THE USE AND ABUSE OF ENTROPY BASED SEGREGATION INDICES

Martin Watts
The University of Newcastle
NSW
Australia

1. INTRODUCTION

The objective of a study of segregation is the rigorous documentation of its pattern at a point in time and its changes (differences) through time (across space). This is typically undertaken by index measurement, where the properties of the chosen index should be relevant to the research questions being investigated. The task is challenging.

Methodological and conceptual debates about segregation indexes have persisted, since the publication of the seminal article by Duncan and Duncan (1955). These index wars have been punctuated by comprehensive reviews (see James and Taeuber, 1985; White, 1985; Massey and Denton, 1988; Watts, 1998; and Reardon and Firebaugh, 2002). The focus of these debates has been mainly evenness indexes in which gender and/or race cohorts have different distributions across organisations (e.g. schools and occupations), but recently Mora and Ruiz-Castillo (2003) and Van Puyenbroeck et al (2012) examine occupational segregation with the addition of another dimension, namely educational attainment and Guinea-Martin et al. (2013) explore the impact of race. Despite the strictures of Charles and Grusky (2004), academic interchanges about the properties of indexes have continued.

An important methodological debate is whether an axiomatic approach to the design of indexes is appropriate (Hutchens, 2001, 2004; Chakravarty and Silber, 2007; Alonso-Villar and Del Río, 2010; Frankel and Volij, 2011), as opposed to less structured approaches (Flückiger and Silber, 1999; Reardon and Firebaugh, 2002; Watts, 1998; Charles and Grusky, 2004; Mora and Ruiz-Castillo, 2003, 2009, 2011). The adoption of a rigorous axiomatic approach has its pitfalls in that the demands of mathematical tractability can dictate the choice of properties for the index (for example, Hutchens, 2004), rather than the actual requirements of the research topic, which differ according to the structural features of the type of segregation being studied.

A number of conceptual issues which impact on the choice of index also remain contentious. Jerby et al. (2005a,b; 2006) argue that changing margins are integral to the process of occupational gender segregation, so that its measurement over time would also be influenced by changes in the occupational structure and changes in the gender shares of employment.[1] This view is challenged by researchers, including Watts (1998); Charles and Grusky (2004); Grusky and Levanon (2006); and Mora and Ruiz-Castillo, (2009, 2011) who advocate the derivation of margin free measures of change in segregation.

[1] The latter would reflect the take up of employment by working age women and men, as represented by their respective labour force participation rates and unemployment rates.

A second conceptual issue is (are) the form(s) of index decomposability which should be exploited to enhance the understanding of the segregation process (Reardon and Firebaugh, 2002; Mora and Ruiz Castillo, 2009, 2011). Most researchers would reject the view that 'universal segregative and integrative forces dwarf occupation specific forces', which would justify the derivation and interpretation of a single index magnitude per time period, as noted by Weeden (1998:4), but there is no unanimity as to the appropriate form of index decomposition. In this paper, to avoid confusion, *Internal Decomposability* is defined as a decomposition which is based on the extant datasets (segregation matrices). On the other hand, *External Decomposability* requires the derivation of new segregation matrices and thus is associated with some form of transformation of the data.

There has been growing support for using entropy based indexes to measure segregation in schools and occupations, in particular. In working papers and published articles over the past decade, Mora and Ruiz-Castillo, hereafter MRC, have advocated the use of the Mutual Information (MI) Index for investigating occupational segregation, after comparing its properties with those of the entropy information index, H, and the related normalized index H*, as well as non-entropy based indexes (MRC, 2009, 2011).

Empirical papers on multi-level school and occupational segregation by Reardon, Yun and McNulty Eitle (2000) and Frankel and Volij (2011) exploit *Internal Decomposability* in the form of *Weak Decomposability by Group*, using the Theil Index, H. On the other hand, MRC (2003), Van Puyenbroeck et al (2012) and Guinea-Martin et al. (2013) utilise the MI index and exploit *Strong Decomposability* in their two dimensional analyses of occupational segregation, but Van Puyenbroeck et al (2012) employ a different index decomposition to Guinea-Martin et al. (2013). However the MI index can also be locally decomposed by race and/or gender based cohorts and/or (groups of) organizations (Alonso-Villar and Del Río, 2010) but this form of *Internal Decomposition* is rejected by Reardon and Firebaugh (2002).

This paper argues that the MI index has many desirable properties which make it suited to the time series analysis of school and occupational segregation, but i) the *Strong Group Decomposability* property must be exploited with caution, because, under some circumstances, exploiting *Local Decomposability* is more useful in generating meaningful comparisons of index magnitudes; ii) margin dependence should be addressed by the application of the Deming and Stephan (DS) algorithm; and iii) the solution provided by MRC (2009, 2011) to the issue of the margin dependence of MI is flawed, and exploitation of the *Shapley decomposition* cannot usually be justified and has limited value anyway.

In Section 2 we explore the key properties of segregation indexes. We then address the reservations outlined above about the application of the MI index in Sections 3 and 4. An empirical illustration is provided in which trends in US occupational segregation by race and gender across 505 detailed occupations are computed over the period 2000-2010, using the MI index. Our results are compared to those generated to address margin dependence which are developed by MRC (2009, 2011). The three-way additive approach of Van Puyenbroeck et al (2012) is employed to examine the extent of intersectionarity with respect to race and gender (see also Guinea-Martin et al., 2013). Concluding comments complete the paper.

2. INDEX PROPERTIES

We define the segregation matrix, X as

$$X = \begin{bmatrix} t_{11} \quad t_{12}, \ldots, t_{1N} \\ t_{21} \quad t_{22}, \ldots, t_{2N} \\ \ldots, \ldots\ldots, \ldots, \ldots \\ t_{G1}\, t_{G2}, \ldots, t_{GN} \end{bmatrix} \tag{1}$$

where t_{gn} is the number of individuals from cohort g in organisation, n (g =1,2,…,G; n= 1,2,….N).

The Mutual Information Index as applied to segregation by cohort can be written:

$$MI(G,N) = \sum_{n=1}^{N}\sum_{g=1}^{G} p_{gn} \log\left(\frac{p_{gn}}{p_g p_n}\right) = \sum_{i=1}^{I} p_i \sum_{n\in C_i} (p_n / p_i) \sum_{g=1}^{G} p_{g|n} \log\left(\frac{p_{g|n}}{p_g}\right)$$
$$= \sum_{k=1}^{K} p_k \sum_{g\in k} (p_g / p_k) \sum_{n=1}^{N} p_{n|g} \log\left(\frac{p_{n|g}}{p_n}\right) \tag{2}$$

where n, g denote organisation and cohort, respectively. All the p terms denote probabilities expressed in terms of the total population, T, so, for example, p_{gn}, for is the probability that a randomly selected individual belongs to cohort g and is in organisation n.

