CofFEE

Centre of Full Employment and Equity

**Conceptualising and measuring spatial segregation:
the state of play**

Martin J. Watts[1]

August 2005

I   Introduction

There has been a long history of research into the manifestations of different forms of segregation, primarily in the US sociological literature. Segregation has been defined as 'the extent to which individuals of different groups occupy or experience different social environments' (Reardon and O'Sullivan (2004, p.122). The study of segregation is generally motivated by the understanding that individuals' location in social space, however defined, is linked to their access to resources, which include economic rewards, contact via social networks, social status, quality schools etc (Reardon and Firebaugh, 2002b, p.85). Academic research must not only be able to adequately document the pattern of this segregation but also test relevant theoretical relationships which will contribute to an understanding of these links.

Indexes are often employed in studies of segregation. An index provides a summary measure of the particular phenomenon under study without compromising the integrity of the data via simple aggregation. The use of an appropriate index to measure quantitative changes in segregation over time informs and provides focus for complementary forms of analysis, including case studies, other descriptive statistics and econometric analysis. The properties of the index must be well understood, so that its changes over time or differences across space can be readily interpreted.

A majority of studies have analysed occupational sex segregation and residential segregation by race and gender, often using index measurement based on two population groups. While residential segregation has an explicit spatial dimension, the treatment of the social/physical geography of segregation has, until recently, been largely neglected.

Debates in the literature over the measurement of both occupational and residential segregation have a long history which commenced with the influential paper by Duncan and Duncan (1955) that outlined the properties of five indexes and promoted the use of the Index of Dissimilarity (ID) in studies of segregation. The pre-eminent role of ID in empirical studies of race and gender segregation by occupation and residence has been periodically challenged when index 'wars' have broken out and the deficiencies of ID and /or the properties of other indexes have been outlined.

James and Taeuber (1985), White (1986), Watts (1992, 1998) and Grusky and Charles (1998) mainly reviewed the properties of aspatial binary, indexes of segregation. Massey and Denton (1988, p.309-310) argued that the urban spatial structure was complex and that residential segregation was the outcome of the interplay of social and economic processes which had five dimensions. They defined *Evenness* as the degree to which populations are distributed uniformly across areal units. *Exposure* measures the extent to which the different population groups share common areas. *Clustering* identifies the extent to which members of a minority group are located close to each other. *Concentration* refers to the degree to which a group agglomerates in space. *Centralisation* measures the extent to which a group resides close to the centre of an urban area.

While the last three measures depend on the location and size of census tracts, the traditional measures of evenness and isolation which are based on the differentiation of the areal units, fail to explicitly take into account their social and/or physical geography and hence the proximity of population groups within and between areal units. Such indexes give rise to the so called checkerboard problem (White, 1983). If each square denotes a neighbourhood (areal unit) with say a particular racial

composition of population, a rearrangement of the squares will have no impact on an aspatial measure of segregation, because interaction is not assumed to operate across neighbourhoods. On the other hand, spatial measures of segregation take explicit account of the social/physical geography, by recognising the potential for interaction between individuals belonging to different areal units based on the social/physical distance between them.

In addition, neither spatial nor aspatial indexes typically address the Modifiable Areal Unit Problem (MAUP), however, even though the former take account of the social/physical geography. The MAUP arises because the areal units which are the basis for the collection of the population and related data, are typically administratively determined and population based[2], rather than being based on meaningful social/spatial divisions. Thus individuals in different areal units may be closer to each other than those within the same areal unit, yet index calculations only take account of interaction between individuals within the same areal unit. Thus all spatial and aspatial measures which rely on population counts across specific areal units are sensitive to the chosen boundaries of those units, even though the spatial measures incorporate interactions across boundaries (Reardon and O'Sullivan, 2004, pp.123-124). Strictly the MAUP incorporates two effects, namely the sacle effect which reflects the degree of disaggregation and hence the number of areal units[3], and zoning, which reflects the choice of the boundaries for the areal units. The zoning effect occurs in the absence of the scale effect, when the number of units is fixed.

Reardon and Firebaugh (2002a) attempt to place the measurement of aspatial segregation on a stronger analytical foundation by using and extending criteria developed by James and Taeuber (1985). In addition, they generalise some of the common measures of segregation to incorporate more than two groups. There had been little attempt to overcome the inherent deficiencies of utilising pairwise comparisons of employment by race and gender in empirical studies of occupational segregation (e.g. King, 1992 and Figart and Mutari, 1993) even though Theil defined the information theory or entropy index in 1972 and generalisations of the ID index were developed by Morgan (1975) and Sakoda (1981). Also, Silber (1992), Boisso et al (1993) and Watts (1997) had explored the properties of some multi-group measures.

Watts (2005) argues that, while Reardon and Firebaugh (2002a) provide conceptual rigour to the interpretation of aspatial indexes, they fail to impose the requirement that meaningful time series or cross-section comparisons of index magnitudes be made. While acknowledging the insights of Reardon and Firebaugh (2002a), Grannis (2002) notes that the indexes under consideration were aspatial which was a major deficiency, particularly in the analysis of levels and trends in residential segregation, as well as income segregation. Grannis provides a comprehensive list of formulas for both aspatial and spatial segregation indexes.

Subsequently, drawing on criteria applied to aspatial multi-group measures of segregation by Reardon and Firebaugh (2002a), Reardon and Firebaugh (2002b) tentatively suggested a set of criteria for spatial indexes. These criteria were further developed by Reardon and O'Sullivan (2004) who used them to investigate the properties of a number of multi-group spatial indexes of segregation. Until these recent papers, the development of spatial indexes of segregation had been fragmented and lacked a coherent conceptual basis. Authors, including Wong (2003b), had attempted to address the complexities for index measurement resulting from the

spatial interdependence of areal units in close proximity to each other, but failed to recognise the need to make meaningful comparisons between index magnitudes over time or across space.

