Measures of disassortativeness and their application to directly transmitted infections

H. J. Whitaker\textsuperscript{1}, C. P. Farrington\textsuperscript{1*}, J. Wallinga\textsuperscript{2} and P. Manfredi\textsuperscript{3}

\textsuperscript{1}Department of Mathematics and Statistics, The Open University, Milton Keynes, MK76AA, UK.

\textsuperscript{2}RIVM, Bilthoven, The Netherlands.

\textsuperscript{3}Department of Statistics and Mathematics Applied to Economics, University of Pisa, 56124 Pisa, Italy.

November 27, 2007

SUMMARY

We propose a measure of disassortativeness to summarize contact patterns relevant to the transmission of directly transmitted infections. We discuss the properties of this measure, describe standardization relative to homogeneous mixing, and generalize it to multivariate contact structures. We explore some of its properties and apply our methods to serological surveys of close contact infections and surveys of self-reported social contacts obtained in several European countries.

KEY WORDS: assortativeness, contact, contact survey, infectious disease, serological survey.
This study was funded as part of POLYMOD, a European Commission project funded within the Sixth Framework Programme, Contract number: SSP22-CT-2004-502084.

*Corresponding author: Paddy Farrington, Email: c.p.farrington@open.ac.uk.
1 Introduction

For infections directly transmitted from person to person, the contact rates between individuals in a population are a key determinant of the epidemiological features of the infection in that population. Together with the probability of infection from contact with an infectious person, the duration of the infectious period, and the level of immunity conferred by the infection, the contact rates determine the values of quantities such as the basic reproduction number and the critical immunisation threshold (Anderson & May 1991). For mathematical transmission models, it is essential to have accurate estimation of contact rates between and within groups. The estimation of contact rates is thus a central preoccupation of infectious disease epidemiology. Such rates may be estimated indirectly from serological survey data or case reports (Anderson & May (1991), Farrington et al. (2001), Farrington & Whitaker (2005)), or directly through surveys (Renton et al. (1995), Edmunds et al. (2006), Wallinga et al. (2006)). Further approaches are to infer contact rates from other relevant data on social activity, for example by simulation methods conditioned on marginals determined by census, land use and mobility data (Eubank et al. (2004), Del Valle et al. (2007)), or more directly using time-use data (Zagheni et al. (2007)).

It has long been known that the epidemiological features of an infection are sensitively dependent on the contact structure (Jacquez et al. (1988), Greenhalgh & Dietz (1996)). In particular, it is of interest to know the degree of assortativeness of the contacts, that is, the extent to which contacts occur between individuals who share characteristics that are relevant to the transmission of infection. Disassortativeness is the converse, namely the extent to which contacts occur between individuals with different characteristics. Much of this focus has been limited to sexually transmitted infections, some of which require a high degree of assortativeness within a ‘core group’ of individuals with high levels of partner change to maintain the infection at endemic levels. However, the degree of assortativeness
is important for other infections as well, in particular those transmitted by close contact via the respiratory route, which include the main infections of childhood. Even though what constitutes a contact in such circumstances is rather nebulous (Diekmann & Heesterbeek (2000)), there is evidence that infection is transmitted more often between pairs of persons that have similar age than between pairs of persons that differ in age (Wallinga et al. (1999), Kanaan & Farrington (2005)).

In this paper we describe an index of disassortativeness that can be applied to any type of contact, with respect to any relevant individual characteristic, and which can be computed for any type of data from which contact rates may be derived. In section 2 we briefly review the measures and methods for evaluating assortativeness and disassortativeness that have been proposed in the literature. In section 3 we describe our proposed new measure and derive some of its properties. In section 4 we give some examples. In section 5, we apply this measure to two contrasting types of data, collected in various European countries: serological survey data and data on self-reported at-risk contacts for infection. We compare the disassortativeness of the contact structures in different countries. We finish with a brief discussion in section 6.

2 Measures of assortativeness

Three approaches to quantifying assortativeness as it relates to infectious diseases have been proposed in the literature, primarily in the context of sexually transmitted infections. The first relates to networks, and particularly sexual networks. In such networks, individuals are represented by nodes and contacts between individuals by edges. The degree of a node is its number of edges. The question then arises as to whether edges tend to link nodes of high degree. If so, the network is said to be assortative. A natural measure of
assortativeness in this context is the Pearson correlation coefficient $\rho$ between the degrees of linked pairs of nodes, or a simple modification of it. Such a measure has been described by Newman (2003).

The correlation coefficient can be used to quantify assortativeness in contact structures other than networks, and in particular in contact structures described by a contact function $\beta(s, t)$. Such descriptions are used to describe contacts in large populations, for which it is impossible to document the entire network. The contact function is defined in the next section; at this stage it suffices to say that it represents the rate at which contacts occur between individuals of types $s$ and $t$. The contact function determines a bivariate contact density for contact pairs $(s, t)$. A natural extension of the Pearson correlation for networks is thus to use the correlation of $S$ and $T$ with respect to this density. High values of the correlation are then interpreted as indicating a high degree of assortativeness, values close to 0 indicate proportional mixing (Dietz & Schenzle (1985)), where $\beta(s, t) = \gamma(s)\gamma(t)$, and values close to $-1$ indicate highly disassortative mixing.