The MI index measures evenness, with segregation being perceived as cohorts being differently distributed across organisations In the second right hand term, organisations are allocated to a series of clusters (subscript I), with p_i being the share of all individuals in the i^{th} cluster, C_i (i=1,2,…I). Segregation in the I^{th} cluster is expressed as a fraction of the associated population and weighted by its share of the total population. On the other hand, in the third term, cohorts are grouped (subscript k) and p_k is the population share of the kth group of cohorts, which could be, for example female Hispanics. MI exhibits *Transpose Invariance* (see below) so it is also an index of representation which measures the extent to which organisations have different cohort shares than the overall economy (Alonso-Villar and Del Río, 2010). Thus, in the third term, MI is expressed as the sum of cohort based measures of segregation weighted by their population shares. These two decompositions exploit the property of *Local Decomposability* (see below).

As shown in equation (2), the MI index measures the reduction in uncertainty about an individual's cohort after discovering the organization to which the individual belongs, but also measures the reduction in uncertainty about an individual's organisation after discovering the cohort to which an individual belongs.

Scale Invariance, *Symmetry in Cohorts* and *Symmetry in Organisations* are desirable properties of segregation indexes. Non-linear binary indexes, including MI, exhibit a strong version of the *Transfers* and *Exchanges* property. A priori, it is unnecessary to require the strong version of the property to hold. Also its meaning is unclear with 3 or more cohorts (Frankel and Volij, 2011), but Reardon and Firebaugh (2002) try to overcome this ambiguity.

Under the *Organisation Division Property* (ODP), the breakup of an organisation into two separate organisations cannot reduce segregation, but it remains constant if the two organisations have equal cohort compositions or one organisation has a zero population. *Organisations Equivalence* (OE) refers to index invariance when combining two or more organisations with identical cohort compositions (Frankel and Volij, 2011:6-7).[2] These two properties imply that the contribution of each organisation to the index measure is weighted by its relative population (Watts, 2003). The MI index satisfies this property.

Under *Organisational Invariance* (OI), an index, S, is invariant to changes in the numbers of individuals, when the cohort composition of each organisation remains constant. This means that

$$S(X\Gamma) = S(X) \tag{5}$$

where Γ is a N*N diagonal matrix with $\gamma_j > 0$, (j = 1,2,….N) (Watts, 2003).

However *OI* is inconsistent with *OE* (and thus ODP) because OI in combination with *Symmetry in Organisations* implies that the organisation weights in the index calculation are constant and equal to the inverse of the total number of organisations. On the other hand, OE requires that each organisation weight reflects its population.

Under the *Cohort Division Property* (CDP) the index is unchanged if a cohort is divided into two sub-cohorts and either one sub-cohort is empty or the sub-cohorts have the same distributions across organizations. The MI index satisfies the CDP.

Composition Invariance (CI) requires that the index is invariant to an equi-proportionate change in the number of cohort g (g=1,..,G) individuals in each organisation:

$$S(\Lambda X) = S(X) \tag{6}$$

where Λ is a G*G diagonal matrix, with $\lambda_g > 0$ (g =1,2,…,G). The MI index does not conform to CI. Again the *Composition Invariance* property and the corresponding *Division Property* are inconsistent because they have different implications for the index weights.

CI and OI jointly ensure that changes in segregation are not an artefact of differences (changes) in overall cohort shares and the distribution of individuals across organisations. An index conforming to both CI and OI is Margin Free (MF) in that the uniform adjustment of either row or column elements of X has no impact on the index measure.[3] Most researchers pay lip service to CI and ignore OI (see, for example, Frankel and Volij, 2011).

Watts (2003) notes that two approaches can be adopted to achieve a MF measure: (a) use an index of segregation which satisfies these and other criteria; and (b) develop a form of index decomposition which reveals the change (difference) in segregation, after the artefactual changes, associated with margin dependence, have been removed. However, a

[2] Frankel and Volij (2011:7) demonstrate that, in the binary case, *Organisational Equivalence* and the *Transfer Principle* jointly imply the *School Division Property*.
[3] Such margin-free changes in the actual segregation matrix are unlikely in practice, but the presence of these invariance properties conditions the measured changes in the index magnitudes in general.

MF index will neither satisfy *OE* (*Organisation Division Property*) nor *CDP*, so neither organisations nor cohorts would be weighted by their respective population shares. The index calculation would then be highly sensitive to the choice of the numbers of both organisations and cohorts.

An index exhibits *Strong Additive Organisational Decomposability* (SAD), since

$$S(X) = S(\tilde{X}^1 \cup \ldots \cup \tilde{X}^i \ldots \cup \tilde{X}^I) + \sum_{i=1}^{I} p_i S(i) \quad (7)$$

where S(X) denotes the Segregation index, X^i (i=1,2,…I) is the ith organisational cluster, each with G cohorts. $\tilde{X}^i$denotes the ith cluster, with all organisations being combined into one. P_i. is the population share of individuals in cluster i. $\cup$ denotes union across organisations, so that X can be denoted as

$$X = X^1 \cup \ldots \cup X^i \ldots \cup X^I \quad (8)$$

Thus overall segregation is internally decomposed into the sum of between and weighted within segregation components (Reardon and Firebaugh, 2002, Hutchens, 2004, Mora and Ruiz-Castillo, 2009, 2011 and Frankel and Volij, 2011).

Likewise, if G cohorts are clustered into K super-cohorts, then, under *Strong Additive Cohort Decomposability* (SGD), the segregation measure can be internally decomposed into independent within- and between- super-cohorts components, so that

$$S(X) = S(\tilde{X}_1 \cup \ldots \cup \tilde{X}_k \ldots \cup \tilde{X}_K) + \sum_{k=1}^{K} p_k S(\tilde{X}_k) \quad (9)$$

where p_k *(k=1,2,…K)* is the proportion of individuals in super-cohort k; and $\widetilde{X_k}$ is the configuration that results when each super-cohort, X_k is treated as a group, so that within-supergroup ethnic differences are ignored (MRC, 2011:169; Frankel and Volij, 2011:13). MI satisfied both these forms of decomposability.

A *Transpose Invariant* (TI) index can be specified both in terms of evenness and representativeness. Thus it takes the same value when the G*N segregation matrix is replaced by its transpose (Mora and Ruiz-Castillo, 2009). MI satisfies this property, as demonstrated in equation (2).

Under *Local Decomposability* (LD) an index can be expressed as the weighted sum of 'local indexes' defined across the cohorts (Alonso-Villar and Del Río, 2010). This form of decomposition identifies the contribution of each cohort to overall segregation, and over time indicates whether or not integration is occurring for that cohort in the context of the overall pattern of change in the extent of segregation.