Spatial research can take three broad forms: i) a time series study based on a given social/physical geography; (ii) a cross section study across different regions; and iii) an intra-regional study based on local measures of segregation. The first two approaches lend themselves to the use of aggregate index measures, whereas intra-regional studies require local segregation indexes (Wong, 2002). In this paper we critically assess the recent developments in the measurement of aggregate spatial, multi-group indexes and local measures of segregation. We argue that the set of criteria to assess spatial measures of segregation need to be carefully reviewed and aligned with the particular objectives of the empirical research.

In the following section, we outline the concepts necessary to define multi-group, spatial segregation. In Section III we outline and critically assess the criteria advocated by researchers for the assessment of both spatial and aspatial indexes, with particular reference to the recent work of Reardon and Firebaugh (2002a) and Reardon and O'Sullivan (2004). In addition, we briefly explore local measures of segregation. In the next section, a set of indexes are assessed according to these criteria, and a decomposition procedure is outlined for time series studies. Concluding comments can be found in the final section.

II   Multi-group Segregation

We consider M population groups which are located in J discrete organisational units where the population groups could be differentiated by race and gender and/or age and the organisational units could denote occupations, ranges of income or specific locations in physical or social space, etc.

Using the notation of Reardon and Firebaugh (2002a), the total population is denoted as T and the proportion of the population in unit p which consists of group m is denoted by π_{pm} (p = 1,2,…,P), (m = 1, 2,…,M), so that

(1) $\sum_{m=1}^{M} \pi_{pm} = 1$ (p = 1,2,…,P)

The share of the total population represented by group m is given by π_m. The number of individuals located in unit p is denoted as t_p.

We now extend these concepts to make them applicable to the measurement of spatial segregation, as outlined by Reardon and O'Sullivan (2004). These P units must now be considered as points located in region, R. τ_p denotes population density at point p, and τ_{pm} denotes the density of group m at point p. The population densities per unit area are calculated by dividing the population count of the areal unit, say a census tract, by its area. Thus a common density is assigned to all points within the tract which is the traditional choroplethic method. This approach, however, is likely to lead to distinct discontinuities across boundaries of the areal units (see, for example, Holt et al, 2004).

We can write the proportion of the population at point p which consists of group m as $\pi_{pm} = \tau_{pm} / \tau_p$, which is independent of the measure of unit area. Spatial

4

segregation requires that spatial proximity be defined between all points in the region R. Reardon and O'Sullivan (2004, p.129) define $\phi(p,q)$ as a non-negative function which defines the proximity of points p and q. It has the following properties $\phi(p,q) = \phi(q,p)$ and $\phi(p,p) = \phi(q,q)$ for all $p,q \in R$, where a larger value of the function denotes closer proximity. We define the measure, Φ_p corresponding to point p as

$$(2)^4 \qquad \Phi_p = \int_{q \in R} \phi(p,q)dq$$

Some spatial measures for the local environment of point p can now be defined. While τ_p denotes the population density at point p,

$$(3) \qquad \tilde{\tau}_p = (1/\Phi_p) \int_{q \in R} \tau_q \phi(p,q)dq$$

denotes the population density in the local environment or the spatially weighted average population density.[5] Likewise $\tilde{\tau}_{pm}$ can be defined with respect to group m by replacing τ_q by τ_{qm} in the integral. These 'tilde' terms denote measures based on a spatially smoothed population surface.

Corresponding values of the group proportions $\tilde{\pi}_{pm} = \tilde{\tau}_{pm} / \tilde{\tau}_p$ can be defined which satisfy

$$(4) \qquad \sum_{m=1}^{M} \tilde{\pi}_{pm} = 1$$

In the early studies of segregation, no consideration was given to the spatial relationships. The use of areal units was merely a form of differentiation. Aspatial segregation can be understood to mean that interaction is confined to within each areal unit so that

$$(5) \qquad \phi(p,q) = 0 \ \forall p \neq q$$

and the population densities defined in (3) simplify to the areal unit densities, τ_p.

Reardon and O'Sullivan (2004, pp.130) note that the spatial proximity function $\phi(p,q)$ can take a variety of values which reflects the nature of the local environment. If proximity declines with Euclidean distance, then the environment of point p is influenced more by populations which are located close by. The spatial proximity function should identify the potential for social interaction, but, by assumption, this potential is the same for all population groups.

Reardon and O'Sullivan, 2004, p.154) acknowledge that there are other density estimation procedures including pycnophylactic (mass preserving) smoothing and dasymetric mapping (see, for example, Mennis 2003; and also Holt et al, 2004). Also smoothing can be undertaken by kernel density estimation which yields cell densities within a regular grid (Martin et al, 2000). These approaches address the MAUP

problem by representing population in small homogenous spatial units (Martin et al, 2000, p.344).

III   Criteria for Indexes of Segregation

(a) Aspatial Indexes

To place the analysis of spatial indexes into context, it is necessary to briefly examine the debate about criteria for aspatial indexes. Reardon and Firebaugh (2002a) argue that multi-group segregation indexes have been constructed in an *ad hoc* manner and are not underpinned by a consistent set of principles. They list the four criteria developed by James and Taeuber (1985) for indexes based on two population groups, namely Organisational Equivalance, Size Invariance, the Principle of Transfers and Composition Invariance.

These criteria can be defined as follows:

(i)   <u>Organisational Equivalence</u>
This criterion is satisfied if the magnitude of an index is unaffected either by the combination of two organisational units which have the same population composition or by the division of a single group of units into sub-units each with identical patterns of segregation;

(ii)   <u>Size Invariance</u>
The index magnitude is invariant when all populations are increased uniformly across the organisational units;

(iii)   <u>Transfers</u>
If an individual from group m is moved from unit i to unit j, where the proportion of persons in group m are higher in unit i than unit j ($\pi_{im} > \pi_{jm}$), then segregation is reduced.