One difficulty with correlation-based measures is that correlation properly describes only the extent of linear association. Thus if individuals of type $s$ only ever contact individuals of type $2s$, or individuals of type $1 + s$, then the correlation is 1 even though contacts are not assortative (in the sense that there are no contacts between individuals of the same type). A further difficulty is the apparent symmetry between ‘assortative’ and ‘disassortative’ implied by the correlation coefficient. In fact, whereas complete assortativeness is easily defined (individuals of type $s$ only make contact with other individuals of type $s$), disassortativeness has no such unique definition in general. This is because assortativeness is always defined with respect to some characteristic. If this characteristic is represented using an unbounded scale, then differences $|s - t|$ may be unbounded, in which case there is no ‘maximum’ dissimilarity.
In many applications, the contact function $\beta(s, t)$ is a low dimensional matrix. An ad-hoc measure of assortativeness, denoted $Q$, has been proposed based on the proportion of contacts that occur along the main diagonal of the matrix (Jacquez et al. (1988), Gupta et al. (1989), Ghani & Garnett (1998)). This measure is defined in such a way that $Q$ takes the value 1 when all contacts are on the diagonal, corresponding to complete assortativeness, the value 0 when contacts are evenly spread (corresponding to homogeneous mixing), and takes values $< 0$ when contacts are disassortative. This measure removes the inappropriate symmetry between assortativeness and disassortativeness implied by the correlation coefficient $\rho$. However, one problem with the measure $Q$ is that it relies on a binary categorization of contacts as either assortative or not, and hence its value depends rather strongly on the groupings used to define the diagonal of the contact matrix. This is not a major problem for the intended application of $Q$, to frequencies of sexual partners, but may be a difficulty in other contexts, particularly when the individual types $s$ can be described on a continuous scale, as is the case with age.

The third approach is to build a model for the contact function, with a model parameter to quantify assortativeness. One such approach, described by Hethcote (1996), is based on the preferred mixing model of Nold (1980) and Newman (2003). The contact function $\beta(s, t)$ is modelled as a mixture of a completely assortative component and a proportional mixing component:

$$
\beta(s, t) = \epsilon \delta(s - t) \alpha(s) + (1 - \epsilon) \gamma(s) \gamma(t)
$$

where $\delta(x)$ is the Dirac delta function. The measure of assortativeness is then $\epsilon$, taking the value 1 when contacts are completely assortative and 0 when contacts are proportional. The main drawback of this method is the underlying assumption that the contact function can be modelled using a simple mixture model, which is unlikely to hold in practice. More elaborate models have been constructed for count data from sexual contact surveys,
summarized in square contingency tables (Morris (1991), Morris & Dean (1994)). Such
types of models, commonly used in the social sciences, may be analysed using generalized
linear models (see Koehly et al. (2004) for a review) or Monte Carlo techniques (Smith
et al (1996)). In such models, assortativeness is typically represented by the odds ratio
associated with the groupings on the diagonal of the contingency table. Like the more
heuristic measure described above, this measure requires a binary categorization of contacts
as assortative or not, and hence depends strongly on the groupings used to define the
diagonal of the contact matrix. Furthermore, these modelling techniques are not applicable
to sources of data other than contact surveys.

The measures of assortativeness described above relate strongly to the types of data for
which they were developed. Ideally, an index of assortativeness is required which may be
estimated from any contact surface, and which is not tied to a particular type of data or
model. This is desirable in view of the wide variety of data from which contact rates can
be estimated, as noted in the Introduction.

3 A general index of disassortativeness

In the next subsection we introduce an index of disassortativeness in the important special
case when contact rates vary with age. Later we will generalise this index to other variables,
and extend it to multivariate contact data.

3.1 Age-dependent contact rates

In what follows, contact rates are assumed to vary with age, and other sources of variation
are ignored. Thus, contacts may be represented by ordered pairs of ages $s$ and $t$, which take
values on $[0, \infty)$. The contact rate is a bivariate function $\beta(s, t)$ describing the per-capita
rate (per unit time) at which an individual of age $t$ makes contact with an individual of age $s$. The contact function need not be symmetric, which is why it may be important to consider ordered pairs $(s, t)$. Whether or not the contact rate function is symmetric depends on the transmission route. For example, for infections transmitted by transfer of saliva, and hence with gendered transmission, as is the case with Epstein Barr virus, the age specific contact rates between males and females may not be symmetric (Farrington & Whitaker (2005)). However, for many infections transmitted by the respiratory route it is sensible to assume that the contact rate function is symmetric.

Let $f(s)$ denote the density of the population age distribution, and $\sigma$ its variance. Then the contact pairs within the population have density

$$f_c(s, t) = \frac{f(s)\beta(s, t)f(t)}{\int_0^\infty \int_0^\infty f(u)\beta(u, v)f(v)dudv}.$$  

Many different definitions of contacts and contact matrices exist in the literature. We briefly clarify the relationships between these various definitions. Thus, Wallinga et al. (2006) describe contacts between individuals classified in discrete age groups in terms of a social contact matrix with entries $m_{st}$ that denote the number of individuals in age group $s$ that are contacted by an average individual in age group $t$ per unit time. If $n$ denotes the size of the population and $n_s = nf(s)$, the number of individuals of age $s$, then $m_{st} = n_s\beta(s, t)$. If $q$ is the probability of infection, given that a contact has occurred, and $D$ is the average duration of infectiousness (assumed to be short), then the next generation matrix (Diekmann & Heesterbeek (2000)) has entries $qDm_{st} = qDn_s\beta(s, t)$. Note that when contacts are reciprocal (in the sense that individual A makes contact with individual B exactly when B makes contact with A, and hence $\beta(s, t) = \beta(t, s)$) the social contact matrix and the next generation matrix are symmetric only when the age groups are of equal size. In general, this is not the case.

A special case is homogeneous mixing, in which the contact rate is constant, $\beta(x, y) = \beta$. 

In this case, \( f_c(s, t) = f(s) f(t) \). The extreme case of complete assortativeness is represented by the function \( \beta(s, t) = \gamma(s) \delta(s - t) \) where \( \delta(.) \) is the Dirac delta function. In this case, \( f_c(s, t) \) is zero except on the diagonal \( s = t \). More generally, a randomly selected contact pair \((s, t)\) may be represented as a point on the plane. The perpendicular distance from this point to the diagonal is \(|s - t|/\sqrt{2}\). This may be thought of as this pair’s ‘deviation from assortativeness’, as represented schematically in Figure 1.