If *Transpose Invariance* holds, then LD also applies to individual organisations. The MI index satisfies these properties (Alonso-Villar and Del Río, 2010), as shown in equation (2), and so rates of change in segregation can be computed for both clusters of organisations and groups of cohorts, if margin dependence issues can be addressed (see below).

MRC (2011:189-190) correctly argue that there are benefits of choosing an index which is not normalised to lie between zero and unity. They consider school segregation matrices of dimension 3*3 and 2*2, both of which entail no mixing of the cohorts within organisations. H and H* both exhibit normalisation (NOR), but, in contrast to MI, do not differentiate between the matrices with respect to segregation levels. More information about the cohort (school) of a student from the school attended (her cohort) from the larger matrix should attract a higher level of segregation.

A seemingly plausible justification for an index exhibiting NOR would be that the interpretation of relative index magnitudes is made easier. However if the chosen index were not MF, then cross section or time series comparisons would be meaningless, whether or not NOR holds. On the other hand, if a composition effect is calculated via the DS data transformation, then this measure is margin free, so comparison of index magnitudes is meaningful, even though the constant upper bound of the index is not unity.

Jargowsky and Kim (2009) argue that segregation and inequality can be formally linked by reference to Information Theory. 'Segregation should therefore be measured by comparing the information about inequality in the group summary data to the information about inequality in the individual-level data' (Jargowsky and Kim, 2009:6). Thus, for example, in respect of occupational gender segregation, the individual data consists of employment data by gender, undifferentiated by occupation, whereas grouped data is differentiated by both gender and occupation. Since inequality in the individual-level data represents the upper bound for the magnitude of inequality across grouped data, treating the measure of segregation as the ratio of these two measures represents normalisation. However, in line with the critique underpinning this paper, the conceptualisation of segregation is being given priority over the objectives of an empirical study of segregation which must be informed by its structural features.[4]

In summary, Frankel and Volij (2011) show that MI is the only index which, with the exception of CI, satisfies their substantive group of properties, namely *Scale Invariance, Symmetry*, *Independence*, the *Organisation Division Property* and *Group Division Property*. Also, MI exhibits Strong forms of *Group Decomposability*, *LD* and *TI.* The multi-cohort Karmel-Maclachlan Index (IS) shares many of the properties of MI.[5]

3. STRONG DECOMPOSABILITY

As noted, MI is strongly decomposable with respect to cohort groups, but also organisational clusters (MRC, 2011). Reardon et al (2000) develop a five-level geographical decomposition of metropolitan public school segregation by race, using Theil's Entropy Index, H which is weakly decomposable. Metropolitan Statistical Area (MSA) based

[4] MRC (2011) note that the strong group and organisation decomposability properties of the MI Index are incompatible with NOR (Frankel and Volij, 2011).

[5] Using time series data, Watts and Rich (1993) and Watts (1995) constructed MF measures of segregation by Occupational Group (OG), using the binary Karmel and Maclachlan index and the multi-group version, IS index, respectively, and by exploiting the Local Decomposability property and applying the DS data transformation. The major disparities in rates of change by OG may indicate discriminatory barriers.

segregation was decomposed into within- and between-school district components of segregation. Reardon et al (2000:359, 363) demonstrate that the index can be decomposed to show the nested contributions attributable to multiracial segregation between white and minority students and in turn segregation among the minority groups, namely Blacks, Hispanics and Asians.

Frankel and Volij (2011:18) use the MI Index and decompose total ethnic segregation across US public schools (2007-08) into measures at 4 geographic levels. In turn, these 4 index measures are decomposed into the nested contributions of '(1) Hispanics and non-Hispanics; (2) blacks, on the one hand, and whites and Asians on the other; and (3) whites and Asians' (Frankel and Volij, 2011:19).

Assume that there are I clusters of organisations, subscripted by i (i=1,…,I), then the Mutual Information Index (2) can be written as:

$$MI = \sum_{i=1}^{I} p_i \sum_{g=1}^{G} (p_{ig} / p_i) \log(p_{ig}/p_i p_g) + \sum_{i=1}^{I} \sum_{g=1}^{G} \sum_{n \in G_i} \frac{p_{ing}}{} \log\left(\frac{p_{ing} p_i}{p_{ig} p_{in}}\right) \quad (10)$$

where the first term denotes the between cluster measure, MI^B and the second terms, MI^W denote the within cluster component.

Researchers must be cautious, however, in using the properties of (*Strong) Additive Decomposability* to provide a more complete picture of patterns of segregation across organizational clusters and groups of cohorts and their evolution over time. This property can be exploited if the associated clusters and groups can be viewed as independent and meaningful entities.

For example, it makes sense to measure school segregation across mutually exclusive and exhaustive, catchment areas each of which is associated with a cluster of schools to which most resident children travel.[6] Putting aside parental choices about relocation to a catchment area with good schools, which is largely outside the direct influence of those schools, the benchmark cohort composition of pupils for a school in a particular catchment area, from the perspective of segregation, is given by the largely exogenous composition by cohort of the pupils living within that catchment area. In this case, the exploitation of an index which exhibits strong additive decomposability (SAD) can be justified to analyse both intra- and inter-areal segregation, although typically no adjustment is made for margin dependence, which is relevant for time series comparisons (see next section).

Thus, in the above studies, the computed measures can only be meaningfully interpreted if school districts, the basic spatial units, are largely self-contained with respect to students' school travel.[7] Also any geographical classification of organisations, such as the clustering of school districts into Core Based Statistical Areas (Frankel and Volij, 2011) should have a

[6] Thus few pupils travel to the schools from outside their catchment area and few students travel outside the catchment area to school.

[7] The growth of charter schools in the USA (and likewise academies in the UK) create a distortion with respect to the interpretation of changes in public school segregation over time, both with respect to its impact on the composition of students remaining in public schools and also the possibility of public school closures. Also the claim that school districts are self-contained catchment areas is weakened (Watts, 2013).

behavioural rather than an administrative justification. In addition, there needs to be an *a priori* rationale for the chosen sequence of nested calculations across ethnic groups.

The spatial clustering of schools is an example of a one way classification, as defined by MRC (2003). Another one-way classification is represented by 2 digit occupations and their associated clusters of 3 digit occupations. The within-cluster component of the segregation index can only be meaningfully based on the clustering of 3 digit occupations into each of the 2 digit occupations.

Consider two and three-digit occupations, where the former includes the Blue Collar (BC) cluster of occupations. Then, if black males dominate across BC occupations, as compared to their overall share of employment, the within-BC segregation measure could well be low, reflecting little variation in the respective shares of employment by gender/race cohorts across these occupations. The overall black male share of employment, which is the correct benchmark, would be much lower than its share in the BC occupations. If correct benchmarks were used, BC occupations would be highly segregated, with respect to black males and one or more other cohorts would be under-represented. Thus within-cluster measures would not be the basis for computing integration (resegregation) rates of two digit occupations, which could inform policy design. Researchers could rely instead on the separate components of the between (two-digit) group computations as a guide to possible barriers to occupational attainment, but then the variation of cohort shares of employment across the three-digit occupations would have then been suppressed. In these circumstances it makes sense to exploit the *Local Decomposability* property instead (see equation 2).

