$ such that $p^m \sim Q$ for some $Q \in \Delta_k^m$.

**Proof.** Since every atomic matrix $p \in P_n$ belongs to some $\Delta_k^n$, with $k \in \mathbb{N}$, we prove the lemma by induction on the index $k$ of the smallest class $\Delta_k^n$ to which $p$ belongs. In the course of the proof, and for the sake of clarity, we shall reserve the notation $\bar{P}, \bar{Q}$, etc. to refer to matrices in $\Delta_1$.

**Base:** $p \in \Delta_1^n$. Trivial.

**Step:** $p \in \Delta_{j+1}^n$ ($j \geq 1$). Assuming that the lemma holds for all atomic matrices in $\Delta_j^n$ we show that it holds also for all atomic matrices in $\Delta_{j+1}^n$.

Suppose $p$ is an atomic matrix which belongs to $\Delta_{j+1}^n$ but does not belong to $\Delta_j^n$. Then, $p$ is non-null and, by Axiom 3, all atomic matrices in $\Delta_j^n$ are strictly less than $p$. Let now $q$ be a non-null atomic matrix in $\Delta_j^n$. Then, $q \prec p$ and, by Axiom 4, $q^m + R \sim p^m$ for some $m \geq n$ and some non-null $R \in P_m$. Since, $q^m$ is itself non-null, this implies (again by Axiom 4) that $R \preceq p^m$. Now, we argue that $R$ must be in $\Delta_j^m$. We reason by absurd. Suppose $R \not\in \Delta_j^m$, then $R = r + S$ for some atomic $r \in P_m$ not in $\Delta_j^m$, and some (possibly empty) matrix $S \in P_m$. However (by Axiom 3) $r \succeq p^m$ and (by Lemma 2) $r + S \succeq p^m$. Hence, $R \succeq p^m$ against the conclusion, reached before, that $R \prec p^m$. Thus, $R$ must be in $\Delta_j^m$ and so also $q^m + R$ is in $\Delta_j^m$. By inductive hypothesis, there is an $m' \geq m$ such that $q^{m'} + R^{m'} \sim \bar{P}$ for some $\bar{P} \in \Delta_{j'}^m$. So, since $p^m \sim q^m + R$, by Lemma 1, $p^{m'} \sim \bar{P}$. This concludes the proof of the lemma. \hfill $\Box$
B.2 Proof of the main theorem

Proof. We leave it to the reader to prove that the ordering satisfies Axioms 1–4.

To show that any ordering \( \preceq \) which satisfies the axioms must be of the required form, let \( P, Q \) be two matrices in \( \mathcal{P}_n \). First, recall that \( P \) and \( Q \) can be rewritten as sums of atomic matrices,

\[
P = \sum_{(i,j) \in S(P)} p_{(i,j)} \quad \text{and} \quad Q = \sum_{(i,j) \in S(Q)} p_{(i,j)},
\]

where \( p_{(i,j)} \) is the atomic matrix in \( \mathcal{P}_n \) such that \( S(p_{(i,j)}) = \{(i,j)\} \).

By Lemma 6, for each atomic \( p_{(i,j)} \in \mathcal{P}_n \), there is an \( k \geq n \) such that \( p^k_{(i,j)} \sim \hat{P}_{(i,j)} \) for some \( \hat{P}_{(i,j)} \) in \( \Delta^k_1 \). Observe that, by Remark 1, one can always find, for each \( (i,j) \in S(P) \cup S(Q) \), a suitable matrix \( \hat{P}'_{(i,j)} \) in \( \Delta^K_1 \), for some sufficiently large \( K \), such that (i) \( \hat{P}'_{(i,j)} \) is atomically equivalent to \( \hat{P}^K_{(i,j)} \) (and therefore, by Lemma 1, also to \( p^K_{(i,j)} \)), (ii) all the \( \hat{P}'_{(i,j)} \) such that \( (i,j) \in S(P) \) are mutually disjoint, and (iii) all the \( \hat{P}'_{(i,j)} \) such that \( (i,j) \in S(Q) \) are mutually disjoint.

Thus, by Axiom 2,

\[
P^K = \sum_{(i,j) \in S(P)} p^k_{(i,j)} \sim \sum_{(i,j) \in S(P)} \hat{P}'_{(i,j)},
\]

\[
Q^K = \sum_{(i,j) \in S(Q)} p^k_{(i,j)} \sim \sum_{(i,j) \in S(Q)} \hat{P}'_{(i,j)}.
\]

Hence, for all \( P, Q \in \mathcal{P}_n \)

\[
P \preceq Q \iff P^K \preceq Q^K \iff \sum_{(i,j) \in S(P)} \hat{P}'_{(i,j)} \preceq \sum_{(i,j) \in S(Q)} \hat{P}'_{(i,j)} \quad (4)
\]

Now, let

\[
S^k_{(i,j)} = \{Q \in \Delta^k_1| Q \sim p^k_{(i,j)}\}
\]

and

\[
S_{(i,j)} = \bigcup_{k \in \mathbb{N}} S^k_{(i,j)}
\]

