Jasso, Guillermina

Working Paper

Linking individuals and societies

IZA Discussion Papers, No. 4288

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Jasso, Guillermina (2009) : Linking individuals and societies, IZA Discussion Papers, No. 4288, Institute for the Study of Labor (IZA), Bonn, https://nbn-resolving.de/urn:nbn:de:101:1-20090821100

This Version is available at:
https://hdl.handle.net/10419/36021

DISCUSSION PAPER SERIES

IZA DP No. 4288

Linking Individuals and Societies

Guillermina Jasso

July 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Linking Individuals and Societies

Guillermina Jasso
New York University
and IZA

Discussion Paper No. 4288
July 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4288
July 2009

ABSTRACT

Linking Individuals and Societies*

How do individuals shape societies? How do societies shape individuals? This paper develops a framework for studying the connections between micro and macro phenomena. The framework builds on two ingredients widely used in social science − population and variable. Starting with the simplest case of one population and one variable, we systematically introduce additional variables and additional populations. This approach enables simple and natural introduction and exposition of such operations as pooling, matching, regression, hierarchical and multilevel modeling, calculating summary measures, finding the distribution of a function of random variables, and choosing between two or more distributions. To illustrate the procedures we draw on problems from a variety of topical domains in social science, including an extended illustration focused on residential racial segregation. Three useful features of the framework are: First, similarities in the mathematical structure underlying distinct substantive questions, spanning different levels of aggregation and different substantive domains, become apparent. Second, links between distinct methodological procedures and operations become apparent. Third, the framework has a potential for growth, as new models and operations become incorporated into the framework.

JEL Classification: C02, C16, D31, D6, D7, D8, J71

Keywords: micro-macro link, justice, comparison, status, power, identity, happiness, personal quantitative characteristics, personal qualitative characteristics, deductive theory, probability distributions, lognormal distribution, Pareto distribution, power-function distribution, rectangular distribution, exponential distribution, normal distribution, pooling, matching, multilevel modeling, inequality, race, segregation

Corresponding author:

Guillermina Jasso
Department of Sociology
New York University
295 Lafayette Street, 4th Floor
New York, NY 10012-9605
USA
E-mail: gj1@nyu.edu

* This research was supported in part by the U.S. National Science Foundation under Grant No. SBR-9321019 and partly carried out while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences in 1999-2000. Early versions of portions of this paper were presented at the annual meeting of the American Sociological Association, San Francisco, CA, 1989, the Seventh Annual Group Processes Conference, Los Angeles, CA, 1994, and the annual meeting of the Mathematical Association of America, Washington, DC, January 2009. I am grateful to participants at those meetings and especially to John Angle, Joseph Berger, Barbara Meeker, Eugene Johnsen, Samuel Kotz, Geoffrey Tootell, Murray Webster, the anonymous referees, and the Editor for many valuable comments and suggestions. I also gratefully acknowledge the intellectual and financial support of New York University. This paper is forthcoming in *Journal of Mathematical Sociology*.

[G]overnments vary as the dispositions of men vary. . . . . [T]here must be as many of one as there are of the other. . . . [I]f the constitutions of States are five, the disposition sof individual minds will also be five.

Plato

Republic, Book VIII

Different men seek after happiness in different ways and by different means, and so make for themselves different modes of life and forms of government.

Aristotle

Politics, Book VII

The central problem of social science remains the one posed, in his own language and in his own era, by Hobbes: How does the behavior of individuals create the characteristics of groups? . . . [T]he central problem is . . . the demonstration of how the general principles, exemplified in the behavior of many men and groups, combine over time to generate, maintain, and eventually change the more enduring social phenomena.

George C. Homans

The Nature of Social Science, 1967

1. INTRODUCTION

Individuals and societies are the central actors of social science. By turns they shape each other, merge harmoniously, and burst apart. Individuals create some of the characteristics of societies, and societies create some of the characteristics of individuals, each imprinting on the other crucial aspects of its own character.

Specifying these multiple and bidirectional relationships in a fruitful way is not easy, and

it is useful to explore a variety of approaches, including analytic, computational, and ethnographic approaches. This paper follows an analytic approach, examining the connections between phenomena at the level of the person and phenomena at the various levels of aggregation, that is, in groups of all sizes as well as subgroups and collections of groups and including emergent properties of groups and subgroups.

The starting idea for this paper was the realization that the underlying mathematical structure of many models and theories which connect micro and macro phenomena can be described in a simple way. As in diagraming sentences, the skeletons reveal a familiar set of ingredients and a variety of operations by which variables that inhere in individuals -- micro variables -- are linked to variables that inhere in populations -- macro variables.