MRC (2011:165-66) interpret the decomposition of organisational clusters (equation 7) as enabling the definitive calculation of the contribution to segregation of cluster i which they define as $p_i S(X^i)$, the intra-cluster measure. The impact of the overall cohort composition of cluster i on the *inter-cluster* component of segregation is ignored, despite the overall index incorporating both components. The intra-cluster measure alone is only meaningful if each cluster can be treated separately, as described above. Also, the comparison of the separate within-cluster components of segregation is usually blighted by margin dependence, as well as unequal numbers of organizations in each cluster.

MRC (2003) interpret human capital (HK) attainment and occupational status as defining a two way classification, so that, in contrast to the one way classification, the form of the internal decomposition is not specified *a priori*. As with any study of segregation, the interpretation of the results will reflect the implicit assumptions about the nature of the segregation processes which need to be clearly articulated. For example, gender preferences concerning both HK acquisition and occupational attainment need to be juxtaposed against exclusionary practices. Also for many occupations there is not a one to one mapping with human capital acquisition.[8] However, the problematic nature of the interpretation of the decomposition of the one way classification, described above, is not apparent for a two way classification.

[8] While employment in Professional and Skilled Blue Collar occupations typically requires specific forms of HK attainment, the relationship between HK attainment and other two-digit occupations is less systematic, particularly in the presence of credentialism.

The above examples illustrate the point that the measurement and interpretation of segregation is context specific (see MRC, 2003, fn. 9), so it is questionable to claim that a universal set of properties should characterise its measurement, irrespective of the structural features of the segregation process under study.

When the appropriate benchmarks for a study of segregation, are based on the properties of the overall dataset, then an index, such as MI, which satisfies properties of *Transpose Invariance* and *Local Decomposability* can be used to compute the overall rate of segregation, expressed as weighted sum of index magnitudes defined with respect to either (groups of) cohorts or clusters of organisations, as shown in equation (2).

This form of decomposition based on cohorts has been challenged by Reardon and Firebaugh (2002:48) who argue that 'Segregation is defined by the *relationships* among the groups' distributions across organizational units—*not by the distribution across units of each group in isolation.*' (my emphasis)

As noted above, in their empirical study Reardon et al (2000) compare the distributions across schools of four ethnic groups in an arbitrary sequence of decompositions across the ethnic groups. Forms of decomposition which incorporate comparisons of some of the cohort distributions of employment are partial which multi-group indexes are designed to avoid (see, for example, King, 1992, who in a study of segregation by gender and race adopts pair-wise comparisons of employment distributions based on the binary index of dissimilarity).

Using the multi-cohort IS index, Watts (1995:2) argues that it is appropriate to 'examine each gender/race distribution of employment by occupation in the context of the overall distribution of employment by occupation', which is a coherent approach to measuring multi-racial segregation. The exploitation of *Local Decomposability* enables such calculations to be made. A similar argument can be made in respect of the use of the MI index.

Reardon and Firebaugh (2002:48-49) also allege that when this dichotomous decomposition is applied to the multi-group Index of Dissimilarity, then this index is divorced from its substantive binary interpretation, namely the proportion of individuals who would need to be exchanged between organisations to achieve zero segregation.

This is not a correct interpretation of the binary ID which is the proportion of individuals who would need to be *removed without replacement* to achieve zero segregation (Cortese et al, 1976). Also, neither ID nor the Gini index lends itself to this form of dichotomous decomposition, because each is CI in the binary case, but not CI in the multi-cohort case. CI means that a uniform change in the numbers of a particular cohort leaves the index unchanged. Thus the respective binary indexes cannot be meaningfully expressed as a weighted sum of cohort based indexes, with the weights being the cohorts' shares of the total number of individuals.

Van Puyenbroeck et al. (2012) simultaneously analyse gender segregation by both occupation (sector) and educational status by developing a three-way additive decomposition of their effects on overall segregation. Overall segregation is defined as gender segregation by occupation plus gender segregation by educational status minus an interaction term, which may be positive or negative. Using survey data from Flemish employees, they find that choice of study field has a larger effect on overall segregation than sectoral choice. Their mutual interaction is negative, indicating that sectoral segregation, although low, is still partly explained by educational choices. Using UK data, Guinea-Martin et al. (2013) measure intersectionality by investigating the joint effect of ethnicity and gender on overall segregation, but they employ a different decomposition procedure.

In empirical work conducted in Section 5, we replace educational attainment by race in applying their three-way additive approach to decomposing segregation into its components based on US data. Overall segregation, SEG, can be written as:

$$SEG = MI(G,N) - MI(G,N.R) + MI(R,N) \tag{11}$$

where R denotes race. In turn, the multivariate MI index, MI(G,N,R) is defined as

$$MI(G,N,R) = MI(R,N) - MI(R,N|G) = MI(G,N) - MI(G,N|R) \tag{12}$$

where the first conditional MI measure can be written as:

$$MI(R,N|G) = \sum_g^G \sum_r^R \sum_n^N p_{grn} \log(p_{grn} p_g / p_{gn} p_{gr}) \tag{13}$$

This analysis strictly yields just 4 summary measures as shown in equations (11) and (12), although Van Puyenbroeck et al. (2012:5-6) do report gender segregation for each level of education and sector.[9] The Local Decompositions of the levels of occupational segregation by gender and race potentially provides some deeper insights as to causes of changes, if any, in the relative magnitudes of the summary measures.

4. MARGIN FREE PROPERTY

4.1 *Introduction*

The MF property encompasses both OI and CI (equations 5 and 6), which must be satisfied simultaneously.[10] This property is contentious with claims that changing margins over time are integral to the segregation process (Chakravarty and Silber, 2007, Jerby et al, 2005a,b, 2006), while others ignore the issue (Hutchens, 2004, Jargowsky and Kim, 2009 and Frankel and Volij, 2011:21). On the other hand, Watts (1998), Charles and Grusky (2004) and MRC (2009, 2011) emphasise the importance of MF measures of changes in index magnitudes.

[9] Clearly a discussion of segregation across 46 educational types and 29 sectors based on the figures for each entity is not very useful, particularly when the relative size of the education type or sector is also relevant.

[10] Reardon and Firebaugh (2002:38) claim that CI corresponds to the MF property, but MF is understood more broadly in the literature (Emerek et al, 2003; Grusky and Charles, 1998; MRC, 2009; and Watts, 1998).