A natural index of absolute disassortativeness is the mean squared deviation from assortativeness

\[
I^2 = \frac{1}{2} E_c(S - T)^2
\]  

where \( E_c \) is the expectation with respect to \( f_c \). This index takes values on \([0, \infty)\). It is zero if and only if contact pairs \((s, t)\) with \( s \neq t \) occur with probability zero, that is, exactly when contacts are completely assortative. Increasingly disassortative contacts result in increasingly large positive values of \( I^2 \). As noted in the previous section, there is no uniquely meaningful definition of ‘complete disassortativeness’, and so no special value is assigned to this scenario.

The quantity \( I^2 \) has a long history in statistics. It was first introduced by Gini (1914) who called it ‘quadratic measure of dissimilarity’ (dissomiglianza). In its present form, \( I^2 \) is defined for symmetric and non-symmetric contact functions \( \beta(s, t) \); \( 2I^2 \) is the second moment about zero of \( S - T \), which coincides with the variance for symmetric contact functions, but not otherwise, since we seek to quantify the degree of disassortativeness, and hence the spread around the diagonal \( s = t \) representing complete assortativeness, not the spread around \( E_c(S - T) \).

A special value of \( I^2 \) of particular interest is that corresponding to homogeneous mixing, that is \( \beta(s, t) = \beta \). The (absolute) degree of disassortativeness is then determined entirely
by the density \( f(s) \), and is

\[
I^2 = \frac{1}{2} \int_0^\infty \int_0^\infty (s - t)^2 f(s)f(t)dsdt = \sigma^2,
\]

the variance of the age distribution. In this context, the square root of \( 2I^2 \) was called ‘quadratic mean difference’ by Gini (1912).

To compare the degrees of disassortativeness induced by the contact function \( \beta(s,t) \) in populations with different age densities \( f(s) \), it is desirable to calibrate the index \( I^2 \) relative to homogeneous mixing. This makes sense because the shape of a homogeneous mixing surface (i.e., a flat surface) is the same in all populations; this is not the case for other choices, such as proportional mixing where \( \beta(x,y) = \gamma(x)\gamma(y) \). This leads to the following standardized measure of global assortativeness:

\[
I^2_s = \frac{1}{\sigma^2} I^2
\]

which takes the value 0 when mixing is completely assortative and the value 1 when mixing is homogeneous. Note however that different contact surfaces can have different values of \( I_s > 0 \), thus if \( I^2_s = 1 \) it does not follow that mixing is homogeneous. Another method of calibration, yielding the value 1 for the proportional mixing model based on marginal contact distributions, is discussed in section 4.

### 3.2 Other variables

Contact rates may vary according to variables other than age. For example, for sexually transmitted infections, sexual activity levels are important. For infections transmitted via the respiratory route, individual sociability or occupation type may also play a role. The definitions introduced in the previous subsection can be generalised directly to such
variables, using equations (1) and (2), where the contact distribution now involves the population density of the variable of interest. For example, if the variable of interest is some measure of sociability $s$, then a contact pair $(s, t)$ denotes a contact between a person of sociability $s$ and one of sociability $t$. The density $f(s)$ now describes the distribution of sociability levels in the population, and $\sigma$ their standard deviation.

Note that the absolute index $I^2$ depends on the scale of measurement of the characteristic of interest. This scale dependence can be removed by standardizing the variable of interest. For example, sociability $s$ can be replaced by its standardized value $s/\sigma$ ($s$ could also be centered, but this is not essential). It is easy to show that $I^2$, calculated with equation (1) using standardized variables, is identical to $I^2_s$. Thus, standardization in the statistical sense (so that variables have unit variance) is equivalent to calibration with respect to homogeneous mixing. The standardized index $I^2_s$ is scale free, and for this reason is generally preferable to the absolute index $I^2$.

### 3.3 Multivariate contact structures

More generally, contact rates may vary with several variables, which might include age and individual activity levels. Thus it is natural to extend the definition of contacts from ordered pairs of singletons to ordered pairs of $k$-dimensional vectors $s$ and $t$ in some subset $C$ of $R^k$. For example, $s$ could represent a pair $(x, u)$ where $x$ is age and $u$ is an individual activity level.

Generally, most interest will reside in the component-specific measures of assortativeness, which may be obtained by applying equations (1) and (2) separately to each variable. However, it is also possible to apply the methods described to obtain a global measure of assortativeness across all variables. Usually, the different components will be measured in different units, and so a global measure only really makes sense if the different variables
are standardized.

For the remainder of this section, we assume that the $k$ contact variables are standardized so that they have unit variance. A global measure of assortativeness is then

$$I_{global}^2 = \frac{1}{2k} E_c(|S - T|^2)$$

where $|.|$ denotes Euclidean distance. In this expression, the expectation is taken with respect to the density of $k$-variate contact pairs

$$f_c(s, t) = \frac{f(s)\beta(s, t)f(t)}{\int \int_C f(u)\beta(u, v)f(v)dudv}$$

where $f(s)$ is the joint density of the $k$ standardized variables in the population.

Other norms than the Euclidean could be used, but an advantage of this choice is that the (standardized) index is additive in its components, that is

$$I_{global}^2 = \frac{1}{k} (I_1^2 + \ldots + I_k^2)$$

where $I_i^2 = \frac{1}{2} E_{ci}((S_i - T_i)^2)$, the expectation $E_{ci}$ being taken with respect to the marginal bivariate distribution

$$f_{ci}(s_i, t_i) = \int \ldots \int f_c(s, t)ds_1dt_1 \ldots ds_{i-1}dt_{i-1}ds_{i+1}dt_{i+1} \ldots ds_kdt_k.$$

Further simplification in the form of this distribution ensues when the $k$ components are independent, which occurs when $f(s) = f_1(s_1) \times \ldots \times f_k(s_k)$ and $\beta(s, t) = \beta(s_1, t_1) \times \ldots \times \beta(s_k, t_k)$.

