Assessing Organizational Diversity with the Smith and Wilson Indices

Assessing Organizational Diversity with the Smith and Wilson Indices

By

Salomón Alcocer Guajardo

Cambridge Scholars Publishing

Assessing Organizational Diversity with the Smith and Wilson Indices

Series: Assessing Diversity in Nonprofit, Private, and Public Organizations

By Salomón Alcocer Guajardo

This book first published 2023

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-2858-8
ISBN (13): 978-1-5275-2858-1

This book is dedicated to Zawar Hussain, Akbar Ali Khan, James Lee Ray, J. David Singer, Benjamin Smith, J. Bastow Wilson, and Allen R. Wilcox for creating the diversity indices presented in this book and for contributing to the quantitative measurement of diversity.

Contents

EXHIBITS

NYC DEPARTMENTS

NYC Department	Acronym
Administration for Children's Services	ACS
Board of Corrections	BOC
Board of Election	BOE
Borough President-Bronx	BP-BX
Borough President-Brooklyn	BP-BK
Borough President-Manhattan	BP-MAN
Borough President-Queens	BP-QNS
Borough President-Staten Island	BP-SI
Business Integrity Commission	BIC
Campaign Finance Board	CFB
City Commission on Human Rights	CCHR
Civilian Complaint Review Board	CCRB
Conflicts of Interest Board	COIB
Department for the Aging	DFTA
Department of Buildings	DOB
Department of City Planning	DCP
Department of Citywide Administrative Services	DCAS
Department of Consumer Affairs	DCA
Department of Correction	DOC
Department of Cultural Affairs	DCLA
Department of Design & Construction	DDC
Department of Education	DOE
Department of Environment Protection	DEP
Department of Finance	DOF
Department of Health/Mental Hygiene	DOHMH

Department of Homeless Services	DHS
Department of Info Tech & Telecomm	DOITT
Department of Investigation	DOI
Department of Parks & Recreation	PARKS
Department of Probation	DOP
Department of Records & Information Service	DORIS
Department of Sanitation	DSNY
Department of Small Business Services	SBS
Department of Transportation	DOT
Department of Youth & Community Development	DYCD
District Attorney - Bronx County	DA-BX
District Attorney - Kings County	DA-BK
District Attorney - Manhattan	DA-MAN
District Attorney - Queens County	DA-QNS
District Attorney - Richmond County	DA-SI
District Attorney – Special Narcotics	DA-NARC
Equal Employment Practices Commission	EEPC
Financial Information Services Agency	FISA
Fire Department	FDNY
Housing Preservation & Development	HPD
Human Resources Administration / Social Services	HRA
Independent Budget Office	IBO
Landmarks Preservation Committee	LPC
Law Department	LAW
MAYORALTY	MAYORALTY
Municipal Water Finance Authority	MWFA
New York City Council	COUNCIL
New York City Fire Pension Fund	FDNYPF
New York City Police Pension Fund	NYCPPF
New York City Tax Commission	NYCTAX

NYC Civil Service Commission	NYCCSC
NYC Employees Retirement System	NYCERS
NYC Health + Hospitals	NYCHH
NYC Housing Authority	NYCHA
Office of Administrative Trials & Hearings	OATH
Office of Collective Bargaining	OCB
Office of Emergency Management	NYCEM (OEM)
Office of Payroll Administration	OPA
Office of the Actuary	ACTUARY
Office of the City Clerk	CLERK
Office of the Comptroller	COMPTROLLER
Office of the Public Advocate (PA)	PA
Offices of the Public Administrators	PUBADMIN
Police Department	NYPD
School Construction Authority	SCA
Taxi & Limousine Commission	TLC
Teachers Retirement System	TRS

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to Courtney Dixon, Sophie Edminson, Amanda Millar, Mhairi Nicol, Adam Rummens, and the production staff of Cambridge Scholars Publishing for making this book a reality. I also would like to thank Andre Hall, Tiffany Nichole Peña, and Steve Santiago for their immeasurable and unending encouragement and support. This book would not be possible without them.