Few indexes are MF because OI is inconsistent with OE. The logarithmic index A, which has been used to measure occupational segregation (Charles and Grusky, 1995), is MF (Watts, 2003), but has attracted controversy.[11] MRC (2009) correctly note that there is some virtue in the MI index itself not satisfying either MF property, namely CI and OI, despite the objective of making rigorous MF comparisons of segregation over time.

4.2 *External Decomposition*

Studies of occupational gender segregation by, amongst others, Beller (1985), OECD (1985) and Rubery (1988) attempted to externally decompose gross (ID) index changes into meaningful components, but did not identify an MF component of change (see, for example, the critique of Rubery's structural and sex-composition effects by Watts, 1992).

MRC (2009) refer to the use by Karmel and Maclachlan (1988) of the DS data transformation to generate a new segregation matrix which maintains the original association structure between organisations and cohorts in say period t^1, but closely approximates the period t^2 marginal distributions. The convergent transformation operates through a series of uniform adjustments of the row and column elements of the t^1 segregation matrix, in turn, until margins of the data matrix correspond as closely as required to those of the period t^2 segregation matrix (see also Watts, 1998).

If $X(t^1,t^2)$ denotes the transformed segregation matrix, then the 'forward' MF component of change can be written as $S(X(t^2))-S(X(t^1,t^2))$. The remaining component of the gross change $S(X(t^1,t^2))-S(X(t^1))$ can be decomposed into terms based on the changes of each margins plus an interaction term. [12]

The data must be defined across a consistent set of organisations. The DS data transformation can be increasingly used for cross-national comparisons of occupational segregation with the greater availability of cross-national datasets using a common occupational classification (see Section 5 for an empirical application).

4.3 *Internal Decomposition*

MRC (2009, 2011) acknowledge that entropy based indexes, MI, H and H*, are not MF. MRC (2011:184-185) refer to two forms of temporal decomposition of changes in MI that are outlined in detail in MRC (2009). They concede that their internal decompositions do not simultaneously address both MF dimensions – 'there are ways to decompose pair-wise comparisons using the MI index to isolate either an invariant 1 or an invariant 2 term'.

The first decomposition yields three terms which can be written as:

$$\Delta M = \Delta N(\pi_g) + \Delta SC_g(P_n) + \Delta EG(\pi_g) \quad (10)$$

where

[11] This index is undefined for zero values of the arguments, and exhibits extreme sensitivity for small occupations (Watts, 1998; 2003; Jerby et al, 2005a,b, 2006, but see Grusky and Levanon, 2006).

[12] A symmetric decomposition can be calculated via averaging over the forward and backward components, with the latter based on $S(X(t^2,t^1)$. However, if a consistent time series is required which is associated with a common base, then the series should be $X(t^i,t^b)$ (i=1,2,…,T) and t^b is the base observation (see below).

$$\Delta N(\pi_g) = \sum_{g=1}^{G} \pi_g \sum_{n=1}^{N} \left(p_{n|g}\,(t^2)\log\left(p_{n|g}(t^2)\right) - p_{n|g}(t^1)\log\left(p_{n|g}(t^1)\right)\right)$$
$$\Delta SC_g = T^c(t^2) - T^c(t^1) \qquad (11)$$
$$\Delta EG(\pi_g) = \sum_{t=t^1,t^2} (-1)^{1(t=t^1)} \left\{ \sum_{g=1}^{G} (p_{g.}(t) - \pi_g) \sum_{n=1}^{N} p_{n|g}\,(t)\log\left(p_{n|g}(t)\right) \right\}$$

where $T^c(t)$=$\sum_{n=1}^{N}\{p_{.n} \log(\frac{1}{p_{.n}})\}$ (MRC, 2009: 45-46). ΔSC_g denotes the change in entropy, associated with changes in organisation margins and $\Delta EG(\Pi_g)$, measures the change in the index associated with changes in cohort margins, $P_{g.}$, and arbitrary margins, Π_g.

MRC (2009) consider a value for Π_g of $P_{g.}(t^1)$. Thus, if the segregation matrices at t^1 and t^2 exhibited the same organisation and cohort margins, then $\Delta SC_g(P_n) = \Delta EG(\pi_g) = 0$, and $\Delta N(\Pi_g)$ would measure the total change in the index, which, by definition, would represent a MF computation. However this is a somewhat unusual scenario and essentially assumes away the problems of MF measurement arising from changing margins. $\Delta N(\Pi_g)$ also satisfies I1, so that if $P_{g.}$ is subject to change, but the conditions satisfying the Composition Invariance property are met, i.e. $X(t^2)=\Lambda X(t^1)$, where Λ is a G*G diagonal matrix, then $\Delta N(\Pi_g)$ is zero, because the conditional distribution, $P_{n|g}$ is unchanged.

MRC are conflating the required properties of an index for MF computation, irrespective of changes in organisation and cohort margins, with the alternative of constructing or assuming the presence of a segregation data matrix with unchanged margins. The former requires both CI and OI, but one component of the decomposition, $\Delta N(\Pi_g)$, only satisfies CI. The latter is addressed via the heroic assumption of no changes in the margins so that the other two terms of the decomposition take the value of zero.

MRC (2009:46-47) derive a second decomposition, based on changes in cohort entropy and the conditional distribution, $P_{g|n}$ in which the equivalent terms ΔEG^n and $\Delta SC(\Pi^n)$ are zero when cohort and organisation margins stay constant, respectively. Also $\Delta N(\Pi^n)$ is zero, under the conditions satisfying Organisational Invariance (I2), but again, there is a conflation of objectives. Neither of the decompositions yields an unambiguous measure of the MF change in the index magnitude.

The lack of clarity which is noted above, over the interpretation of $\Delta N(\Pi_g)$, when margins are changing, is addressed, when MRC (2009) consider the work of Deutsch et al (2006) who combine the DS data transformation with the concept of Shapley value. MRC argue that there is a close relationship between these two approaches.

As noted above, using the DS data transformation the 'forward' MF component of change can be written as $S(X(t^2))-S(X(t^1,t^2)$. The marginal distributions based on $X(t^2)$ and $X(t^1,t^2)$, are equal, since the DS transformation of $X(t^1)$ is designed to yield a MF comparison with $X(t^2)$. The MRC approach is premised on the questionable assumption that the conditional distributions, $P_{n|g}$ based on $X(t^1)$ and the one based on $X(t^1,t^2)$, 'closely resemble' each other (MRC, 2009:48). They demonstrate that under these restrictive assumptions, the MF change in the index from t^1 to t^2, can be written as $\Delta N(P_{g.}(t^1))$ and similarly the total margin dependent component can be calculated, since the index magnitude corresponding to $X(t^1)$ can be readily computed (MRC, 2009:48-49). However there is also an I2 based decomposition, which under appropriate assumptions about particular conditional distributions closely resembling each other, yields a forwards MF change equal to $\Delta N(P_{.n}(t^1))$. [13]

Further the calculation of these MF effects do not require the use of the DS data transformation, since the internal decomposition of the index change between MF and non-MF components is not reliant on calculations associated with either $X(t^1,t^2)$ or $X(t^2,t^1)$.