(iv)   <u>Composition Invariance</u>
This property refers to the invariance of the index, following uniform, percentage changes in the number of members of say group m across all units, with the distributions of other groups across the units being unchanged.

Criteria (i) and (ii) are self explanatory. Watts (2005) challenges whether, in the aspatial case, the strong transfers criterion is appropriate, despite it being based on the Pigou-Dalton principle which is used in the income inequality literature (e.g. Shorrocks and Wan, 2005). The strong criterion requires that if, say, group m individuals are over (under)-represented in unit i but less over- (under-) represented in unit j, then a transfer of a group m individual from unit i (j) to unit j (i) will reduce aspatial segregation, even though the over- (under-) representation of group m in one unit has decreased but has increased in the other unit. First the impact of the transfer is to change the distribution of individuals across organisational units, which distorts the computations. Second, there is an explicit requirement that the index be non-linear in population group shares across organisational units for the strong criterion to hold. Multi-group indexes respond differently to an exchange and a transfer, so it is appropriate to add the principle of exchanges as a fifth criterion, as advocated by Reardon and Firebaugh (2002a, p.37).

(v)   Exchanges

If an individual from group m in unit p is exchanged with an individual from group n in unit q, where the proportion of persons in group m are higher in unit p than unit q ($\pi_{pm} > \pi_{qm}$), and the proportion of persons in group n is greater in unit q than unit p ($\pi_{qn} > \pi_{pn}$), then segregation is reduced.

Again the index must be non-linear for the exchange condition to hold. Linear indexes, including the ID, satisfy the weak exchange criterion whereby an exchange which leads to the convergence of group m shares across the organisational units, reduces the measure of segregation.

Empirical research typically takes the form of a time series analysis of individual countries or regions or alternatively cross-country or cross regional studies. Rigorous and meaningful time series or cross-section comparisons need to be made. Composition Invariance is designed to ensure that the index magnitude is not sensitive to the overall shares of the population groups.

Reardon and Firebaugh (2002a, p.38) argue that Composition Invariance corresponds to the margin free criterion discussed in the segregation literature by authors including Grusky and Charles (1998) and Watts (1992; 1994). However 'margin free' is understood more broadly in the extant literature. Using the terminology of Blackburn et al (1993) with respect to occupational segregation, a margin free gross index should be characterised by both Composition Invariance and (Gendered) Occupations Invariance (see also Charles and Grusky, 1995). Emerek et al (2003) and Grusky and Charles (1998), as well as Watts (1998) which is not referenced by Reardon and Firebaugh (2002a), despite appearing in the same issue of *Demograph*y, are quite clear about the need for margin-free measurement.

Thus, in addition to Composition Invariance, the magnitude of a margin free segregation index must be invariant to changes in the relative sizes of the organisational units, if the group composition of these units remains unchanged. We define this criterion as Organisational Unit Invariance.[6] Thus, if λ_p (p = 1,2,...,P) denotes the share of the overall population in the individual organisational unit p, then if, say over time, π_{pm} is constant for all p,m, but there are changes in λ_p, the magnitude of the index remains constant.[7] It is unclear why Composition Invariance would be given logical priority over Unit Invariance as a criterion by Reardon and Firebaugh.

Reardon and Firebaugh (2002a, p.38) add two further properties to the five criteria listed above, namely organisational and group decomposability.

(vi)   Additive Organisational Decomposability

If J organisation units are clustered into K (<J) clusters, then the overall segregation measure can be decomposed into the sum of independent within and between cluster components.

(vii)   Additive Group Decomposability

If M groups are clustered into N supergroups, then the chosen measure of segregation can be decomposed into the sum of independent within and between supergroup components.

Watts (2005) argues that no justification is provided by Reardon and Firebaugh (2002a) for these forms of decomposability. When occupational segregation is being considered, for example, measures of unevenness incorporate the deviation of gender and/or race shares by occupation from the benchmark shares, namely the corresponding overall population group shares. Hence a disaggregation[8] procedure which throws up separate measures of segregation within clusters of organisational units (occupations), based on the associated population shares has limited usefulness, even if it can be argued that some occupations can be grouped together due to some degree of homogeneity and hence reduced social distance based on theoretical considerations, such as patterns of promotion (Reardon and Firebaugh, 2002b, p.90).

For example, consider the highly male dominated Skilled Blue Collar occupations in a study of the evolution of sex segregation (Watts, 2003). In Australia and some European countries Blue Collar occupations are characterised by a low rate of within group segregation, because the low female employment share is being utilised as the benchmark and there is little variation in the female share of employment across these occupations. On the other hand, the between group component will pick up the disparity between the female share of employment across these occupations and the overall female share of employment. This form of disaggregation has no justification within this context.

The rejection of Additive Group Decomposability does not preclude the disaggregation of the aggregate index measure into a weighted sum of dichotomous indices, for groups of organisational units (cf. Reardon and Firebaugh, 2002a, p.54), but each computation should be based on the common benchmark, rather than being decomposed on the basis of between and within group components. In this proposed disaggregation, a reduction in segregation across a group of organisational units, associated with a reallocation of members of the population groups within those organisational units, translates directly into a reduction in the measure of segregation, but segregation cannot be reduced to zero within the group of units, unless the group population shares coincide with the overall population shares.

Such a disaggregation can assist in the identification of possible barriers to integration. Studies of occupational segregation reveal that rates of integration across occupational groups exhibit major disparities (Watts, 2003). Thus the argument that 'universal segregative and integrative forces dwarf occupation specific forces', which justifies the use of a single aggregate index magnitude, is dubious, as noted by Weeden (1998, p. 4). We reconsider the usefulness of this property in a spatial context below.