It is worth re-emphasizing that the multivariate index of assortativeness only makes sense when calculated with standardized variables, in which case each component $I_i^2$ is dimensionless, as is the global index $I_{global}^2$. This takes the value zero when mixing is completely assortative in all its components, and the value 1 when mixing is homogeneous in all its components.
4 Examples

In this section we work through a few examples and further ideas involving the proposed indices $I^2$ and $I_s^2$.

Example 1: Proportional mixing

Suppose that contact rates vary with age alone, with the proportional mixing structure $\beta(s, t) = \gamma(s)\gamma(t)$ (Dietz & Schenzle (1985)). Then the contact density $f_c(s, t)$ is of the form $f_c(s)f_c(t)$, where

$$f_c(s) = \frac{\gamma(s)f(s)}{\int_0^\infty \gamma(t)f(t)dt}$$

is the marginal density of age at contact, that is, the marginal density $\int f_c(s, t)dt$. In this case, the index of disassortativeness is $I^2 = V_c(S)$, the variance of the age at contact. The standardized index is $I_s^2 = V_c(S)/V(S)$ where $V(S) = \sigma^2$ is the variance of age in the population. This is equal to 1 for homogeneous mixing, but in general can be either greater than 1 (for example, if $\gamma(s)$ is largest in the tails of the age distribution) or less than 1 (for example, if $\gamma(s)$ is largest at central ages).

Example 2: Relationship with the correlation coefficient

Suppose that contacts vary with age and that the contact function $\beta(s, t)$ is symmetric. Let $\rho$ denote the Pearson correlation between the ages $S$, $T$ in contact pairs $(S, T)$, that is,

$$\rho = \frac{E_c(ST) - E_c(S)E_c(T)}{\sqrt{V_c(S)V_c(T)}}.$$
The correlation coefficient $\rho$ is related to the index of dis assortativeness by

\[ I^2 = E_c(S^2) - E_c(ST) \]

\[ = E_c(S^2) - E_c(S)^2 - \{E_c(ST) - E_c(S)^2\} \]

\[ = (1 - \rho)V_c(S) \]

and to the standardized index by

\[ I_s^2 = (1 - \rho) \frac{V_c(S)}{V(S)}. \] (4)

Note that $1 - \rho$ is equal to 1 for proportional mixing, whereas $I_s^2$ is equal to 1 only for homogeneous mixing. However, for a given age distribution $f(s)$, functions $\gamma(s)$ of different shapes will generally result in different values of $I^2$ and $I_s^2$. In consequence, proportional mixing can be assortative or disassortative when compared to homogeneous mixing.

**Example 3: An alternative standardization**

We have proposed that the absolute measure of dis assortativeness $I^2$ should be calibrated with respect to homogeneous mixing, which turns out to be equivalent to standardizing the distribution of the characteristic of interest. Thus, for example, the variable age is replaced by age divided by its standard deviation $\sigma$. Then the standardized index $I_s^2$ is equal to 1 when mixing is homogeneous.

This is not the only possible method of standardization. For definiteness we consider age, and suppose that the contact function $\beta(s,t)$ is symmetric. An obvious alternative is to standardize with respect to the marginal age of contacts

\[ f_c(s) = \frac{f(s)\int_0^\infty \beta(s,t)f(t)dt}{\int_0^\infty \int_0^\infty f(s)\beta(s,t)f(t)dsdt}. \]

Let $V_c(S)$ denote the variance of this marginal density. Thus the index of dis assortativeness
standardized with respect to contact age is

\[ I_c^2 = \frac{I^2}{V_c(S)}. \]  

(5)

This is equal to 1 when \( \beta(s, t) \propto \gamma(s)\gamma(t) \), where \( \gamma(s) \) is

\[ \gamma(s) = \int_0^\infty \beta(s, t)f(t)dt. \]

Thus \( I_c^2 = 1 \) when the contact function corresponds to proportional mixing, with the same marginal contact densities as those observed. It is readily shown that \( I_c^2 = 1 - \rho \). Thus measuring disassortativeness using the correlation coefficient is equivalent to our measure, standardized with respect to the marginal contact distribution.

Which standardization is best? Consider two scenarios A and B, each with two age classes, and the following proportional mixing contact functions:

\[
\beta_A(s, t) = \begin{pmatrix}
100 & 10 \\
10 & 1
\end{pmatrix} \quad \text{and} \quad \beta_B(s, t) = \begin{pmatrix}
10 & 10 \\
10 & 10
\end{pmatrix}
\]

With a uniform age distribution, we obtain \( I_c^2 = 1 \) (and \( \rho = 0 \)) for both A and B, indicating equivalent degrees of disassortativeness, whereas \( I_s^2 = 0.33 \) for A and \( I_s^2 = 1 \) for B, indicating that A is more assortative than B. This is clearly appropriate, since 83.5% of contacts for A are on the main diagonal compared to 50% for B. Thus standardization with respect to the age distribution, rather than the contact distribution, arguably produces a more meaningful measure of disassortativeness.

Note from equations (4) and (5) that we can write

\[ I_s^2 = I_c^2 \times \frac{V_c(S)}{V(S)}. \]

Thus the index of disassortativeness \( I_s^2 \) factorizes into two components: the component \( I_c^2 \), which represents the degree of disassortativeness relative to that induced by the marginal contact distribution under independence, and the component \( V_c(S)/V(S) \) which represents
the degree of disassortativeness induced by the marginal contact distribution, relative to homogeneous mixing.