Three issues are raised by this approximation. First, does empirical evidence support the assumption of unchanged conditional distributions, following the application of the DS data transformations? This question will be left to Section 5, but, in their empirical work, MRC (2011) make no attempt to check whether this key assumption has empirical support and, at best, this form of internal decomposition is only suited to MI. Further, the only way to check these crucial assumptions about the conditional distributions is to undertake the DS data transformation, so the approximations proposed by MRC (2009) are then redundant.

Second, it is unclear what MRC (2011:182-186) are claiming about the properties of MI, as opposed to the deficiencies of H and H*, because the abstract notes that '[A]pplied researchers *may* do better using the MI index than using either H or H* in two circumstances: ...(2) if they are interested in a margin-free measurement of segregation changes' (MRC, 2011:159). The criticism of the H, H* indexes then only has merit if the authors can show how they can achieve MF measurement of the change in segregation based on the MI index, yet at best they can construct two different decompositions, neither of which is robust. No further detail is provided in Mora and Ruiz-Castillo (2011).[14]

Third, but less important, is whether it is useful to compute the symmetric gross change in the index and its decomposition, via backwards and forwards effects. If a time series analysis is being conducted, then the construction of a MF annual series requires a constant base year. Then each annual observation can be compared to any other. This requirement rules out the use of both forward and backward calculations.[15]

---

[13] MRC (2009:48-49) treat the MF change between t^1 and t^2 as symmetric, so the DS transformation is also applied to $X(t^2)$. The symmetric MF change in the index is $MF(t^1,t^2)=0.5*(S(X(t^2))-S(X(t^1,t^2)) + S(X(t^2,t^1)-S(X(t^1)))$. MRC (2009) have to assume that the conditional distributions, $P_{n|g}$, based on $X(t^2)$ and $X(t^2,t^1)$, also 'closely resemble' each other. The symmetric MF component based on the I1 decomposition takes the form:
$MF(t^1,t^2) = 0.5 * (\Delta N(P_{g.}(t^1)) + \Delta N(P_{g.}(t^2)))$.

[14] Mora and Ruiz-Castillo (2011:182-184) examine the evolution of the US public school enrolments over the period 1989-2005 and later compute the two net changes in segregation, based on the two decompositions.

[15] Time series data such as real GDP require a base year, which is subject to periodic change too.

5. US OCCUPATIONAL SEGREGATION, 2000 – 2010

5.1 *Results*

Gross and MF values of the MI index were computed for unpublished annual Bureau of Labor Statistics employment data (2000-2010)[16] defined across eight gender/race cohorts, namely male and female White, Black, Asian and Hispanic, respectively and 505 detailed occupations based on the 2002 Census occupational classification. These occupations are divided into four occupational groups (OGs), namely Managerial (52 occupations), Professional (123), Clerical, Sales and Service (131) and Blue Collar (199).[17]

Some occupations have zero employment for a particular cohort. These occupations could be suppressed by combining them with others, but this is somewhat arbitrary. Instead we replaced each zero by unity which indicates employment of 1,000 for the cohort. Given that OE holds, this adjustment would have minimal effect on the index magnitudes.

The variation in the numbers of occupations in each group means that absolute comparisons between index magnitudes for the OGs should be made cautiously. However rates of change in the index magnitudes can be compared. For simplicity the base year was chosen to be 2010 for the two indexes, so that the critical assumptions about the equality of conditional distributions could be readily examined.

Table 1 shows toal segregation and its division into 3 components as defined by Van Puyenbroeck et al. (2012). It is clear that gender segregation dominates occupational segregation by the 4 race categories. Also these contributions to overall segregation are relatively independent as shown by the small interaction term. There is minimal change in overall segregation over the sample period.

TABLE 1

US Occupation Segregation by Gender and Race:

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
SEG	0.261	0.262	0.265	0.265	0.252	0.268	0.269	0.271	0.269	0.264	0.264
MI(G,N)	0.211	0.212	0.213	0.213	0.199	0.215	0.213	0.216	0.214	0.211	0.209
MI(R,N)	0.049	0.050	0.051	0.052	0.036	0.053	0.056	0.055	0.055	0.054	0.055
-I(G,R,N)	0.001	0.000	0.001	0.001	0.018	0.000	-0.001	0.000	0.000	-0.001	0.000

Source: Unpublished Bureau of Labor Statistics data.
Notes: Derivations by author.

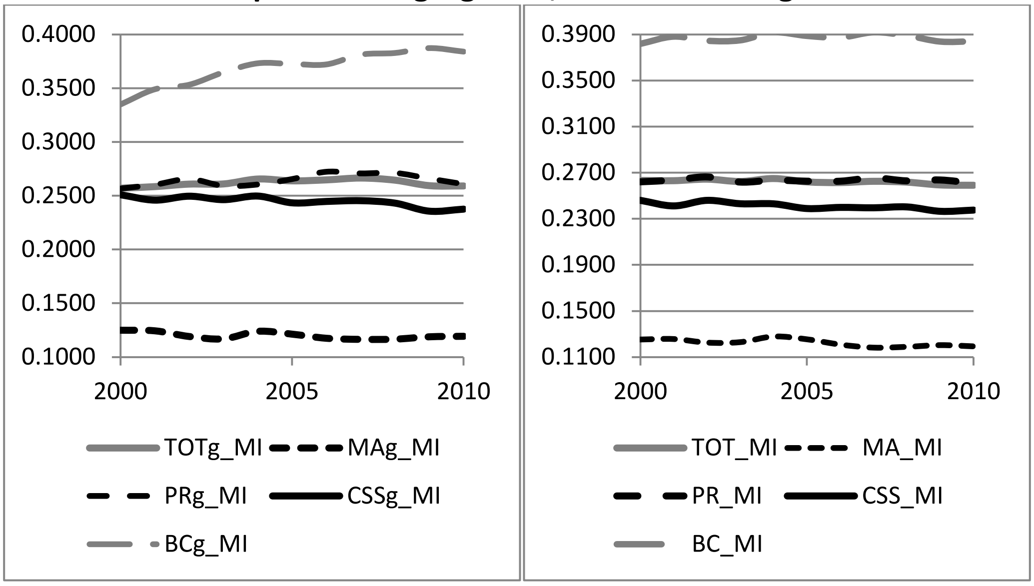
[16] Subsequent annual data are incompatible with the 2000-2010 data. Although the names of the broad- and intermediate-level occupational groups in the 2010 Census remained the same, some detailed occupations were re-classified between the broader groups, under the new 2010 Standard Occupational Classification.