From there it is straightforward to develop a framework that systematically incorporates the basic ingredients and operations. Attentiveness to these ingredients leads in a natural way to methods of analysis which not only illuminate the connections between micro and macro phenomena but also have the potential for generating new questions and seeing both substantive and methodological problems in a new light.

The common set of ingredients may be summarized as follows:

1. Individuals have quantitative characteristics -- such as beauty, intelligence, and wealth -- and qualitative characteristics -- such as gender, ethnicity, and place of residence.

2. Personal qualitative characteristics generate populations, for example, all persons of the same gender or of the same ethnicity or living in the same place.

3. Individuals may be characterized by their magnitudes on quantitative characteristics and, against the backdrop of the population, by their ranks on these variables.

4. Populations generate subpopulations.

5. Populations combine to form superpopulations.

6. Populations, subpopulations, and superpopulations may be described by parameters of the distributions of the individual-level variables and, in the case of superpopulations, of the population-level variables and of their functions.

7. The propositions of social science assert associations between the variables, including associations between variables defined at the same level of aggregation and associations between variables defined at different levels, for example, individual-level and population-level variables.

These ingredients and the framework developed in this paper are grounded in basic ideas from calculus (which distinguishes between functions of one variable and functions of several variables), probability theory (which distinguishes between univariate and multivariate distributions), and demography (which distinguishes between one and several populations), and thus these disciplines provide tools and results which make obvious some necessary relationships and enable derivation of many other results.

Using as building-blocks the basic ideas of population and variable, we begin constructing the framework by considering the case where there is only one population and one variable. Next we systematically add variables and populations until we reach the general case of several populations and several variables. At each step we review some of the models and methods that are pertinent, and, indeed, which stimulated formulation of the framework. This approach enables simple and natural introduction and exposition of such operations as pooling, matching, calculating summary measures, regression, hierarchical and multilevel modeling, finding the distribution of a function of random variables, and choosing between two or more distributions. Illustrations of the procedures are drawn from a variety of topical domains in social science, including the theories and models which led to the framework.

Three useful features of the framework are: First, similarities in the mathematical structure underlying distinct substantive questions, spanning different levels of aggregation and different substantive domains, become apparent. Second, links between distinct methodological procedures and operations become apparent. Third, the framework has a potential for growth, as new models and operations become incorporated into the framework.

However, intense focus on the framework and the underlying mathematical structures – no matter how salutary -- may obscure the substantive richness and the promise for deeper

understanding of micro-macro connections that only theory can provide. Accordingly, we present a final theoretical illustration focusing on residential racial segregation and based on a hypothetico-deductive theory which spans micro and macro worlds. The starting postulates in the theory are Popperian (1963:245) "guesses" about human nature, so that empirical analysis of their many implications and predictions yields knowledge about human nature.

The paper is organized as follows: Section 2 provides exposition of the framework. The theoretical illustration is presented in Section 3. A short note concludes the paper.

2. FRAMEWORK FOR STUDYING THE CONNECTIONS BETWEEN MICRO AND MACRO VARIABLES

2.1. The Simplest Case: One Population and One Variable

We begin with the fundamental concepts -- population and variable -- and show how in the simplest case of one population and one variable it is possible to generate one new micro variable and many macro variables as well as a subpopulation structure with its own new set of micro and macro variables.

First, however, we note that the very idea of a population presupposes a personal characteristic. Whatever the population may be -- a birth cohort, residents of a country or other geopolitical entity, members of a club -- it is defined by a personal characteristic, such as year of birth, country of citizenship, etc. Accordingly, because we wish to start with a population, we take as given a personal characteristic and we do not count it. When we say we begin with the simplest case, that of one population and one variable, we mean one variable beyond the one that defines the population.

For convenience, we define micro variable as a characteristic of a person and macro variable as a characteristic of an aggregate of two or more persons.[1]

---

[1] This usage is not universal. For example, in economics the term "micro" refers to decisionmaking units, which need not be persons, but may be large aggregates such as corporations; and the term "macro" refers to aggregates of decisionmaking units.

Generating Macro Variables from One Population and One Micro Quantitative Variable. Suppose that in a population of persons we know the values of a particular personal quantitative characteristic -- say, income. An operation familiar to all readers is that of calculating a set of summary measures of the distribution of that characteristic in that population. This set includes the mean, standard deviation, range, etc. These summary measures are themselves values of new macro variables. Thus, if all we had were one population and one micro variable, we could nonetheless generate a set of new measures, each a value of a new macro variable, defined on the distribution of the original micro variable. For example, the micro variable "income" generates a large set of macro variables, including "mean income," "income inequality," "minimum income," and "maximum income."