Reardon and Firebaugh (2002a) do not provide any justification for the supergroup decomposition, either, particularly when a supergroup could conceivably consist of one population group, which according to the decomposition is treated in isolation, with the between group component being an aggregate of these components across the supergroups.

Reardon and Firebaugh (2002a) ensure that the multi-group index measures under consideration lie between 0 and 1, by dividing the segregation measure by its maximum value. In the absence of Composition and Unit Invariance, this property, while constraining the index magnitude, does not simplify the interpretation of its changes magnitude through time or differences across different countries or regions.

Watts (2003, p.638) argues that no index can satisfy the first five criteria listed above, plus Organisational Unit Invariance, because the criteria of Organisational Equivalence and Unit Invariance are inconsistent. If an index is to be Unit Invariant then the weight attached to each of the P organisational units in the index calculation must be constant. Necessarily this implies that the unit weights in the index calculation must be equal ie 1/P, even though the units may differ significantly in relative size. This is inconsistent with Organisational Equivalence which permits units to be broken up with no impact on the index magnitude.

The solution to reconciling the properties of the individual indexes with the demanding and restrictive criteria for the measurement of segregation, which include both forms of invariance, but exclude transfers and the strong form of exchanges and criteria (vi) and (vii), is to develop a form of index decomposition which reveals the change (difference) in segregation, after the artifactual changes in the index, associated with changes in the population shares and the relative magnitudes of units have been removed. The choice of the index becomes less constrained, since the two invariance properties do not need to be satisfied.

In their Australian time series study of occupational sex segregation, Karmel and Maclachlan (1988) employ the IP index to which Duncan and Duncan (1955, p.211) make an oblique reference. The IP index is related to the binary ID via the population shares and has a simpler interpretation. It is the total share of the population (employment) that has to be notionally relocated between organisational units (ie. with replacement) to achieve uniform population shares across all units.

The binary ID is the IP index normalised by its corresponding maximum value. This has led to confusion about the interpretation of the ID index in the literature. While the normalisation ensures that the ID lies between zero and unity, this does not enable simple time series (or cross section) comparisons to be made between index values in the absence of the invariance properties.

Karmel and Maclachlan (1988, pp.190-191) employ an iterative procedure, based on principles developed by Deming and Stephan (1940), to transform the initial distribution of the population by organisational unit and group shares into a distribution, which has the same population across each organisational unit and the same overall group shares of the total population as the later (period 2) observation (see also Watts, 2005).

Thus, while the aggregate numbers in each population group and across the units coincide with those prevailing in period 2, there are differences in the composition of each organisational unit with respect to the population groups. Karmel and MacLachlan (1988, p.194) provide an example based on occupational gender segregation. If the data under consideration is time series, then this procedure can be used to re-specify all time series observations according to a common (period 2) base, so that they can be directly compared. By construction, this procedure eliminates the problem of the absence of margin-free properties by ensuring that the change in the index magnitude solely reflects changes in the composition of the population groups across the organisational units. This procedure is not unique, however, because there is no criterion for choosing the base period. Watts (2003) recommends the use of a simple average correlation criterion. Within this framework the approach can be extended to decompose multi-group data across organisational units (see Watts and Macphail, 2005, for a study of Canadian segregation by occupation, industry and sex).

Against this aspatial background, we now turn to the criteria for the measurement of spatial segregation indexes.

(b) Spatial Indexes

In considering the appropriate criteria for a spatial index of segregation, it is necessary to outline at the outset the type of empirical studies which index measurement will support. Rigorous comparisons of measures of segregation across different regions are virtually impossible, because aggregate measures are sensitive to Composition Invariance, but more importantly the areal units across regions will not be comparable and will be unequal in number. Time series studies are possible for regions, but often the classification of areal units changes which means that a reconciliation is required (Holt et al, 2004). Finally a cross-section analysis based on local segregation measures can be undertaken.

Drawing on Reardon and Firebaugh (2002b), Reardon and O'Sullivan (2004, pp.131-136) recast the criteria for aspatial segregation indexes into an equivalent form for aggregate spatial indexes.

(i)   Scale interpretability:
If the group proportions are the same in the local environment of each individual, then the spatial segregation index should take the value zero, since the whole region has uniform group proportions. On the other hand, a segregation index should reach its maximum value (typically normalised to equal 1), if the local environment of each individual is mono-racial. In short, if the proximity of any two members of different groups is zero, then the index takes its maximum value. Reardon and O'Sullivan (2004, p.132) note that a segregation index can assume a negative value if the population is *hyper-integrated*. This means that individuals, on average, experience greater diversity in their local environments than the diversity of the overall population, which can occur at the edges of the space.

(ii)   Arbitrary boundary independence:
A spatial segregation measure should be independent of the definitions of areal units. This demanding criterion requires that all individuals can be precisely located in space and there is a unique and exhaustive set of spatial proximities for all pairs of locations. In this case the MAUP is overcome. Reardon and O'Sullivan (2004, p.154) indicate that they are developing a set of tools which will enable the estimation of smooth population density surfaces.

(iii)   Location equivalence:
Reardon and O'Sullivan (2004, p.132) outline a spatial analogue of the organisational equivalence criterion. Assume that the local environments of two points p and q have the same population composition (i.e $\tilde{\pi}_{pm} = \tilde{\pi}_{qm}$ for all m) and the same proximity to all other points (i.e. $\phi(p,s) = \phi(q,s)$ for all $s \neq p, q$). The two points can then be treated as one point with the population density being equal to the sum of the corresponding densities associated with the two original points, and so segregation is unchanged.