[17] In a US study Watts (1995) divided Blue Collar occupations between Skilled and Unskilled, but the 2002 occupational classification does not lend itself to this division.

The remaining index calculations are based on the 505 occupations, divided into 4 Occupational clusters and 8 gender/race cohorts, so that gender and race are no longer treated separately. This means that the different experiences of say Black Women and Men can be readily established.

Figure 1 shows the gross and MF indexes. Summary statistics are shown in Table 2. The overall change in segregation was small for both measures. The evidence for the Professional Cluster was mixed. The measures exhibited similar trends with declines for Clerical, Sales and Service and Managerial occupations and a rise for Blue Collar occupations.

FIGURE 1
US Occupational Segregation, Gross and Margin Free Index



Source: See Table 1. Author's own calculations.
Notes: TOT, MA, PR, CSS and BC denote Total, Managerial, Professional, Clerical, Sales and Service and Blue Collar occupations. Subscript g denotes gross index measures. The absence of a subscript denotes MF calculations which are based on each Occupational Cluster separately.

The MF calculations led to more modest changes, particularly for BC occupations. The rates of index change by OG for the two indexes are quite similar when the MF transformation is applied to each OG separately. These trends may suggest the presence of exclusionary employment practices against specific cohorts within particular organisations. Further analysis of the results is beyond the scope of this paper.

Examination of Table 3 reveals disparate rates of integration (resegregation) for the gender/race cohorts, with more extreme values typically being associated with the gross index values. There was little change for White and Black Women. There was strong evidence of the integration of Hispanic Women but Asian Women experienced higher segregation. With the exception of Asian men there was little change in the incidence of male segregation.

TABLE 2
Summary Statistics for Gross & MF Indexes by Occupational Group

	TOT	MA	PR	CSS	BC	TOT	MA	PR	CSS	BC
	Gross MI index					Margin Free MI Index				
Growth (%)	0.82	-4.52	1.59	-5.34	14.69	-0.68	-10.2	0.33	-3.54	4.33
						-1.59	-4.74	-0.50	-3.44	0.57

Source: See Table 1. Author's own calculations.
Notes: The first set of MF index growth rates is based on the DS data transformation applied to all occupations. The second set is based on the transformation being applied to each OG, separately.

TABLE 3
Summary Statistics for Gross & MF Indexes by Gender/Race Cohorts

	TOT	WM	WW	BM	BW	AM	AW	HM	HW
	Gross MI Index								
Growth (%)	0.82	-1.55	-1.40	-0.94	1.55	13.00	13.65	-1.83	-6.53
	Margin Free MI Index								
Growth (%)	-0.68	-2.65	0.22	-2.06	0.15	7.12	11.02	-0.40	-5.23

Source: See Table 1. Author's own calculations.
Notes: WM, WW, BM, BW, AM, AW, HM, HW denote White men, White women, Black men, Black women, Asian men, Asian women, Hispanic men and Hispanic women, respectively.

5.2 *MRC Approximation*

Conditional probabilities p(n|g) were calculated for the 8 gender/race cohorts based on the employment data for 2000, X(2000), and the DS transformed segregation matrix X(2000, 2010) with 2010 treated as the base year. Absolute differences in the corresponding probabilities across the cohorts were summed for each of the 505 occupations. This yielded total cohort based differences ranging from 0.187 to 0.225. Thus the assumption that the conditional distribution changes very little when the DS transformation is applied is rejected. The gross and MF index changes for total employment over this 11 year period are relatively modest, otherwise larger differences in the conditional probabilities would have been expected.[18]

The approximation used by MRC (2009), was applied to the calculation of the forward MF effect, using each year as a base year for all years of observations to avoid any claims of data selectivity. This procedure generated 11 consistent sets of time series data in the sense that each time series was based on data exhibiting the same margins.

The estimated change of the MF index magnitude over the sample period ranged from declines of 5.18% to 6.00% across the time series data based on different base years. On the other hand the equivalent calculation using the DS data transformation yields MF index declines ranging from 0.19% to 0.77%. This is a very marked disparity, which means that the approximation should not be utilised. Also, as noted, it is necessary to calculate the robust

[18] The DS transformation was also applied to the 2010 data to equalise the margins to those corresponding to 2000, i.e. X(2010, 2000). The summed absolute differences in the conditional probabilities ranged from 0.189 to 0.226 for the eight cohorts.

MF index changes using the DS data transformation in order to assess whether the MRC approximation is accurate, which defeats the purpose of using the approximation.

The Shapley decomposition entails the computation of the MF index change, via the DS data transformation, but the interaction effect is incorporated into the components of the gross index change, associated with the changing gender and occupation margins (see MRC, 2011: Given the focus on the MF change in segregation research, little is gained by the computation of the Shapley decomposition.

6. CONCLUSION

The Mutual Information index has a number of desirable properties, including *Transpose Invariance* and *Local Decomposability*. Also the index has not been normalised. In addition, MI exhibits *Strong Cohort* and *Organisational Decomposability*. Ironically it does not exhibit *Composition Invariance* or *Organisation Invariance*, yet Margin Free changes are typically required in time series research.

This paper has argued first that the strong decomposability properties must be used with considerable caution, so rigorous insights can be drawn. In some circumstances the exploitation of *Local Decomposability* can be justified, instead, depending on the structure of segregation being analysed and the type of research questions being asked. Second, MF measures of change require the use of the DS data transformation to exploit the equalisation of margins. The internal decomposition of time series changes in the MI index outlined in Mora and Ruiz-Castillo (2009, 2011) requires the untenable assumption that sets of conditional probabilities are equal. The problematic nature of the MRC approximation was illustrated by the application of these techniques to unpublished annual Bureau of Labor Statistics employment data (2000-2010), by occupation, gender and race.

The empirical work drew on the three-way additive decomposition designed by Van Puyenbroeck et al. (2012) to explore the degree of intersectionality and found that there was little overlap in the effects of gender and race on occupational segregation. The remaining empirical work demonstrated the insights gained from exploiting the MI properties of Transpose Invariance and Local Decomposability to establish Margin Free time series of segregation by Occupational Cluster and also race/gender cohorts, which can inform the design of policy..

7. REFERENCES

Alonso-Villar, O. and Del Río, C. 2010. “Local versus overall segregation measures.” *Mathematical Social Sciences* 60: 30-38.
Beller, A. 1985. “Changes in the Sex Composition of US Occupations.” *Journal of Human Resources* 235-250.
Chakravarty, S.R. and Silber, J. 2007. “A Generalised Index of Employment Segregation.” *Mathematical Social Sciences* 53: 185-195.
Charles, M. and Grusky, D.B. 1995. “Models for Describing The Underlying Structure of Sex Segregation.” *American Journal of Sociology* 1004:931-971.