Generating Macro Variables from One Population and One Micro Qualitative Variable. Similarly, if the one micro variable is qualitative, it gives rise to a new set of macro variables -- the proportions in each category of the qualitative variable. For example, the micro qualitative variable “mother tongue” generates the new macro variables, “proportion whose mother tongue is Hindi,” “proportion whose mother tongue is Amharic,” proportion whose mother tongue is Spanish,” and so on.

Note that if we had a collection of populations like the original population -- call such a collection of populations a superpopulation – then it would be possible to investigate the distributions of these new macro variables. Just as "income" and “mother tongue” describe the units of the original population (viz., the persons), "income inequality," “proportion whose mother tongue is Hindi,” and “proportion whose mother tongue is English” describe the units of the superpopulation (viz., the original populations).

Generating a Subpopulation Structure. The distribution of one individual-level characteristic in a population also generates a subpopulation structure (or even several subpopulation structures) and thus generates for the population an additional set of values of macro variables.

If the characteristic is a qualitative variable, then the subpopulation structure is inherent

in the categories of the variable; for example, the population "U.S. citizen" and the variable "sex" jointly generate the two subpopulations "male U.S. citizen" and "female U.S. citizen." Note that qualitative characteristics generate both new macro variables – as in "proportion whose mother tongue is Arabic" – and a subpopulation structure, including, for example, the subpopulation consisting of all persons whose mother tongue is Arabic.

If the characteristic is a quantitative variable, then two kinds of subpopulation structures can be generated. The first is a subpopulation based on ranks -- formally, censored subdistributions. For example, the distribution of the variable "income" in any population can lead to the subpopulation structure, "the top half and the bottom half of earners". The second is a subpopulation based on values of the variables -- formally, truncated subdistributions – as in "the group with incomes below $100,000 and the group with incomes above $100,000."[2]

In both the censored and truncated cases, many subpopulation structures can be generated. For example, "top half and bottom half" is not the only possible set of censored subdistributions; one can generate any number of nonoverlapping subdistributions, and these may or may not be equally-sized. To illustrate, a campaign to solicit donations may send different kinds of letters and appeals to each of four groups based on previous donations: the top 5 percent, the next 25 percent, and the two halves of the bottom 70 percent. Similarly, in the case of truncated subdistributions, one can generate any number of nonoverlapping subdistributions, and the endpoints can be any numbers. To illustrate, the fundraising campaign of the previous example may target separately three groups, those with incomes above ten million dollars, those

---

[2] The distinction between censoring and truncation is by now standard (Gibbons 1988:355; Kotz, Johnson, and Read (1982a:396). The term "censoring" refers to selection of the units in a subdistribution by their ranks or percentage (or probability) points in the parent distribution; truncation refer to selection of the units in a subdistribution by values of the variate. Thus, the truncation point is the value x separating the subdistributions; the censoring point is the percentage point α separating the subdistributions. For example, the subpopulations with incomes less than $35,000 or greater than $90,000 each form a truncated subdistribution; the top 5 percent and the bottom 10 percent of the population each form a censored subdistribution.

with incomes between one million and ten million dollars, and everyone else.

Generating Macro Variables from a Subpopulation Structure. Qualitative characteristics automatically generate both new macro variables – proportions variables – and a new subpopulation structure. Quantitative characteristics, via the operations of censoring and truncation, generate new macro variables. The population can now be described in terms of "proportion male," "proportion with incomes below $100,000", etc.

Generating SubMacro Variables. In the case of a quantitative variable, the subdistributions in each of the subpopulations give rise to new macro variables such as "mean income of the top half" or "dispersion among the group with incomes above ten million dollars." Formally, the new macro variables are summary measures describing the subdistributions. For convenience, let us adopt the convention of referring to a macro variable defined on a subpopulation as a submacro variable.

Generating New Micro Variables. As noted above, a single micro quantitative variable automatically gives rise to a new quantitative variable, namely, the variable formed by the individuals' ranks on the values of the original variable. Thus, each individual not only has a measurement on the original micro variable but, along with it, a measurement on the rank variable. For example, an individual not only has an income amount; he/she also has an income rank. Further, whenever a subpopulation structure is generated from a quantitative variable, new rank variables are created in each of the subpopulations. Examples of such submicro variables include "percentile rank within the top truncated subdistribution."

Additionally, there exists another procedure by which it is possible to generate new micro variables from this simplest case of one population and one micro variable, provided again that the micro variable is quantitative. This procedure requires that the distribution of the one micro variable be mathematically specified and, in so doing, makes allusions, as it were, to other members of the distributional family. To the extent that these other members of the distributional family may be thought of as other variables, the procedure has left the realm of one-population/one-variable. Accordingly, we defer description to section 2.2.2.4 below.

Nonetheless, it should be noted that the procedure is applicable to any theoretical analysis or modeling exercise in the one-population/one-variable case; and hence it would appear that this simplest case can generate new micro variables beyond the rank variables.