(iv)   Population density invariance:
Reardon and O'Sullivan (2004, p.132) develop a spatial generalisation of the aspatial size invariance criterion (James and Taeuber 1985). If the population density of each group m at each point p, τ_{pm} is multiplied by a constant factor, then segregation is unchanged.

(v)   Composition invariance:
Reardon and O'Sullivan (2004, pp.133-134) employ a spatial analog of the James and Taeuber (1985) definition of composition invariance to maintain consistency. This states that if the number of individuals in a particular group increase uniformly across all locations within the region, and the numbers and spatial distribution of all other groups remain unchanged, then segregation is unchanged.

(vi)   Transfers:
Reardon and O'Sullivan (2004, pp.134-135) suggest a spatial analog of the Reardon and Firebaugh (2002a) multi-group transfer. If an individual of group m is transferred from point p to q, then segregation is reduced if the proportion of group m in the local environments of all points closer to p than q is greater than the proportion of group m in the local environments of all points closer to q than p.

The restrictions on the functional form of an index are compounded in the spatial case, since the impact of a transfer of individuals between p and q is influenced by the respective spatial proximity functions, $\phi(p,s)$, $\phi(q,s), \forall s$. All the indexes under consideration by Reardon and O'Sullivan fail the criterion of transfers (p.152).

(vii)   Exchanges:
Reardon and O'Sullivan (2004, p.135) identify two forms of exchange:
Exchanges (Type 1): *If an individual of group m from point p is exchanged with an individual of group n from point q, and if the proportion of group m in the local environments of all points closer to p than q is greater than the proportion of group m in the local environments of all points closer to q than p, and if the proportion of group n in the local environments of all points closer to q than p is greater than the proportion of group n in the local environments of all points closer to p than q, segregation is reduced. In simpler terms, if an exchange moves two individuals of different groups to locations where they are less likely to encounter members of their own group (and hence, more likely to encounter members of other groups), then segregation should be reduced. In the aspatial case, this reduces to the usual exchange criterion (James and Taeuber 1985; Reardon and Firebaugh 2002a).*

The authors define a second exchange criterion (p.135), which is again quoted in full:
Exchanges (Type 2): *If an individual of group m from point p is exchanged with an individual of group n from point q, and if the proportion of group m is greater than the proportion of group n in the local environments of all points closer to p than q, and if the proportion of group n is greater than the proportion of group m in the local environments of all points closer to q than p, segregation is reduced.*

Again the implicit non-linear assumption is being made here, which is required if the local proportions for the affected population groups both exceed or are less than their overall share across the region and are made more equal by the exchange (Reardon and O'Sullivan, 2004, p.135). Thus the ID and other linear indexes, at best satisfy the criteria only weakly. In addition, with the exception of the Exchange type 2, the Information Theory, Relative Diversity and Dissimilarity Indexes only satisfy the two spatial exchange criteria, if spatial symmetry is assumed (Reardon and O'Sullivan, 2004, pp. 156-157).

Spatial symmetry imposes significant constraints on the type of geographical area which can be subject to analysis. The authors note (pp.156-157) that an infinite space with a uniform population density, and spatial proximity based just on Euclidean distance will satisfy this condition but it is impossible to fulfill. The authors outline a second form of spatial geography which satisfies the condition, but it is again overly restrictive.

Finally Reardon and O'Sullivan (2004, pp. 151-152) appear to be implying that the Exchange Type 1 criterion is relevant to a particular index measure, namely social exposure, yet this criterion is not satisfied with such an index.

(viii)   Additive Spatial Decomposability
Reardon and O'Sullivan also impose additive spatial decomposability, so that the aggregation of areal units into fewer, larger spatial areas will enable total spatial segregation to be represented by the sum of within group and between group components. As noted above, Watts (2005) argues that additive organisational decomposability has little resonance within the aspatial segregation literature (cf. Reardon and Firebaugh, 2002a), despite the possibility of theoretical linkages between organisational units, since the chosen groupings of organisational units then provide the data for the local benchmarks against which the composition of each unit is compared.

Additive spatial decomposability has an even weaker conceptual foundation. By definition, spatial proximity $\phi(p,q)$ which underpins the measurement of spatial segregation captures the proximity between population groups located at different points (areal units) and hence reflects the physical or social geography. An (arbitrary) exhaustive and non-overlapping grouping of points (areal units) over continuous space would then override the degree of proximity between the areal units as measured by the spatial proximity function. Reardon and O'Sullivan (2004, pp.147-149) acknowledge this shortcoming in their disaggregation of the Information Theory (Theil) index, which has three components, including an interaction term, so that the simple disaggregation based on between and within group components does not hold.

Reardon and Firebaugh (2002b, p.90) suggest that, in both the occupational and school (aspatial) segregation literature, the organisational units are treated as socially distant from each other. While housing segregation is properly understood in a (social) spatial context, the majority of the occupational segregation studies rely on aggregate region or country level employment data, rather than employment by enterprise so that the incorporation of linkages between different

occupations may be overstated due to the spatial diffusion of the corresponding employees and the associated lack of mobility.

(ix)   Additive Grouping Decomposability
If M groups are clustered in N supergroups, then a segregation measure should be decomposable into a sum of independent within- and between-supergroup components.

No justification is given for this property.

IV   Multi-Group Spatial Indexes

a) Aggregate Measures

Reardon and O'Sullivan (2004, p.136) outline some general principles for the construction of spatial indexes. Spatial exposure will reflect the group composition of the local environment for individual members of each group, whereas spatial evenness picks up the extent of variation of the composition of the population of local environments across the region. By replacing the actual group proportions π_{pm} by the local environment measures, $\tilde{\pi}_{pm}$, Reardon and O'Sullivan (2004, pp.137-144) construct spatial analogs of the aspatial multi-group measures, developed by Reardon and Firebaugh (2002a), namely the Spatial Information Theory Index, H; Relative Diversity Index, R, the Index of Dissimilarity D, the Spatial Proximity Index, SP and Spatial Exposure Index R. They claim that the first four indexes are measures of unevenness, but no supporting algebra is provided to demonstrate that the index of clustering, SP, is a measure of unevenness.[9] We focus on the spatially weighted entropy index which is favoured by the authors and the spatial version of the generalised Index of Dissimilarity.