Example 4: A bivariate model

Farrington et al. (2001) and Farrington & Whitaker (2005) fit models to serological survey data using a bivariate contact surface of the form

$$\beta(x, u; y, v) = uv\beta_0(x, y)$$

(6)

where \(u, v\) are individual activity levels of mean 1 and \(x, y\) denote age. The individual activity levels and the ages are assumed to be independent. Let \(f(x)\) and \(g(u)\) denote the marginal densities for age and individual activity levels, respectively. Then

$$f_c(x, u; y, v) = uv g(u)g(v) \frac{f(x)\beta_0(x, y)f(y)}{\int_0^\infty \int_0^\infty f(x)\beta_0(x, y)f(y)dxdy},$$

where \(\beta_0(x, y)\) is the per-capita age-specific contact rate.

Suppose that the (absolute) disassortativeness indices for age and activity levels take values \(I^2_a\) and \(I^2_i\), respectively. The global disassortativeness index is then

$$I^2_{global} = \frac{1}{2} \left( \frac{I^2_a}{\sigma^2_a} + \frac{I^2_i}{\sigma^2_i} \right)$$

where \(\sigma^2_a\) and \(\sigma^2_i\) are the variances of age and individual activity levels in the population, respectively. The value of \(I^2_a\) depends on \(\beta_0(x, y)\):

$$I^2_a = \frac{1}{2} \sum_x \sum_y (x - y)^2 \frac{f(x)\beta_0(x, y)f(y)}{\sum_x \sum_y f(x)\beta_0(x, y)f(y)}.$$

If the activity levels \(U, V\) are distributed gamma with mean 1 and variance \(\sigma^2_i\), then

$$I^2_i = \frac{1}{2} \int_0^\infty \int_0^\infty (u - v)^2 uv g(u)g(v)du dv
= \sigma^2_i \left( 1 + \sigma^2_i \right)$$
and hence the standardized measure of disassortativeness for individual activity levels is 
$1 + \sigma_i^2$. Since this is always greater than 1, this particular model for individual activity 
levels is dissasortative relative to homogeneous mixing.

**Example 5: Ridge plots**

A criticism of any numerical index is that a single number inevitably loses valuable information. We therefore propose to complement the indices $I^2$ or $I_s^2$ by an easily interpretable plot which may provide additional insight.

We consider a single variable, which for definiteness we shall take to be age, denoted $s$ and $t$. The parameterization $(s,t)$ may be replaced by $(u,v)$, as described in Farrington & Whitaker (2005), where

$$u = (s + t)/\sqrt{2}, \quad v = (s - t)/\sqrt{2}.$$ 

The Jacobian of the transformation from $(s,t)$ to $(u,v)$ coordinates is 1, hence the index $I^2$ may be written

$$I^2 = \int v^2 w(v) dv$$

where the $w(v)$ are the weights

$$w(v) = \int f_c \left( \frac{u + v}{\sqrt{2}}, \frac{u - v}{\sqrt{2}} \right) du.$$ 

The weight $w(v)$ is the density of contacts on the subdiagonal parallel to the complete assortativeness line $s = t$, lying $v$ units from it along the perpendicular. In particular, when the variable of interest is defined on $[0, \infty)$, as with age, then

$$w(v) = \int_{|v|}^{\infty} f_c \left( \frac{u + v}{\sqrt{2}}, \frac{u - v}{\sqrt{2}} \right) du.$$ 

Now graph $w(v)$ for all values of $v$; the graph will generally be more easily interpretable if $w(v)$ is graphed against $\sqrt{2}v$, which equals the age difference $s - t$ between contacted
and contacting individual, rather than against $v$. When the contact function is symmetric, $w(v)$ is symmetric about $v = 0$. Ridges of high contact density for specific age differences $s - t$ will appear as bumps in the plot. Such ridges are important features, as they might reveal inherent preferences for contacting individuals within similar age cohorts (class mates, friends, life partners) and for contacting individuals of different generations (parents, children).

5 Applications

We describe the estimation of measures of disassortativeness from two distinct data types: serological surveys and contact surveys carried out in several European countries, primarily as part of the POLYMOD project. We describe each model and results in turn, then comment on the results of both sets of analyses in a separate subsection.

5.1 Serological survey data

Paired serological data on Varicella Zoster and Parvovirus were obtained in Italy in 2003/4 ($n = 2444$) and Finland in 1997/8 ($n = 2470$) (Mossong et al. (2007)). We also used data from a 1986 serological survey in the UK, yielding paired serological data on mumps and rubella for 4575 boys (Morgan-Capner et al. (1988)).

5.1.1 The model

The model for the contact surface is that described in equation (6). The individual activity levels and the ages are assumed to be independent; we assume that individual activity levels are gamma with variance $\theta^{-1}$. We follow Farrington & Whitaker (2005) and set $\beta_0(x, y) = \gamma(m) \times b(d|m)$ where $m = (x + y)/\sqrt{2}$ and $d = (x - y)/\sqrt{2}$. The functions $\gamma$
and $b$ are proportional to gamma and symmetric beta densities, respectively. The function $\gamma$ represents the variation in contact rates along the main diagonal $x = y$ of the contact surface, and the function $b$ represents the spread perpendicular to the diagonal. Thus

$$\gamma(m) \propto m^{\nu-1} \exp \left( -\frac{vm}{\sqrt{2\mu}} \right), \quad b(d|m) = \frac{(m + d)^{\alpha-1}(m - d)^{\alpha-1}}{m^{2\alpha-2}}.$$ 

The procedure for fitting this model to the paired serological survey data by maximum likelihood is described in detail in Farrington & Whitaker (2005). Briefly, using an endemic equilibrium assumption, this model for the contact surface leads to a frailty model for the force of infection, which comprises six parameters: the parameters $\theta$, $\nu$, $\mu$ and $\alpha$ for the contact surface, and two scaling parameters (one for each infection). For each year of age, the bivariate seropositivity counts (-/-, -/+), (+/-, +/-) form a multinomial vector of index 4, with probabilities derived from the frailty model. The multinomial likelihood is then maximised with respect to the six parameters. The absolute, standardised and global indices of dissassortativeness were calculated numerically from the fitted contact surface, and 95% confidence intervals were obtained using non-parametric bootstrapping (399 bootstrap samples) within age groups.