Thus, even in the simplest of all social-science cases -- where a single micro variable is observed in a single population -- a focus on the building blocks of population and variable immediately yields (i) values of macro variables describing the population, (ii) a subpopulation structure, (iii) values of submacro variables describing the subpopulations, and (iv) in the case of a quantitative micro variable, a new micro variable defined on the ranks, further new submicro variables defined on the ranks within subpopulations, and possibly additional new rank and nonrank micro variables defined by the procedure to be described in section 2.2.2.4.

2.2. Introducing More Variables to a Population with One Quantitative Variable

2.2.1. Introducing A Single New Quantitative Variable

We continue with a single population with one quantitative variable, and begin by introducing a single new quantitative variable. Now that we have two quantitative variables, there are (at least) three new kinds of models for linking individuals and populations that we can investigate; these are the regression model, the change-of-variable model, and the binary-choice model.

2.2.1.1. The Simple Regression Model

Whenever there are two quantitative micro variables, assessment of their association generates a new set of macro variables. The population can now be described by the parameters of the regression line – intercept and slope – magnitude of linear correlation, various sums of squares, and so on.

2.2.1.2. The Change-of-Variable Model: Finding the Distribution of a Function of a Random Variable

The general problem in the change-of-variable model is to establish the connections among three things, (i) the relationship between two variables, (ii) the distribution of one

variable, and (iii) the distribution of the second variable.[3] The problem is usually stated in the form: Given the function $y = h(x)$ and given the distribution of X, what is the distribution of Y?

The change-of-variable problem arises in the study of diverse phenomena. For example, in the study of distributive justice it arises in the investigation of two distinct questions. The first is the question posed by Brickman, Folger, Goode, and Schul (1981): Suppose that each individual member of a society is judged to be perfectly justly rewarded in the distribution of a socially valued good; does it necessarily follow that the observer making the judgment will also judge the resultant probability distribution of that good to be a just distribution? The second is the question posed by Jasso (1980): Suppose that the sense of being justly or unjustly treated in the distribution of a valued good is a function of the actual reward (and a constant just reward), what will the distribution of the sense of being justly or unjustly treated look like in the population? As stated, both of these questions are simplest-case (bivariate) versions of more general questions which are inherently multivariate; in this section we examine the bivariate versions, and in section 2.2.2.2 the multivariate versions.

The tools usually collected under the rubric of "finding the distribution of a function of a random variable" enable precise statement and solution of the two questions above. Statement of these questions requires defining the variables X and Y, the function h that connects them, and the distribution of X; the solution is the distribution of Y. More generally, given any two of the three elements of the problem -- the distribution of X, the distribution of Y, and the function $y = h(x)$ -- it is possible to solve for the third element. We now illustrate two techniques, one utilizing the probability density function (PDF) and the other utilizing the quantile function (QF), using as

[3] The general problem is known as the problem of finding the distribution of a function of a random variable. One of the methods that is used in solving this problem is called the change-of-variable technique. Because the phrase "change-of-variable" is compact and immediately signals the associated operation, in this paper we refer to the general model involving the distribution of the function of a random variable as the change-of-variable model.

examples the two distributive-justice questions mentioned above.[4]

PDF-Based Technique for Finding the Distribution of a Function of a Random Variable. In the problem of finding the distribution of the justice evaluations (Jasso 1980), the given elements are the justice evaluation function (viz., the function *h* which connects the actual reward *X* and the justice evaluation *Y*) and the distribution of the actual reward, and the problem is to find the distribution of the justice evaluation. The justice evaluation function is given by

$$J = \theta \ln\left(\frac{A}{C}\right), \tag{1}$$

where *J* denotes the justice evaluation, *A* denotes the actual reward, *C* denotes the just reward, and θ is the signature constant which in this version of the problem is fixed at +1.

Suppose that the resource under consideration is income and that the distribution of income is a two-parameter Pareto, whose probability density function, denoted $f(x)$ is written:

$$f(x) = \left[\frac{\mu(c-1)}{c}\right]^c c x^{-c-1}, \qquad x \geq \frac{\mu(c-1)}{c}, \tag{2}$$

where the two parameters are the mean μ and a shape constant c which must be greater than unity.[5] Suppose further that the just reward is fixed at the mean of the distribution of income. Then the probability density function of the distribution of justice evaluations, denoted $g(y)$, is

---

[4] For comprehensive introduction to probability distributions, see Johnson, Kotz, and Balakrishnan (1994, 1995) and Stuart and Ord (1987). Mathematically specified distributions have associated with them a variety of functions. The most basic is the cumulative distribution function (CDF), which is defined as the probability α that the variate *X* assumes a value less than or equal to *x* and is usually denoted $F(x)$. Probably the best known of the associated functions is the probability density function (PDF), denoted $f(x)$, which in continuous distributions is the first derivative of the distribution function with respect to *x* (and which in discrete distributions is sometimes called the probability mass function). One of the most useful is the quantile function (QF), which among other things provides the foundation for whole-distribution measures of inequality such as Pen's Parade. The quantile function, variously denoted $G(\alpha)$ or $Q(\alpha)$ or $F^{-1}(\alpha)$, is the inverse of the distribution function, providing a mapping from the probability α to the quantile *x*.