The entropy index at point p is defined by Reardon and O'Sullivan (2004, p.139) as

(6) $\tilde{E}_p = -\sum_{m=1}^{M} \tilde{\pi}_{pm} \log(\tilde{\pi}_{pm})$

which incorporates information about the composition of the population in the local environment of point p. The overall index can be written as:

(7) $\tilde{H} = 1 - (1/TE) \int_{p \in R} \tau_p \tilde{E}_p dp$

where E is the overall regional entropy of the total population given by

(8) $E = -\sum_{m=1}^{M} \pi_m \log(\pi_m)$

The authors (p.139) explain that the entropy index measures the extent to which local environments are less diverse than the total population of the region. If each local environment is mono-racial, then the index will take the value of unity. If there is a

spatially uniform composition, then $\tilde{E}_p = E$ and the index takes the value of zero, signifying complete integration.

The spatial dissimilarity index is defined as

$$(9) \qquad \tilde{D} = (1/2TI)\sum_{m=1}^{M} \int_{p \in R} \tau_p \left| \tilde{\pi}_{pm} - \pi_m \right| dp$$

It measures the extent of the weighted difference between the overall population composition and that of the local environments. These evenness measures all incorporate a point weight, τ_p, which ensures that they are not Spatial Unit Invariant, the spatial analogue of Organisational Unit Invariance.

The authors argue that, since the aspatial criteria, described in Reardon and Firebaugh (2002a) are special cases of the spatial criteria, the spatial measures will not meet the spatial criteria, listed above, unless the aspatial measures meet the corresponding aspatial criteria.

In Table 1, which is drawn from Reardon and O'Sullivan (2004, pp.151-152), the five indexes which they outline are judged against the criteria, to which the Unit Invariance criterion has been added. Putting aside the first four standard criteria, the indexes perform poorly against the remaining criteria. As suggested above, satisfaction of the transfer and exchange criteria is highly problematic due to the necessary restrictions on the geographical data. Neither of the invariance criteria is satisfied by any of the indexes.

In their reviews of aspatial and spatial indexes, respectively, Reardon and Firebaugh (2002a) and later Reardon and O'Sullivan (2004) bring some rigor to the evaluation process, by drawing on criteria developed by authors, including James and Taeuber (1985) and Massey and Denton (1988). There is, however, little attempt to review the relevance of the underlying criteria through engagement with both the extant theoretical literature and empirical studies that are undertaken by sociologists, geographers and economists. In particular, the results in Table 1 show that any time series analysis of spatial segregation across a region would be difficult to interpret due to the impact of changing overall population shares and areal unit or point population densities. The use of the spatial analog of a margin free index would confront the incompatibility between Locational Equivalence and Unit Invariance.

Table 1: Properties of Spatial Indexes of Segregation[a]

	Information Theory ($\tilde{H}$)	Relative Diversity ($\tilde{R}$)	Dissimilarity ($\tilde{D}$)	Spatial Proximity (SP)	Spatial Exposure ($\tilde{P}*$)	Silber/KM Index ($\tilde{I}\tilde{S}$)
Scale	YES	YES	YES		YES	YES
MAUP	YES	YES	YES	YES	YES	YES
Locational Equivalence	YES	YES	YES		YES	YES
Population Density Invariance	YES	YES	YES	YES	YES	YES
Compositional Invariance					-	
Areal Unit Invariance					-	
Transfers					-	
Exchange (Type 1)	YES[c]	YES[c]	NO[bc]		-	NO[bc]
Exchange (Type 2)	YES	YES[c]	NO[bc]		-	
Additive Spatial Decomposability	YES	YES			-	
Additive Grouping Decomposability	YES				-	

[a] Based on Reardon and O'Sullivan (2004, pp.151-152). The authors differentiate between binary and multi-group spatial indexes, but the results are identical, except that the binary Dissimilarity Index exhibits Composition Invariance.

[b] The index exhibits weak satisfaction of type 1 and type 2 exchange under this spatial specification.

[c] The spatial symmetry of the Region, R with respect to the proximity function, ϕ, is a necessary condition for this criterion to hold.

Earlier in the paper the Karmel and Maclachlan decomposition procedure for an aspatial binary segregation index, IP was outlined. Drawing on the work of Karmel and Maclachlan and Silber (1992), Watts (1997, p.470) extended the IP index to the multi-group aspatial analogue which can be written as:

$$(10) \qquad IS = (1/2T)\sum_{p=q}^{P}\sum_{m=1}^{M}\left|\tau_{pm} - \tau_p\pi_m\right|$$

where τ_{pm} denotes the number of group m at point p, and the other symbols are as defined. This aspatial index can be readily respecified as a spatial index, along the lines described by Reardon and O'Sullivan (2004, p.137).

$$(11) \qquad \tilde{IS} = (1/2T)\sum_{m=1}^{M}\int_{p\in R}\tau_p\left|\tilde{\pi}_{pm} - \pi_m\right|dp = (2I)\tilde{D}$$

The relationship between the multi-group (spatial) measures of IP and ID is via the interaction index, I, which is equivalent to the product of the overall population shares in the two group case. The spatial IS index has a simpler interpretation, namely the proportion of the total population which must relocate to achieve a uniform

composition of the population across the whole region. Its properties are also noted in Table 1

An analogous decomposition is required to address the need to obtain meaningful comparisons of regional spatial segregation over time. If the data are defined across common areal units and associated measures of proximity, then the spatial transformation of the population data is relatively straightforward and is as described earlier.