ABOUT THE AUTHOR

Salomón Alcocer Guajardo, Ph.D., is a management consultant and researcher. He is a former associate professor of nonprofit and public management. Dr. Guajardo has held executive budgeting and finance positions in municipal government. He graduated from the University of California, Los Angeles (UCLA) with a Bachelor of Arts in political science. Dr. Guajardo received graduate degrees in public management and educational psychology from the University of Pittsburgh. He received a doctorate degree in public policy research and analysis from the Graduate School of Public and International Affairs, University of Pittsburgh. His research on compensation, diversity, and organizational efficiency has been published in leading peer-reviewed journals.

PREFACE

As presented in the companion book *Assessing Organizational Diversity with the Simpson Index* (Guajardo, 2023a), Simpson's (1949) index of diversity ($1 - \sum p^2$) is used frequently to assess demographic (or social) diversity in nonprofit, private, and public organizations. Since its development, Simpson's diversity index has been modified to calculate alterative measures of heterogeneity. For instance, Hussain and Khan (2019) developed a diversity index of evenness based partially on Simpson's index. The Hussain-Khan index is expressed as follows: $D_{HK} = \frac{1}{n-1}\left(\frac{1}{f_m} - 1\right)$, where n represents the number of different categories (or classifications) and f_m represents the category (or classification) with the largest percent. Ray and Singer (1973) also modified Simpson's index of diversity to develop an index of concentration. When the Ray and Singer index is subtracted from 1, it yields a diversity coefficient. The Ray-Singer index of diversity is represented as follows: $D_{RS} = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{1 - \frac{1}{n}}}$, where n represents the number of different categories (or classifications) and $\sum p^2$ represents Simpson's dominance index. In addition to these Simpson-based diversity indices of evenness, Smith and Wilson (1996) developed several indices of diversity based on $\sum p^2$. For purposes of this book, the following Smith-Wilson indices of diversity are discussed:

- $D_{SW1} = \frac{1 - \sum p^2}{1 - \frac{1}{n}}$, where n represents the number of different categories (or classifications) and $\sum p^2$ represents Simpson's dominance index;
- $D_{SW2} = \frac{\left(\frac{1}{\sum P^2}\right)}{n}$, where n represents the number of different categories (or classifications) and $\sum p^2$ represents Simpson's dominance index; and,
- $D_{SW3} = \frac{\ln \sum p^2}{\ln n}$, where n represents the number of different categories (or classifications), $\sum p^2$ represents Simpson's dominance index, and ln is the natural logarithm of a number.

Wilcox (1967) also modified Simpson's index of diversity to develop the following index: $D_W = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{\left(\frac{n-1}{n}\right)}}$, where n represents the number of different categories (or classifications) and $\sum p^2$ represents Simpson's dominance index. Although the indices are based on $1 - \sum p^2$ and use n to calculate evenness scores, the indices produce diversity scores which differ in size.

This book focuses on the assessment of Simpson-based indices of evenness when they are applied to demographic employment data to obtain measures of evenness. Consistent with the preceding companion books (Guajardo, 2023a, 2023b, and 2023c), this book addresses fundamental analytical and measurement issues and questions that arise when Simpson-based indices of evenness are applied to demographic and employment data to obtain measures of heterogeneity. The issues and questions addressed in this book include the following:

- How is measurement bias addressed by a particular diversity index?
- How is the number of categories used for a demographic (or social) characteristic addressed by a particular diversity index?
- What are the statistical properties of a distribution of scores of a particular diversity index when it is applied to demographic and employment data?
- What is the appropriate statistical method to use based on the distribution of scores obtained by a particular diversity index?
- What is the maximum value of diversity that is obtainable by a particular diversity index?