[5] The shape constant operates as the general inequality parameter specified by Jasso and Kotz (2008).

found by applying the change-of-variable formula to the PDF of the income distribution:

$$g(y) = f(x)\left|\frac{dx}{dy}\right|, \tag{3}$$

where x is stated in terms of its relation to y (Hoel 1971:244-248). The formula yields:

$$g(y) = \left[\frac{c-1}{c}\right]^{c} c\, e^{-cy}, \qquad y \geq \ln\left(\frac{c-1}{c}\right). \tag{4}$$

The new PDF is recognized as a one-parameter negative exponential. Thus, when the distribution of the actual reward is Pareto, the distribution of justice evaluations is negative exponential. Panels A.1 and A.2 of Figure 1 depict graphs of the probability density functions of the Pareto and the negative exponential, respectively.[6]

-- Figure 1 about here --

QF-Based Technique. To illustrate the technique based on the quantile function, we draw on the Brickman et al. (1981) problem. In this case, the three elements are the just reward function (viz., the function *h* which connects a reward-relevant characteristic *X* and the just reward *Y*), the distribution of the reward-relevant characteristic, and the distribution of the just reward (Jasso 1983a).

Suppose that in a given observer's just reward function there is only one argument, education (ed), and that the just earnings function is given by:

$$\textit{just earnings} = 8(1.05)^{ed}, \tag{5}$$

where just earnings is in thousands of dollars. Suppose further that the schooling distribution is lognormal, with mean 12 and shape parameter 0.2, whose quantile function is written:

$$Q(\alpha) = 12 \exp\left[.2\, Q_N(\alpha) - .02\right], \tag{6}$$

where $Q_N(\alpha)$ denotes the quantile function of the standard normal variate. Then the quantile function of the just earnings distribution is found by applying the formula:

---

[6] In this problem, when the distribution of the actual reward is lognormal, the distribution of justice evaluations is normal, and when the distribution of the actual reward is either rectangular or power-function, the distribution of justice evaluations is positive exponential.

$$Q_Y(\alpha) = h[Q_X(\alpha)], \quad (7)$$

where h denotes the just earnings function (Hastings and Peacock 1974:18). The resultant formula for the quantile function of the just earnings distribution is:

$$Q_Y(\alpha) = 8 \times 1.05^{\{12 \exp[.2 Q_N(\alpha) - .02]\}}. \quad (8)$$

Panels B.1 and B.2 of Figure 1 depict graphs of the quantile functions for the schooling and just earnings distributions, respectively. As shown, the distributional form of the just earnings variable is unnamed; though the form resembles somewhat that of the extreme-value variate, further work is needed before we can either classify it as belonging to an already known family or else conclude that it belongs to a new variate and give the new variate a name.[7]

As noted above, because the three elements in the change-of-variable process are connected by mathematically necessary relationships, any one element can be solved for from the other two. In both our examples, we began with the distribution of X and the function $y = h(x)$, and solved for the distribution of Y. Thus, given the distribution of the reward-relevant characteristic (schooling in our example) and given the just reward function, the shape of the just reward distribution is entailed; similarly, given the distribution of the actual reward and given the justice evaluation function, the shape of the justice evaluation distribution is entailed. But these are not the only problems that can be solved. For example, in the Brickman et al. (1981) problem, it is possible to begin with the schooling distribution and the desired just earnings distribution, and solve for the entailed just earnings function; similarly, in the Jasso (1980) problem, it is possible to begin with the income distribution and the justice evaluation distribution, and solve for the entailed justice evaluation function.

A wide variety of results illuminates the change-of-variable problem, including the extremely useful Jensen's inequality, which establishes the connection between the convexity of

---

[7] In a recent similar case, theoretical analysis of status led to a new variate which did not resemble any known variate and which was termed "Unnamed" (Jasso 2001:122). Subsequently, Jasso and Kotz (2007) named it the ring(2)-exponential and generalized it to two new families of probability distributions, which they named the mirror-exponential and the ring-exponential.

the h function and the relative magnitudes of the mean of the outcome variable, $E(Y)$, and the functional transformation of the mean of the input variable, $h[E(X)]$, in the special case where X is nonnegative.[8] To illustrate, consider the justice evaluation function (Jasso 1980, 1999). In the special case in which the just reward C is a constant k, the justice evaluation function, $J = \ln(A/k)$, is logarithmic and hence nonconvex, so that the mean of the justice evaluations must be less than or equal to the logarithm of the mean of the argument (A/k):

$$E[\ln(A/k)] \leq \ln[E(A/k)] . \quad \textbf{(9)}$$

By algebraic manipulation, this reduces to the statement that the mean of log A must be less than or equal to the log of the mean of A:

$$E[\ln(A)] \leq \ln[E(A)] . \quad \textbf{(10)}$$

Finally, it can be shown that, in the further special case where the constant k is equal to the mean of A, the mean of the justice evaluations must be negative.