5.1.2 Results

The model deviances, degrees of freedom and $p$ values for goodness of fit are shown in Table 1. Figure 2 shows the observed and fitted values for the Italian data; similar plots (not shown) were obtained for the other data sets. The increased spread of the data at older ages is due to lower sample sizes in those age groups. We conclude that the models fit well. Figure 3 shows the fitted contact functions $\beta_0(x,y)$ for the three data sets, and Table 2 shows the global ($I^2$) and standardised ($I^2_s$) indices of dissassortativeness with respect to both age and activity levels, as well as the combined, global measure $I^2_{\text{global}}$. 

19
5.2 Contact survey data

Survey participants in eight countries were asked to record all the individuals they came into contact with within one assigned day (Mossong et al. (2007)). A contact was defined by several criteria, the most important of which was a two-way conversation. The participants recorded the ages of contacts (children’s contacts were recorded by their parents). Depending on country surveyed, only between 20% and 36% (average 30%) of contact ages were recorded exactly, the others being quoted within a range, typically within 5 years. Accordingly, to avoid introducing spurious accuracy, we assigned the ages of the participants and their contacts to sixteen 5-year age bands: 0-4, 5-9, 10-14, ..., 70-74, 75+. For each country sixteen subtables were created corresponding to the participants’ age groups (indexed by $i$, $i = 1, \ldots, 16$). Each subtable comprises $N_i$ rows, one for each of the $N_i$ individuals sampled of age $i$ (indexed by $k_i = 1, \ldots, N_i$; the subscript $i$ is retained to make the age group explicit) and 16 columns (indexed by $j$, $j = 1, \ldots, 16$) for the ages of their contacts. The cells contain the number of people contacted, which we denote $r_{jk_i}$: this is the number of persons of age $j$ contacted by the $k_i$th individual of age $i$ (where $k_i$ ranges from 1 to $N_i$).

5.2.1 The model

These contact data were modelled within a Bayesian framework. We assumed that contacts are bivariate, the contact function varying with age and with individual activity levels, as specified in equation (6). Here the age-specific contact rate $\beta_0(x, y)$ takes the form of a contact matrix $\beta_{ij}$.

No parametric form need be assumed for the matrix $\beta_{ij}$, since the ages of the contacts were documented. Information on the distribution of the activity levels $u$ is obtained from between-individual variation in the number of contacts. However, since data on the contact
patterns of the contacts (specifically, their activity levels) were not collected, it is necessary to assume a parametric form for how the u’s and v’s combine. Model (6) is perhaps the simplest such model, and can be shown to correspond to a frailty model for the forces of infection (Farrington et al. (2001)).

We used an over-dispersed Poisson model for the counts $r_{jk_i}$, with mean $\alpha_{jk_i}$ and overdispersion parameter $\psi$:

$$r_{jk_i} \sim \text{Poisson}(\lambda_{jk_i})$$

$$\lambda_{jk_i} \sim \text{Gamma}(\text{mean} = \alpha_{jk_i}, \text{precision} = \frac{\psi}{\alpha_{jk_i}^2}).$$

The mean $\alpha_{jk_i}$ was related to the contact parameters of interest as follows:

$$\alpha_{jk_i} = u_{k_i} \beta_{ij} f_j$$

where $f_j$ is the proportion of the population in age class $j$, obtained from life tables for each country. The individual activity levels for the survey participants were assumed to be gamma distributed:

$$u_{k_i} \sim \text{Gamma}(\text{mean} = 1, \text{precision} = \theta).$$

Note that these activity levels account for the hierarchical nature of the data, and the likely dependence between counts within individuals. Finally we assumed that the age-specific contact matrix is symmetric:

$$\beta_{ji} = \beta_{ij},$$

thus survey participants of age $i$ and their contacts of age $j$ contact each other at the same rate. Weak priors were placed on $\beta_{ij}$, $\theta$ and $\psi$.

5.2.2 Results

The model parameters were estimated using WinBUGS. We also fitted a non-symmetric model (with no allowance for overdispersion), and obtained very similar results. We con-
clude that the symmetry assumption, which must hold in practice, is not invalidated by
biased responses. The results were also insensitive to the choice of prior.

Table 3 shows the global ($I^2$) and standardised ($I^2_s$) indices of dissassortativeness with
respect to both age and activity levels, with 95% credible intervals, as well as the combined,
global measure $I^2_{global}$. Figures 4 and 5 show the estimated contact matrices $\beta_{ij}$ (summarized
by the posterior means) for Italy and Poland respectively. These matrices are scaled such
that each entry represents the percentage of all contacts:

$$100 \times \frac{\beta_{ij}}{\sum \beta_{ij}}.$$ 

Ridge plots were obtained for all eight countries; they show similar features. The plots for
Poland, Italy and Belgium are shown in Figure 6, with the age differences $s - t$ on the
horizontal scale.