Charles, M. and Grusky, D.B. 2004. “*Occupational Ghettos: The Worldwide Segregation of Women and Men*, Stanford: Stanford University Press.
Cortese, C.F, Frank, R.F. and Cohen, J. 1976. “Further Considerations on the Methodological Analysis of Segregations Indexes.” American Sociological Review 41:630-637.
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: 427-44.
Deutsch, J. Flückiger, Y. and Silber, J. 2009. “Analyzing Changes in Occupational Segregation: The Case of Switzerland (1970-2000).” in Y. Flückiger, J. Silber and S. Reardon (eds.), *Research on Economic Inequality,* Volume 17, *Occupational and Residential Segregation*, Emerald Books, 173-204.
Duncan, O.D. and Duncan, B. 1955. “A Methodological Analysis of Segregation Indexes.” *American Sociological Review* 20:210-217.
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.
Flückiger, Y. and Silber, J. 1999. *The Measurement of Segregation in the Labour Force*. Heidelberg Physica-Verlag.
Frankel, D. and Volij, O. 2011. “Measuring School Segregation,” *Journal of Economic Theory* 146:1-38.
Grusky, D.B. and Charles, M. 1998. “The Past, Present and Future of Sex Segregation Methodology.” *Demography* 35:497-504.
Grusky, D.B. and Levanon, A. 2006. “Describing Occupational Segregation in Sparse and Incomplete Arrays.” *Sociological Methods & Research* 34(4):554-572.
Gunea-Martin, D. Mora, R. and Ruiz-Castillo, J. (2013) ‘The joint effect of ethnicity and gender on occupational segregation. An approach based on the Mutual information Index’ Economic Series, Working Paper 11-40, Universidad Carlos III de Madrid, May.
Hutchens, R. 2001. “Numerical Measures of Segregation: Desirable Properties and Their Implications.” *Mathematical Social Sciences* 42:13—29.
Hutchens, R. 2004. “One Measure of Segregation.” *International Economic Review* 45(2):555-578.
James, D.R. and Taeuber, K.E. 1985. “Measures of Segregation.” In Tuma, N.B.(ed.) n *Sociological Methodology*, 14:1-32, San Francisco, CA: Jossey-Bass.
Jargowsky, P. and Kim, J. 2009. “The information theory of segregation: Uniting segregation and inequality in a common framework.” In Y. Flückiger, J. Silber and S. Reardon (eds.), *Research on Economic Inequality, Vol. 17, Occupational and Residential Segregation*, Emerald Books, 3-31.
Jerby, I., Semyonov, M. and Lewin-Epstein, N. 2005a. “Capturing Gender-Based Microsegregation: A Modified Ratio Index for Comparative Analyses.” *Sociological Methods & Research* 34(1):122-136.
Jerby, I., Semyonov, M. and Lewin-Epstein, N. 2005b. “A Two Dimensional Index for Comparative Analyses of Gender Segregation Regimes: Revisiting the Margin Free Concept.” *mimeo*, The Horowitz Research Institute on Science and Economy, Tel-Aviv University, Israel.
Jerby, I., Semyonov, M. and Lewin-Epstein, N. 2006. “On Measures of Gender Occupational Segregation: Statistical and Conceptual Considerations (a response to Gusky and Levanon).” *Sociological Methods & Research* 34(4):573-586.

Karmel, T. and Maclachlan M. 1988. “Occupational Sex Segregation - Increasing or Decreasing.” *Economic Record* 64:187-195.
King, M.C. 1992. “Occupational Segregation by Race and Sex: 1940-88.” *Monthly Labor Review* 30-37.
Massey, D. S. and Denton N. A. 1988. “The Dimensions of Residential Segregation.” *Social Forces* 67(2): 281-315.
Mora, R. and Ruiz-Castillo, J. 2003. “Additively Decomposable Segregation Indices. The Case of Gender Segregation by Occupations and Human Capital Levels in Spain.” *Journal of Economic Inequality* 1:147-179.
Mora, R. and Ruiz-Castillo, J. 2009. “The Invariance Properties of the Mutual Information Index of Multigroup Segregation.” In Y. Flückiger, J. Silber and S. Reardon (eds.), *Research on Economic Inequality*, 17, Occupational and Residential Segregation, Emerald Books, 33-53.
Mora, R. and Ruiz-Castillo, J. 2011. “Entropy-Based Segregation Indices.” *Sociological Methodology* 41:159-194.
OECD (1985) *The Integration of Women into the Economy*. Paris, OECD.
Reardon, S. F., Yun, J.T. and McNulty Eitle, T. 2000. “The changing structure of school segregration: Measurement and evidence of multiracial metropolitan area school segregation, 1989-1995.” *Demography* 37(3):351-364.
Reardon, S. F., and Firebaugh, G. 2002. “Measures of Multigroup Segregation.” In Stolzenberg, R.M. (ed.) *Sociological Methodology*, 32, 33-67, Boston, MA: Blackwell Publishing.
Rubery, J. (ed) 1988. *Women and Recession*, London: Routledge Kegan Paul.
Van Puyenbroeck, T. De Bruyne, K. and Sels, L. 2012. “More than ‘Mutual Information’: Educational and sectoral gender segregation and their interaction on the Flemish labor market”, *Labour Economics* 19: 1–8.
Watts, M.J. 1992. “How Should Occupational Segregation be Measured?” *Work, Employment and Society* 6(3):475-87.
Watts, M.J. 1995. “Divergent Trends in Gender Segregation by Occupation in the USA: 1970-92.” *Journal of Post Keynesian Economics* 17(3): 357-379.
Watts, M. J. 1995. “Trends in Occupational Segregation by Race and Gender, 1983-92: A Multi-Dimensional Approach.” *Review of Radical Political Economics* 27(4):1-36.
Watts, M.J. 1998. “Occupational Gender Segregation: Index Measurement and Econometric Modelling.” *Demography*, 35(4):489-96.
Watts, M.J. 2003. “The Evolution of Occupational Gender Segregation in Australia: Measurement and Interpretation.” *Australian Journal of Labour Economics*, 6(4):631-655.
Watts, M.J. 2013. “Socioeconomic segregation in UK (secondary) schools: are index measures still useful?” *Environment and Planning A* 45(7):1528-1535.
Watts, M.J. and Rich, J. 1993. “Occupational Sex Segregation in Britain, 1979-89: The Persistence of Sexual Stereotyping.” *Cambridge Journal of Economics* 17(2):159-177.
Weeden, K. 1998. “Revisiting Occupational Sex Segregation in the United States, 1910-1990: Results from a Log-Linear Approach.” *Demography* 35(4):475-88.
White, M.J. 1985. “Segregation and Diversity Measures in Population Distribution.” *Population Index* 52(2):198-221.