These issues are addressed throughout this book because little empirical research has been devoted to examining the adaptation and use of diversity indices to measure and analyze demographic (or social) diversity in organizations. Although the issues and questions addressed in this and its companion books are fundamental to carrying out empirical research, practitioners and researchers alike often ignore or take the analytical (or measurement) issues and questions for granted.

As stated in preceding companion books (Guajardo, 2023a, 2023b, and 2023c), the book series consists of 9 books. They are the following:

- *Assessing Organizational Diversity with the Simpson Index* applies the Simpson diversity index to demographic and employment data reported by New York City (NYC) departments for fiscal year 2019. This book focuses on the application and analysis of Simpson diversity formulas for calculating biased and unbiased measures of demographic heterogeneity.

- *Assessing Organizational Diversity with the Shannon Index* applies the Shannon diversity index to the same demographic and employment data used in the first book. This book focuses exclusively on the application and analysis of Shannon diversity formulas for calculating biased and unbiased measures of demographic heterogeneity.
- *Assessing Organizational Diversity with the Heip index* applies the Heip, Sheldon, and other Shannon-based diversity indices to the data used in the first and second books. The Heip, the Sheldon, and the other Shannon-based diversity indices presented in the book are modifications of the Shannon index of diversity. From a statistical standpoint, the Heip and Sheldon indices possess statistical properties that are superior to the original Shannon index. Like the first and second books, this book focuses on the application and analysis of the indices in terms of measuring demographic heterogeneity in organizations.
- *Assessing Organizational Diversity with the Smith and Wilson Indices* applies the Smith and Wilson (SW) indices to the same data used in the previous companion books. In addition to applying the SW indices, other Simpson-based indices such as the Ray and Singer (RS) index of concentration are presented in the book. The SW and RS indices are modifications of the Simpson ($D = 1- \sum p^2$) diversity index and assess demographic heterogeneity as well. This book applies Simpson-based indices to the same data used in previous books to measure demographic heterogeneity in organizations.
- *Assessing Organizational Diversity with the McIntosh Index* applies the McIntosh evenness index to the same demographic and employment data used in the preceding companion books. This book focuses on the analysis of diversity scores obtained by the McIntosh index. Because the index incorporates the number of groups used to categorize a demographic (or social) characteristic of interest and the size of the workforce simultaneously, the diversity scores contain less measurement bias and have a greater degree of compatibility in comparison to the diversity indices covered in preceding companion books.
- *Assessing Organizational Diversity with the Index of Qualitative Variation* (IQV) applies the Mueller and Schuessler IQV to the same

demographic and employment data used in the previous companion books. Because the IQV is not invariant to ordering sequences, this book focuses on the application and analysis of heterogeneity scores obtained from the different ordering sequences of the data. Like the McIntosh evenness index presented in the 5th book, the IQV incorporates jointly the number of groups used in the categorization of the demographic (or social) characteristic of interest and the size of the workforce.

- *Assessing the Validity of Diversity Indices* compares the indices used in each book jointly and uses factor analysis to determine whether they assess the same (or different) aspects of demographic (or social) diversity. Pearson pairwise correlation analyses are also performed to assess the statistical associations amongst the diversity indices. Statistical analyses for equality of means are performed as well.
- *Assessing Organizational Diversity with Quantile Regression* applies quantile regression analysis to each of the diversity indices presented throughout the book series. This book performs quantile regression analyses at the 25th, 50th, 75th, and 90th percentiles for age, ethnic, and gender diversity.
- *Assessing Organizational Diversity with Structural Equation Modeling* (SEM) focuses exclusively on causal modeling. This book focuses on the development and analysis of a SEM for specific diversity indices discussed in the book series. In so doing, the analyses treat age, ethnic, and gender diversity as intervening (or mediating) variables of organizational performance.

For purposes of continuity and compatibility, each diversity index is subjected to the same statistical analyses. The IQV, McIntosh evenness, Shannon, Simpson, and SW indices are of special focus in this book series because they have been used in previous research on demographic (or social) diversity in nonprofit, private, or public organizations.