2.2.1.3. The Binary-Choice Model

Another kind of model -- the binary-choice model -- arises when the individuals in a population express their preferences for one or another distribution. Suppose that each individual has, as in the previous sections, an amount of income; now suppose that each individual is offered an alternative amount of income. There are now two micro variables in the population, current income and alternative income. If each person chooses between current income and alternative income, then in the aggregate the choice is equivalent to choosing between the current income distribution and the alternative income distribution. As a result of the choice, a new subpopulation structure is generated, plus new macro and submacro variables. The two new subpopulations consist of the subpopulation who chose current income and the subpopulation who chose alternative income; the new variables include macro variables defined on the population, such as "proportion who chose current income," and submacro variables defined on the subpopulations, such as "mean current income among those who chose alternative income."

---

[8] For exposition of Jensen's inequality, see Stuart and Ord (1987).

Also as a result of the choice, new micro variables are generated. These include the categorical preference variable and a quantitative strength-of-preference variable that can be defined as a function of the difference between current income and alternative income.

In this section we present procedures for studying binary-choice problems, using two examples. In the simplest version of the problem, each individual's rank α is the same on both the old and the new micro variable. Let A denote the distribution of the old micro variable and B that of the new micro variable. Then, for the *i*th individual,

$$\alpha_{i\mathrm{A}} = \alpha_{i\mathrm{B}} . \tag{11}$$

It follows that the quantile function associated with each of the two micro variables represents the values of the two variables corresponding to an individual of rank α. Thus, for example, $Q_{\mathrm{A}}(.25)$ represents the value of variable A corresponding to the individual whose rank is 0.25, and $Q_{\mathrm{B}}(.25)$ represents the value of variable B corresponding to the same individual. The variable whose value is larger at a given point α is said to dominate at that point. Accordingly, graphs of the two quantile functions, superimposed on the same grid, show which variable dominates over particular regions. Points at which the two quantile functions intersect are endpoints of regions of dominance.

To recover the subpopulation structure in a binary-choice problem, it is only necessary to solve the equation

$$Q_{\mathrm{A}}(\alpha) - Q_{\mathrm{B}}(\alpha) = 0 . \tag{12}$$

Each subpopulation consists of all the persons who favor a candidate and who lie contiguously on the income continuum. The number of subpopulations is at least two, but may range much higher; the size of the subpopulations may also vary greatly. Thus, each instance of a binary choice may generate a highly distinctive subpopulation structure, and with it new micro, macro, submicro, and submacro variables.

To illustrate the binary-choice model, we present two examples. The first concerns electoral choice; the second concerns the question of how individuals choose the goods they value. Both questions are of interest in several social science disciplines and subdisciplines.

Choosing an Income Distribution. Consider an election in which there are two candidates and they can be represented by the income distribution that would result from their policies; voting for a candidate is tantamount to voting for an income distribution. This type of model is analyzed in Jasso (1983b). As already noted, in the simplest version of the problem, both distributions produce the same ordering and each person votes for the candidate whose associated distribution maximizes his/her own income. The relevant questions include: What proportion of the population favor each of the two candidates? Do all the supporters of a given candidate lie in the same region of the income distribution?

Suppose that the two candidate-distributions are both drawn from the two-parameter Pareto family; they have the same mean, but one is more unequal than the other. To answer the questions posed, we solve equation (12). The number of solutions to this equation, excluding solutions at the extreme endpoints 0 and 1, is one less than the number of distinct regions of supporters. The proportion of the population in each distinct region is equal to the value of α at the upper endpoint of the region minus the value of α at the lower endpoint of the region. The total support for each candidate is calculated by summing the proportions in each of the regions supporting that candidate.