5.3 Commentary on the results

The contact surfaces for Finland, Italy and the UK estimated from serological survey data
(Figure 3) share key features. Most contacts occur, highly assortatively, in childhood,
whereas contact rates in adulthood are much lower and much less assortative. The indices
of age-related assortativeness in Table 2 are broadly similar in the three countries, and
substantially less than 1, indicating that contacts are highly assortative with age. Assor-
tativeness associated with activity levels also differs little between countries, as indicated
by the confidence intervals for $I^2$ and $I^2_s$ in Table 2. Note however that, unlike Italy
and the UK, the Finnish data are consistent with zero activity level variance (and hence
independence within individuals of seropositivity to VZV and rotavirus infection).

The assortativeness indices estimated from contact survey data in Table 3 also indicate
great similarity between countries in assortativeness associated with both age and activity
levels. The estimated age-related indices $I_s^2$ lie in the range $0.44 - 0.52$, with narrow credible intervals, indicating substantial assortativeness compared to homogeneous mixing. Figures 4 and 5 show that contact rates are generally higher on or close to the main diagonals than far from the main diagonals, thus confirming the observation that contacts are more assortative than under homogeneous mixing. The main feature of the ridge plots in Figure 6 is a very large peak at 0, indicating a strong assortative feature. Secondary peaks at age differences of $20 - 30$ most likely correspond to parent-child contacts; these are particularly marked for Belgium. In some countries (not shown) a further faint peak is discernible at age differences of about 55 years, perhaps corresponding to contacts between children and grandparents.

It is instructive to contrast the results obtained using serological survey data with those obtained using contact surveys. The contact surfaces associated with both age and activity level are more assortative when estimated from serological survey data than from contact survey data, as witnessed informally by the non-overlap of interval estimates. One reason is that the gamma-beta parametric model used with the serological survey data precludes secondary peaks such as those observed in Figure 6. However, the main reason is that age-related contact rates are much lower in adulthood generally when estimated with this model from serological data than when estimated from contact survey data. Also, individual variability associated with activity levels is greater when estimated from contact surveys that from serological data, even though in this instance the same contact surface model was used for the two sets of data. The contrast may be due to exclusion of non-conversational contacts from the contact surveys, which may be less dependent on individual characteristics, yet may play an important role in transmission.

The two estimation methods have complementary advantages and disadvantages. The main advantage of serological data is that the estimated contact rates relate directly to
contacts relevant to the transmission of infection; its main disadvantage is that strong parametric assumptions are required to define a suitable family of contact surfaces. The main advantage of contact surveys is that the age-related contact surface may be estimated largely non-parametrically (though the activity-level surface still depends on parametric assumptions); its main disadvantage is that there is no guarantee that the contacts surveyed correspond to those relevant to the transmission of infection.

6 Discussion

We have proposed a measure of disassortativeness to summarize contact patterns in a population, in the context of infectious disease modelling for infections directly transmitted from person to person. This measure of disassortativeness, with respect to any characteristic, is the mean squared deviation of contact pairs from the diagonal representing complete assortativeness. Its standardized value, when calibrated relative to homogeneous mixing, defines a measure that may be compared meaningfully between populations with different marginal distributions of the characteristic of interest. This standardized index takes the value 0 for completely assortative mixing, and the value 1 for homogeneous mixing; disassortative mixing corresponds to values greater than 1. It can be generalized to contacts described by more than one variable.

The indices of disassortativeness considered here are quite general, and may be estimated from a variety of types of data, including network data, serological survey data (provided a parametric model for the contact surface is available) and contact survey data. Further detail, particularly on inter-generational contacts, may be obtained using ridge plots. We have described applications to serological and contact survey data collected in several European countries; these show that contacts are assortative with respect to age
and activity levels.

The index we have described is only a summary measure, and thus cannot hope to capture all aspects of a contact surface that are relevant to the transmission of infection. In particular, contact surfaces with different shapes may share the same index. It is tempting to seek to relate the degree of disassortativeness to other summary quantities relevant to the transmission of infectious diseases, such as the basic reproduction number $R_0$. This, however, depends primarily on the absolute values of the per-capita contact rates $\beta(s, t)$, whereas measures of disassortativeness are determined by the relative variation in $\beta(s, t)$, that is, by the shape of the contact surface rather than its mean value. Nevertheless, it is interesting to note that the impact on $R_0$ of modelling individual variation using a gamma distribution (see Farrington et al. (2001)) is to increase its value by a factor of about one plus the variance of the individual effect, which is the standardized disassortativeness index derived in the present paper. The degree of disassortativeness is related to the connectedness of a population and hence to the speed at which an infection will spread throughout the population. Thus it is likely that the index of disassortativeness is related to the ratio of the dominant to subdominant eigenvalues of the next generation operator. More work is required to elucidate further such connections.

The degree of disassortativeness of at-risk contacts for spreading infections is an important characteristic of contact patterns. We have shown that the degree of disassortativeness can usefully be summarized by a single index, and have found that, for contacts associated with age and activity level, this index is remarkably constant across several countries in Europe. This is helpful for constructing mathematical models of transmission that capture realistic human contact patterns. The index also facilitates comparisons between different estimation methods, and hence the continuing search for improvements in such techniques.
Acknowledgements

We thank the POLYMOD consortium for access to the contact data and the VZ and parvovirus serological survey data, and the HPA Centre for Infections for the UK mumps and rubella data.