Changes over time in the boundaries of the areal units which are the basis of data collection compounds the difficulty of interpretation of changes in aggregate spatial segregation. A dasymetric approach can address the problem of inconsistent boundaries, assuming that suitable external data are available (Holt et al, 2004). Also while both centroid based approaches using kernel estimation and the dasymetric approach yield discontinuous population distributions, they both moderate the omnipresent MAUP problem.

The above observations clearly demonstrate that the comparison of measures of spatial segregation across different regions is largely meaningless. A decomposition of the type described is impossible in this case, because there are not common areal units across the regions.

(b) Local Segregation Measures

Wong (2002) explores a number of local (area based) measures of segregation which would be suited for cross-sectional analysis. First it should be noted that for any study the overall region which is being analysed must be coherent in the sense that it is a self contained area with respect to the phenomenon under study. Otherwise some local measures of segregation may be inconsistent, with the majority of measures, which could undermine a cross-section analysis but would be less significant in a study based on the aggregate segregation measure. There is no formal discussion of the requisite properties in Wong's paper, but since the population composition of each areal unit is conditioned by and contributes to the overall population composition, the invariance properties are not required. For local measures to be meaningfully compared they must be scale invariant so that a uniform increase in the density at point p of all population groups which leaves population shares unchanged at point p leaves the local segregation measure unchanged. It could be argued that scale interpretability is required, but for the comparison of local measures, this property is unnecessary. For example, Wong (2002, pp.85-87) considers the entropy index at point p (equation 6), through the incorporation of its composite population based on point p and adjacent areal units. The local measure as well as this measure standardised by its maximum diversity and also its aggregate diversity are calculated. The relative local measures are unaffected by the common denominators and it is the relative measures which are relevant for the purposes of empirical analysis. Wong (2002) also considers interaction measures but the analysis will be confined to measures of evenness.

V   Concluding Comments

The measurement of spatial segregation is contested because first there is no unanimity about its conceptualisation with writers emphasising different dimensions of segregation (White, 1986 and Massey and Denton, 1988 and Reardon and O'Sullivan, 2004). Second, differences persist with respect to both the appropriateness and importance placed on different criteria, which reflect, in part, the absence of a clear articulation as to the empirical application for the measure of spatial segregation. Third, the adoption of new measures of spatial segregation requires the parallel development of GIS software for empirical implementation, so that further developments in measurement have been slowed by absence of empirical studies. Also in the absence of a continuous population density surface, there were the characteristic spatial problems of the checkerboard and MAUP.

Thus the introduction of space has meant that the difficulties of interpretation that were found in aspatial index measures of segregation have been accentuated, as is revealed in Table 1. There is no scope for cross-section comparisons of aggregate spatial segregation measures and time series analysis of an aggregate (regional) index may require the estimation of a common discontinuous population surface for each time period and the adoption of a (Karmel/Maclachlan) decomposition procedure to enable meaningful time series comparisons. Finally, while the cross-sectional comparison of intra-regional local segregation measures is less sensitive to the choice of index, but MAUP may well undermine the interpretation of the local measures. This is inevitable if a finite number of locations underpins the research.

VI   References

Blackburn, R.M., Jarman, J. and Siltanen, J. (1993) 'The Analysis of Occupational Gender Segregation over Time and Place: Considerations of Measurement and Some New Evidence', *Work, Employment and Society*, 7(3), pp.335-362.

Boisso, D., Hayes, K., Hirschberg, J. and Silber, J. (1993) 'Occupational Segregation in the Multi-Dimensional Case: Decomposition and Tests of Significance', *Journal of Econometrics,* 61, pp.161-171.

Charles, M. and Grusky, D.B. (1995) 'Models for Describing The Underlying Structure of Sex Segregation', *American Journal of Sociology*, 100(4), pp.931-971.

Deming, W. and Stephan, F. (1940) 'On a least squares adjustment of sampled frequency table when the expected marginal totals are known', *Annals of Mathematical Statistics*, 11, pp.427-44.

Duncan, O.D. and Duncan, B. (1955) 'A Methodological Analysis of Segregation Indexes', *American Sociological Review*, 20, pp.210-217.

Emerek, R., Figueiredo, H., Gonzalez, M.P., Gonas, L. and Rubery, J. (2003) *Indicators on Sex Segregation*. DP 2003 – 02, Report for the European Commission's Expert Group on Gender and Employment, Bruxelles: European Commission.

Figart, D.M. and Mutari, E. (1993) 'Gender Segmentation of Craft Workers by Race in the 1970's and 1980's', *Review of Radical Political Economics*, 25(2), pp. 50-66.

Grannis, R. (2002) 'Discussion: Segregation Indices and Their Functional Inputs', in Stolzenberg, R.M. (ed.) *Sociological Methodology*, 32, pp.69-84, Boston, MA: Blackwell Publishing.

Grusky, D.B. and Charles, M. (1998) 'The Past, Present and Future of Sex Segregation Methodology', *Demography* 35, 497-504.

Holt, J.B., Lo, C.P. and Hodler, T.W. (2004) 'Dasymetric Estimation of Population Density and Areal Interpolation of Census Data', *Cartography and Geographic Information Science*, 31(2), pp.103-121.

James, D.R. and Taeuber, K.E. (1985) 'Measures of Segregation.' In Tuma, N.B.(ed.) n *Sociological Methodology*, 14, pp.1-32, San Francisco, CA: Jossey-Bass.

Karmel, T. and Maclachlan M. (1988) 'Occupational Sex Segregation - Increasing or Decreasing', *Economic Record* 64, pp.187-195.