The equation to be solved is given by:

$$\frac{(c_A - 1)}{c_A(1-\alpha)^{1/c_A}} - \frac{(c_B - 1)}{c_B(1-\alpha)^{1/c_B}} = 0 . \qquad \textbf{(13)}$$

This equation equals zero at the critical point, denoted α^*:

$$\alpha^* = 1 - \left[\frac{c_A(c_B - 1)}{c_B(c_A - 1)}\right]^{\frac{c_A c_B}{c_A - c_B}} . \qquad \textbf{(14)}$$

Investigation of the critical point reveals that it occurs only once, so that there are only two distinct regions of dominance. Moreover, further investigation reveals that the critical point must lie between the point $1 - 1/e$ (approximately 0.632) and the point one. Thus, the candidate who wins, wins by a large majority, at least 63 percent of the population, and all supporters lie in a contiguous region of the income distribution. Additionally, the winning candidate is the

candidate that dominates over the leftmost region, thus benefiting the poorer individuals in the population. Panel A.1 of Figure 2 depicts the binary-choice model for the case where the two Pareto distributions have shape parameters equal to 1.5 and 2.0.

-- Figure 2 about here --

Suppose now that the two candidate-distributions are not both drawn from the Pareto variate; although one is Pareto, the other is lognormal. In this case, shown in Panel A.2 of Figure 2, the resulting subpopulation structure has three distinct regions of dominance, with the very interesting result that the candidate favored by the poorest individuals -- the Pareto -- is also the candidate favored by the richest individuals, with a large middle segment favoring the lognormal.

Note that in both cases, the new variables that are generated include macro variables defined on the original population, such as "number of distinct regions formed by the choice process," submacro variables defined on the new subpopulations, such as "mean income among the poorest segment favoring candidate A," micro variables, such as "strength of preference for candidate A," and submicro variables, such as "income rank within the subpopulation supporting candidate A."

In the binary-choice problem, the precise character of the subpopulation structure is determined jointly by the two candidate-distributions. Thus, this is a most intriguing problem, filled with interesting sociological questions. At one extreme, we may conceive of the candidate-distributions being completely constrained by the society's resource base and technological level (i.e., there are only two feasible income distributions), so that contextual and other inherently macro variables exert considerable influence on the outcome. At the other extreme, we may imagine a large set of feasible income distributions so that the candidates' decisionmaking assumes great importance.

Choosing a Good. For our second illustration, we draw on the problem of how individuals choose the goods they value, a problem discussed in the theory of distributive justice and now seen to be a problem in the broader new unified theory of sociobehavioral forces (to be discussed in Section 3 below). In one set of models designed to address this question in justice

theory, individuals choose to value the good from which they derive the highest magnitudes on the justice evaluation variable (Jasso 1987). In a special case of this set of models, the choice is between valuing the income amount and valuing the income rank. The problem is solved by solving equation (12) for the difference between the quantile function associated with the positive exponential which results from an ordinal good and the quantile function associated with the justice evaluation distribution which results from the income distributional pattern of the given cardinal good. For example, if income is lognormally distributed, then the distribution of justice evaluations is normal; and, as discussed in section 2.2.1.2 on the change-of-variable model, if income is Pareto distributed, then the distribution of justice evaluations is negative exponential. In this problem of choosing between income amount and income rank, there is no analytic solution for equation (12) for either the case where income is Pareto distributed or the case where income is lognormally distributed; that is, there is no analytic solution for the binary choice problem when one distribution is the positive exponential that arises from an ordinal good and the second is either normal or negative exponential. However, it appears that three distinct regions are generated and that the middle region always commands a majority. Panels B.1 and B.2 of Figure 2 present the graphs for these two cases.

Note that when the binary-choice problem involves a choice between the value of a variable and the associated rank, as in the special case just discussed, the problem contains only one variable and thus could be considered in the one-population/one-variable class. However, in the general case the binary-choice problem requires two distinct variables, and hence its logical place is in this two-variable section. Nonetheless, it should become increasingly apparent that a single variable can go a very long way, generating new micro and submicro variables as well as new macro and submacro variables.

Finally, note that it is possible to obtain an analytic solution to equation (12) in certain cases in which the two variates are drawn from different distributional families. A prime example involves the case where the two variates are Pareto and power-function, both with the same value of the shape parameter c. In this case, equation (12) produces a quadratic equation

whose roots are two points equidistant from the median. Thus, there are three regions of dominance, the leftmost and rightmost containing equal shares of the population, and the middle region containing a majority, of never less than 67.7 percent, favoring the power-function alternative. If there were a contest between two candidates whose visions of society generated, respectively, a power-function and a Pareto distribution of income, the candidate associated with the power-function would win by a landslide; and both the poorest and the richest individuals would be worse-off than if the opponent had won.

2.2.2. Introducing Several New Variables

This section presents extensions of the change-of-variable and binary-choice problems to the case where the number of micro quantitative variables is greater than two. As well, this section presents the procedure mentioned in section 2.1, involving only one explicit variable but many implicit variables.