References


Edmunds J. Social contacts and mixing patterns relevant to the spread of infectious diseases. *Submitted for publication.*


Figure 1: Schematic depiction of a contact surface. The diagonal $s = t$ is the line of complete assortativeness.
Figure 2: Observed (○) and fitted (lines) proportions with specified varicella zoster virus (VZ) and parvovirus (P) serologic status (+ seropositive, - seronegative) by age.
Figure 3: Contour plots for contact surfaces with respect to age estimated from serological survey data.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>3.1</td>
<td>0.8</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>5-9</td>
<td>0.8</td>
<td>6.7</td>
<td>1.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.7</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>10-14</td>
<td>0.2</td>
<td>1.1</td>
<td>1.1</td>
<td>1.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.7</td>
<td>0.6</td>
<td>0.8</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>15-19</td>
<td>0.1</td>
<td>0.2</td>
<td>1.4</td>
<td>1.4</td>
<td>1.0</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>1.2</td>
<td>1.0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>20-24</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>1.0</td>
<td>3.2</td>
<td>1.9</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.6</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>25-29</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>1.9</td>
<td>2.0</td>
<td>1.0</td>
<td>0.7</td>
<td>0.5</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>30-34</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>35-39</td>
<td>0.5</td>
<td>0.6</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.7</td>
<td>0.9</td>
<td>1.1</td>
<td>0.9</td>
<td>0.7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>40-44</td>
<td>0.5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.7</td>
<td>0.9</td>
<td>1.2</td>
<td>0.8</td>
<td>0.8</td>
<td>0.4</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>45-49</td>
<td>0.4</td>
<td>0.4</td>
<td>0.6</td>
<td>1.2</td>
<td>0.6</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>50-54</td>
<td>0.2</td>
<td>0.4</td>
<td>0.8</td>
<td>1.0</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>55-59</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>60-64</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>65-69</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>70-74</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>75+</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Figure 4: Contact matrix for Italy estimated from contact survey data.
Figure 5: Contact matrix for Poland estimated from contact survey data.
Figure 6: Ridge plots of the contact functions of three European countries estimated from contact survey data.
Table 1: Fit for serological survey models.

<table>
<thead>
<tr>
<th>Country</th>
<th>Deviance</th>
<th>d.f.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>215.3</td>
<td>231</td>
<td>0.76</td>
</tr>
<tr>
<td>Italy</td>
<td>229.5</td>
<td>231</td>
<td>0.52</td>
</tr>
<tr>
<td>UK</td>
<td>289.2</td>
<td>263</td>
<td>0.13</td>
</tr>
</tbody>
</table>
Table 2: Disassortativeness measures from serological survey data for 3 European countries: absolute ($I^2$) and standardized ($I^2_s$) measures with 95% percentile bootstrap confidence intervals (CI), and global measures ($I^2_{global}$).

<table>
<thead>
<tr>
<th>Country</th>
<th>age ($I^2$)</th>
<th>activity levels ($I^2_s$)</th>
<th>global ($I^2_{global}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>112.5</td>
<td>0.21</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(63.3, 182.2)</td>
<td>(0.12, 0.34)</td>
<td>(0.00, 0.18)</td>
</tr>
<tr>
<td>Italy</td>
<td>92.1</td>
<td>0.16</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(36.6, 187.2)</td>
<td>(0.07, 0.33)</td>
<td>(0.02, 0.25)</td>
</tr>
<tr>
<td>UK</td>
<td>131.9</td>
<td>0.25</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(69.9, 196.2)</td>
<td>(0.13, 0.37)</td>
<td>(0.09, 0.26)</td>
</tr>
</tbody>
</table>
Table 3: Disassortativeness measures from contact survey data for 8 European countries: absolute ($I^2$) and standardized ($I^2_s$) measures with 95% credible intervals (CI), and global measures ($I^2_{global}$).

<table>
<thead>
<tr>
<th>Country</th>
<th>$I^2$ (95% CI)</th>
<th>$I^2_s$ (95% CI)</th>
<th>$I^2$ (95% CI)</th>
<th>$I^2_s$ (95% CI)</th>
<th>$I^2_{global}$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>234.7 (224.0, 245.9)</td>
<td>0.46 (0.44, 0.49)</td>
<td>0.52 (0.44, 0.62)</td>
<td>1.37 (0.44, 0.62)</td>
<td>0.92 (1.32, 1.44)</td>
</tr>
<tr>
<td>Finland</td>
<td>223.6 (214.7, 233.4)</td>
<td>0.45 (0.43, 0.47)</td>
<td>0.30 (0.26, 0.37)</td>
<td>1.25 (0.26, 0.37)</td>
<td>0.85 (1.21, 1.28)</td>
</tr>
<tr>
<td>Germany</td>
<td>236.1 (227.1, 245.1)</td>
<td>0.47 (0.45, 0.49)</td>
<td>0.30 (0.25, 0.36)</td>
<td>1.23 (0.25, 0.36)</td>
<td>0.85 (1.21, 1.28)</td>
</tr>
<tr>
<td>Italy</td>
<td>257.4 (246.7, 268.2)</td>
<td>0.52 (0.50, 0.55)</td>
<td>0.35 (0.29, 0.42)</td>
<td>1.28 (0.29, 0.42)</td>
<td>0.90 (1.23, 1.32)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>215.3 (207.1, 223.1)</td>
<td>0.45 (0.43, 0.47)</td>
<td>0.40 (0.35, 0.48)</td>
<td>1.30 (0.35, 0.48)</td>
<td>0.88 (1.28, 1.35)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>224.7 (206.5, 243.6)</td>
<td>0.46 (0.42, 0.50)</td>
<td>0.37 (0.26, 0.55)</td>
<td>1.30 (0.26, 0.55)</td>
<td>0.88 (1.21, 1.39)</td>
</tr>
<tr>
<td>Poland</td>
<td>205.1 (197.1, 212.6)</td>
<td>0.44 (0.42, 0.45)</td>
<td>0.44 (0.37, 0.50)</td>
<td>1.32 (0.37, 0.50)</td>
<td>0.88 (1.30, 1.37)</td>
</tr>
<tr>
<td>UK</td>
<td>241.8 (232.9, 250.8)</td>
<td>0.49 (0.47, 0.50)</td>
<td>0.31 (0.26, 0.37)</td>
<td>1.25 (0.26, 0.37)</td>
<td>0.86 (1.21, 1.30)</td>
</tr>
</tbody>
</table>