King, M.C. (1992) 'Occupational Segregation by Race and Sex: 1940-88', *Monthly Labor Review*, pp.30-37.

Martin.D., Tate, N.J. and Langford, M. (2000) 'Refining Population Surface Models: Experiments with Northern Ireland Census Data', *Transactions in GIS* 4(4), pp.343-360.

Massey, D. S., and Denton, N.A. (1988) 'The Dimensions of Residential Segregation', *Social Forces* 67, pp.281–315.

Mennis, J. (2003) 'Generating Surface Models of Population Using Dasymetric Mapping', *Professional Geographer* 55, pp.31–42.

Moon Z.K. and Farmer F.L (2001) 'Population Density Surface: A New Approach to an Old Problem', *Society and Natural Resources*, 14(1), pp.39-51.

Morgan, B.S. (1975) 'The Segregation of Socioeconomic Groups in Urban Areas: A Comparative Analysis', *Urban Studies*, 12, pp.47-60.

Reardon, S. F., and Firebaugh, G. (2002a) 'Measures of Multigroup Segregation', in Stolzenberg, R.M. (ed.) *Sociological Methodology*, 32, pp.33-67, Boston, MA: Blackwell Publishing.

Reardon, S. F., and Firebaugh, G. (2002b) 'Response: Segregation and Social Distance—A Generalized Approach to Segregation Measurement', in Stolzenberg, R.M. (ed.) *Sociological Methodology*, 32, pp.85-101, Boston, MA: Blackwell Publishing.

Reardon, S.F. and O'Sullivan, D. (2004) 'Measures of Spatial Segregation', in Stolzenberg, R.M. (ed.) *Sociological Methodology*, 34, pp.121-162, Boston, MA: Blackwell Publishing.

Sakoda, J. (1981) 'A Generalized Index of Dissimilarity', *Demography*, 18, pp.245-250.

Shorrocks, A. and Wan, G. (2005) 'Spatial Decomposition of Inequality', *Journal of Economic Geography*, 5, pp.59-81.

Silber, J. (1992) 'Occupation Segregation Indices in the Multidimensional Case: A Note', *Economic Record*, 68(202), pp.276-277.

Watts, M.J. (1992) 'How Should Occupational Sex Segregation be Measured?', *Work, Employment and Society*, 6, pp.475-487.

Watts, M.J. (1997) 'Multi-Dimensional Indexes of Occupational Segregation: A Critical Assessment', *Evaluation Review* 21(4), pp.461-82.

Watts, M.J. (1998) 'Occupational Gender Segregation: Index Measurement and Econometric Modelling', *Demography*, 35, pp.489-96.

Watts, M.J. (2003) 'The Evolution of Occupational Gender Segregation in Australia: Measurement and Interpretation', *Australian Journal of Labour Economics*, 6(4), pp.631-655

Watts, M.J. (2005) 'Criteria for the Measurement of Multi-Group Segregation Indexes: A Further Examination' , *mimeo*, Centre of Full Employment and Equity (CofFEE), The University of Newcastle, NSW.

Watts, M.J. and MacPhail, F. (2005) 'Occupational Sex Segregation in Canada, 1987-2003', *mimeo*, Centre of Full Employment and Equity (CofFEE), The University of Newcastle, NSW.

Weeden, K. (1998) 'Revisiting Occupational Sex Segregation in the United States, 1910-1990: Results from a Log-Linear Approach', *Demography*, 35, pp.475-88.

White, M.J. (1983) 'The Measurement of Spatial Segregation', American Journal of Sociology, 88, pp.1008-1018.

White, M.J. (1986) 'Segregation and Diversity Measures in Population Distribution', *Population Index* 52, pp.198–221.

Wong, D.W.S. (2002) 'Modelling Local Segregation: a Spatial Interaction Approach', *Geographical and Environmental Modelling*, 6(1), pp.81-97.

Wong, D.W.S. (2003a) 'Spatial Decomposition of Segregation Indices: A Framework Toward Measuring Segregation at Multiple Levels', *Geographical Analysis*, 35(3), pp.179–194.

Wong, D.W. S. (2003b) 'Implementing Measures of Spatial Segregation in GIS.' *Computers, Environment and Urban Systems* 27, pp.53–70.

---

[1] The author is Deputy Director of the Centre of Full Employment and Equity and Associate Professor of Economics at the University of Newcastle, Australia.

[2] For example in the USA, census tracts incorporate 2500 to 8000 residents, with an average of about 4,000 residents. Given the variation in population density, the areas covered by census tracts (and other sub-areas) differ considerably.

[3] The aggregation issue arises in the measurement of occupational segregation, which while arguably aspatial, (but see Reardon and Firebaugh, 2002b, p.90), is sensitive to the degree of occupational disaggregation. Increased disaggregation unambiguously raises or leaves unchanged the index magnitude (see also Wong, 2003a), but the impact on the rate of change of the index over time is ambiguous.

[4] While integration occurs over two dimensional space Reardon and O'Sullivan (2004) adopt a single integral which is associated with integrating over points.

[5] It could be argued that the weighted population density be based on spatially weighted population counts divided by the corresponding area.

[6] It is important not to conflate measures of unevenness with some notion of social or economic progress associated with an increase of a minority (female?) share of population, which is say employed. Then Composition Invariance would not be warranted.

[7] The binary ID index is Composition Invariant, but it is not Unit Invariant. The standardized ID Index exhibits Unit Invariance, but no longer exhibits the property of Composition Invariance (Charles and Grusky, 1995, p.935).

[8] We use the term *disaggregation* rather than *decomposition* to avoid later confusion with respect to terminology.

[9] Reardon and O'Sullivan (2004, p.125) state that, in the absence of arbitrary boundaries, the major dimensions of spatial segregation can be reduced to evenness (clustering) and isolation (exposure). No formal proof is provided.