This book series is written for practitioners and researchers in human resources and other fields that are interested in measuring and analyzing demographic, occupational, or social heterogeneity in organizations. The purpose of the book series is to addresses measurement and analytical issues that practitioners and researchers alike are likely to face when they apply a particular diversity index to demographic and employment data provided by a nonprofit, private, or public organization. As such, this book series should serve as a reference for selecting the diversity index that is best suited for measuring and analyzing heterogeneity in an organizational setting. This book series also should serve as a reference for selecting the statistical method that is best suited for analyzing the distribution of scores obtained by the diversity index of choice.

References

Guajardo, Salomón A. 2023a. *Assessing organizational diversity with the Simpson index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023b. *Assessing organizational diversity with the Shannon index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023c. *Assessing organizational diversity with the Heip index*. UK: Cambridge Scholars Publishing.

Hussain, Zawar and Khan, Akbar Ali. 2019. "A new index for measuring evenness". *Communications in Statistics - Theory and Methods*, Vol 48: 354 – 367.

Ray, James Lee and Singer, J. David. 1973. "Measuring the concentration of power in the international system". *Sociological Methods and Research*, Vol. 1: 403 – 437.

Simpson, Edward Hugh. 1949. "Measurement of diversity". *Nature*, Vol. 163: 688.

Smith, Benjamin and Wilson, J. Bastow. 1996. "A consumer's guide to evenness indices". *Oikos*, Vol. 76: 70 – 82.

Wilcox, Allen. R. 1967. *Indices of Qualitative Variation*. Oak Ridge, TN: Oak Ridge National Laboratory, U.S. Atomic Energy Commission.

CHAPTER 1

INTRODUCTION

As stated in previous companion books (Guajardo, 2023a, 2023b, and 2023c), researchers have used diversity (or *integration*) indices to assess the level of demographic (or social) *heterogeneity* in nonprofit, private, or public organizations since the early 1970s (e.g., Akram, Abrar ul Haq, Natarajan, and Chellakan, 2020; Boehm, Kunze, and Bruch, 2014; Choi, 2010; Gazley, Chang, and Bingham, 2010; Grabosky and Rosenbloom, 1975; Guajardo, 2014; Moon and Christensen, 2020; Nachmias and Rosenbloom, 1973). Of the plethora of indices of diversity that have been developed to assess heterogeneity (or variation), researchers use the Simpson (1949) and Shannon (1948) indices the most frequently. For the most part, the Simpson and Shannon indices have been applied to aggregate demographic employment data to measure age, ethnic, or gender heterogeneity. More recently, diversity indices have been used to assess concepts such as educational and occupational diversity. In most of the previous studies, workforce diversity has served as a *dependent variable*. More recent studies, however, have treated workforce diversity as an *independent variable* which influences organizational performance (e.g., Gazley, Chang, and Bingham, 2010; Khan, Khan, and Senturk, 2019; Lee-Kuen, Sok-Gee, and Zainudin, 2017; Pitts, 2005). Consistent with the preceding companion books, this book takes the position that workforce diversity such as age, ethnic, and gender heterogeneity is an *intervening variable* that influences organizational performance (e.g., Guajardo, 2014; Pitts, 2006).

Simpson-based indices of diversity and evenness

Several Simpson-based indices of *evenness* have been developed since Simpson (1949) created the index of diversity (i.e., $S = 1 - \sum p^2$). They include the following:

- Hussain and Khan (2019) index of evenness: $D_{HK} = \frac{1}{n-1}\left(\frac{1}{f_m} - 1\right)$

- Ray and Singer (1973) index of evenness: $D_{RS} = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{1 - \frac{1}{n}}}$
- Smith and Wilson (1996) indices of evenness:
  $D_{SW1} = \frac{1 - \sum p^2}{1 - \frac{1}{n}}$
  $D_{SW2} = \frac{\left(\frac{1}{\sum P^2}\right)}{n}$
  $D_{SW3} = \frac{\ln \sum p^2}{\ln n}$
- Wilcox (1967) index of evenness: $D_W = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{\left(\frac{n-1}{n}\right)}}$

Except for the D_{HK} index, the evenness diversity indices use the Simpson dominance index ($\sum p^2$) to calculate a score of how well employees are distributed across the demographic (or social) characteristic of interest such as age, ethnicity, and gender.