We show that any two matrices in \( S_{(i,j)} \) contain the same number of non-null atomic matrices, and therefore this number depends only on \( i \) and \( j \). Let \( Q_1 \) and \( Q_2 \) be any two
matrices in $S(i,j)$. Then, for some $k, k', Q_1 \in \Delta^k_1, Q_2 \in \Delta^{k'}_1$, $Q_1 \sim p_{(i,j)}^k$ and $Q_2 \sim p_{(i,j)}^{k'}$. We assume without loss of generality that $k' \geq k$. By Lemma 1, $p_{(i,j)}^{k'} \sim Q_1^k$, and so $Q_1^{k'} \sim Q_2$. Since these two matrices are both in $\Delta^{k'}_1$, by Lemma 5, they must contain the same number of non-null atomic matrices. Moreover, the number of non-null atomic matrices contained in $Q_1^{k'}$ is the same as the number of those contained in $Q_1$. Thus, all the matrices in the set $S(i,j)$ contain the same number of non-null atomic matrices which depends only on $i$ and $j$. Let us denote it by $n(i,j)$ and let $f$ be the function $\mathbb{N} \mapsto \mathbb{N}$ such that, for every $i, j$, $f(|i - j|) = n(i,j)$. So, since $P'_{(i,j)}$ belongs to $S(i,j)$, the number of non-null atomic matrices contained in $P'_{(i,j)}$ is equal to $f(|i - j|)$.

Now, the matrices $\sum_{(i,j) \in S(P)} \hat{P}'_{(i,j)}$ and $\sum_{(i,j) \in S(Q)} \hat{P}'_{(i,j)}$ in (4) are in $\Delta^{k'}_1$. So, by Lemma 5, they can be compared by simply counting the number of non-null atomic matrices contained in them. This is equal to the sum of the numbers of non-null atomic matrices contained in each $P'_{(i,j)}$ which is, in turn, equal to $f(|i - j|)$. Therefore:

$$\sum_{(i,j) \in S(P)} \hat{P}'_{(i,j)} \preceq \sum_{(i,j) \in S(Q)} \hat{P}'_{(i,j)} \iff \sum_{(i,j) \in S(P)} f(|i - j|) \leq \sum_{(i,j) \in S(Q)} f(|i - j|). \quad (5)$$

Finally, from (4) and (5) it follows that:

$$P \preceq Q \iff \sum_{(i,j) \in S(P)} f(|i - j|) \leq \sum_{(i,j) \in S(Q)} f(|i - j|).$$

It is obvious, by Axiom 3, that $f$ must be strictly increasing. To show that $f$ must be strictly convex, consider, for any $k \geq 0$, the matrices $P$ and $Q$ such that:

$$S(P) = \{(1, k + 1), (2, k + 2), (k + 3, k + 3), (k + 4, k + 4), \cdots \}$$
$$S(Q) = \{(1, k + 2), (2, k + 1), (k + 3, k + 3), (k + 4, k + 4), \cdots \}.$$ 

Hence, by Axiom 1, we must have that, for all $k$, $2f(k) < f(k + 1) + f(k - 1)$. 

C Proof of Proposition 2

Proof. Within each class of similar matrices, to prove that $\preceq^C$ implies $\preceq^M$, given Definition 3 and Proposition 1, it suffices to show that for any partial matrices $P$ and $Q$ such that $P \prec Q$, we have $\sum_i d^P_i \geq \sum_i d^Q_i$ for all $j \leq k$, where $k$ denotes the number of non-zero entries in $P$ and $Q$ (the cardinality of $S(P)$ and $S(Q)$). Now, notice that when $P \prec Q$, $d^P$ and $d^Q$ differ only by two elements, say $d^P_m \neq d^Q_m$ and $d^P_n \neq d^Q_n$. The result follows since, as it is easily checked, $\max\{d^P_m, d^P_n\} > \max\{d^Q_m, d^Q_n\}$ and $d^P_m + d^P_n \geq d^Q_m + d^Q_n$. 

\[\square\]
D  Proof of Theorem 3

Proof. Given Theorem 2, we can concentrate only on Axiom 5. The reader can easily check that if \( f(k) = k^2 \) the ordering satisfies Axiom 5.

To show that, in order to satisfy Axiom 5, \( f \) must be quadratic, suppose \( P, Q, R \) are matrices in \( \mathcal{P}_n \) such that:

\[
S(P) = \{(0,0),(1,1),(0,k),(1,k+1)\}
\]
\[
S(Q) = \{(0,1),(1,0),(0,k),(1,k+1)\}
\]
\[
S(R) = \{(0,0),(1,1),(0,k+1),(1,k)\}.
\]

Thus, \( P \triangleleft Q \) and \( P \triangleleft R \), since the inversions that lead from \( P \) to \( Q \) and from \( P \) to \( R \) are both minimal. Then, by Axiom 5, \( Q \sim R \) and therefore:

\[
\sum_{(i,j)\in S(Q)} f(|i-j|) = 2f(1) + 2f(k) = \sum_{(i,j)\in S(R)} f(|i-j|) = 2f(0) + f(k+1) + f(k-1).
\]

Observe that, by definition of \( f \), \( f(0) = 0 \), and \( f(1) = 1 \) (see above, Appendix B.2). Therefore, to satisfy Axiom 5, since \( k \) is arbitrary and \( f \) is fixed for all \( n \), we must have that, for all \( k \),

\[
f(k+1) - f(k) = f(k) - f(k-1) + 2.
\]

This difference equation has a unique solution, i.e. \( f(k) = k^2 \).

\[ \square \]
References