As discussed in *Assessing Organizational Diversity with the Simpson Index* (Guajardo, 2023a), the following formula is used to obtain an evenness (or standardized) Simpson diversity score: $s_E = \frac{1 - \sum p^2}{\left(\frac{n-1}{n}\right)}$, where *n* is the number of distinct groups in the sample (e.g., n = 3), *p* is the percent of the total workforce for each group in the sample, and p^2 is the product of squaring the percent for each group. The distribution of Simpson evenness scores ranges from 0 to 1. An evenness score of 0 indicates that the workforce is concentrated in one group of the demographic (or social) characteristic of interest. An evenness score of 1 indicates that the workforce is distributed evenly across each group of the demographic (or social) characteristic of interest. As a standardized score, the index measures the proportion of the *empirical maximum value* (EMV) that is attained in terms of the demographic (or social) characteristic of interest.

Briefly, Hussein and Khan (2019) developed an evenness index of diversity ($D_{HK} = \frac{1}{n-1}\left(\frac{1}{f_m} - 1\right)$) based on the number (or frequency) of individuals with different characteristics in a community. Although Hussein and Khan (2019) did not discuss the inclusion of categories with missing data explicitly, they included such categories when applying their index. In developing their index, however, they wanted to enhance the precision of measuring diversity and evenness. As developed by Hussein and Khan

(2019), the evenness index ranges from 0 to 1. When the index is applied to organizations, a score of 0 indicates a lack of evenness due to employees being concentrated in one group. A score of 1 indicates that employees are distributed evenly across each group.

Ray and Singer (1973) proposed an index of concentration ($C_{RS}= \sqrt{\frac{\Sigma p^2 - \frac{1}{n}}{1-\frac{1}{n}}}$) to measure the aggregation of political power amongst groups or states. Regardless of the size of the group or population (N), C_{RS} yields a distribution of scores ranging from 0 to 1 (Ray and Singer, 1973; Taagepera and Ray, 1977). According to Ray and Singer (1973), a score of 0 indicates a lack of concentration of political power among the groups (or states) under analysis. A score of 1 indicates that political power is concentrated in one group (Ray and Singer, 1973). Although the scores of C_{RS} range from 0 to 1, the scores fluctuate as the number of groups (or states) in the analysis changes (see Exhibit 1-1).

Exhibit 1-1. Measurement of political power index

States	Percent (p)	p^2
A	0.90	0.810
B	0.06	0.004
C	0.04	0.002
Total	1.00	0.815
n = 3		
	C_{RS} =	0.850

States	Percent (p)	p^2
A	0.90	0.810
B	0.06	0.004
C	0.04	0.002
A	0.00	0.000
D	0.00	0.000
E	0.00	0.000
Total	1.00	0.815
n = 5		
	C_{RS} =	0.877

Source: Ray, James Lee and Singer, J. David. 1973. "Measuring the concentration of power in the international system". *Sociological Methods & Research*, Vol. 1: 403-437.

As Exhibit 1-1 illustrates, each group (or state) is included in the calculation of the C_{RS} score regardless of the share of political power. Stated differently, C_{RS} includes each case regardless of whether data exists for the group. When the index is applied to organizations, C_{RS} maintains the comparability of the diversity scores amongst the organizations being analyzed due to including each group. Exhibit 1-2 illustrates how C_{RS} yields comparable diversity scores for age when some of the age groups lack employment data. As will be discussed in Chapter 5, several indices yield incompatible evenness diversity scores amongst NYC departments due to excluding age groups with missing employment data when the heterogeneity coefficients are calculated.