2.2.2.1. The Multiple Regression Model

Extension of the simple regression model to the case of several micro quantitative variables is straightforward. One of the variables is designated the dependent variable and its regression on the others yields several slopes as well as an intercept, R^2, and all the associated quantities (such as the various sums of squares). The obtained estimates are new macro variables which characterize the population.

2.2.2.2. Distribution of a Function of Several Random Variables

Generalization of the change-of-variable problem yields the new problem studied under the rubric of "distribution of a function of several variables." While the techniques are more complicated -- e.g., convolution techniques -- the logic is similar, and many results can be obtained.

An example of this problem is provided in Jasso (1999), where the problem of finding the distribution of justice evaluations is extended to the case where the justice evaluation varies as a function of two variables, the actual reward *A* and the just reward *C*. By assigning distributional forms to both arguments and a specified relation between them, a range of possible justice

evaluation distributions is obtained. In general, there are six ideal types of interest in this kind of theoretical analysis. The six ideal types arise as the combinations of two dimensions: (i) whether the actual reward *A* and the just reward *C* are identically or differently distributed; and (ii) whether *A* and *C* are independently distributed, perfectly positively associated, or perfectly negatively associated.[9]

A variety of special tools enable derivation of the distribution of the outcome variable in this kind of problem. In particular, techniques based on the probability density function are useful for the two cases in which *A* and *C* are independently distributed, and techniques based on the quantile function and its inverse are useful for the four cases in which *A* and *C* have perfect association.

An important objective in this kind of problem is to obtain distribution-independent results. We have obtained four relatively simple but nonetheless substantively meaningful results, collected in panel A of Table 1. First, if the actual reward and the just reward are independently and identically distributed, then the justice evaluation distribution is symmetric about zero -- that is, exactly equal numbers of the population are judged to be (or judge themselves to be) underrewarded and overrewarded, and the underrewarded subdistribution is a mirror image of the overrewarded subdistribution. Second, if the actual reward and the just reward are identical and perfectly negatively associated, then the justice evaluation distribution is symmetric about zero -- as in the first result. Third, if the actual reward and the just reward are identical and perfectly positively associated, then the justice evaluation is degenerate at zero -- that is, every individual is judged to be (or judges him/herself to be) perfectly justly rewarded. Fourth, if the actual reward and the just reward have different distributions, then, regardless of their association, the justice evaluation distribution can be symmetric or asymmetric about any

[9] Perfect positive association denotes the case in which each individual has the same relative rank on both *A* and *C*. Perfect negative association denotes the case in which the rank ordering in *A* is exactly the reverse of the rank ordering in *C*; thus one ranking is the conjugate ranking of the other (Kotz, Johnson, and Read 1982b:145).

number.

-- Table 1 about here --

Of course, distribution-independent results tend to be scarce. Hence, much of theoretical analysis in this kind of problem involves obtaining results for variates from specified distributional families. Table 1, panel B, reports a number of results, for the special case where both the actual reward *A* and the just reward *C* are drawn from the same variate family. As expected from the distribution-independent results, the justice evaluation distributions are Equal in the cases where *A* and *C* are identical and perfectly positively associated. Also as expected, when *A* and *C* are identical and either independent or perfectly positively associated, the justice evaluation distributions are symmetric. Figures 3 and 4 depict graphs of the justice evaluation distributions for some special cases. Figure 3 highlights the case where *A* and *C* are independent, and utilizes the PDF representation of the distributions; Figure 4 highlights the cases where *A* and *C* are perfectly associated, and utilizes the QF representation of the distributions.

-- Figures 3 and 4 about here --

As shown in Table 1, when the justice evaluation is experienced about a personal attribute that is ordinal (such as beauty or athletic skill), and the reward distribution is thus constrained to the rectangular, the two possible outcome distributions for the independent case and the perfect-negative-association case are, respectively, the Laplace and the logistic. The rectangular distribution is a special case of the power-function variate (the case where the shape parameter *c* is equal to unity), and hence it is no surprise that the corresponding outcomes for the power-function variate are also the Laplace and logistic. Moreover, the Pareto variate is related to the power-function variate and, here, too, it is no surprise that the counterpart distributions of justice evaluations are also Laplace and logistic. Graphs of the Laplace and logistic appear, respectively, in the upper-lefthand panel of Figure 3 and the lower-lefthand panel of Figure 4.

When the actual reward *A* and the just reward *C* are different (but drawn from the same family), the distributions of justice evaluations, for both the Pareto and the power-function case, are (i) asymmetrical Laplace, in the case of independence, (ii) positive or negative exponential, in

the case of perfect positive association, and (iii) quasi-logistic, in the case of perfect negative association (Table 1). Graphs of these variates appear in the upper-righthand panel of Figure 3 and the two righthand panels of Figure 4. In the case of perfect positive association, the pattern of shape parameters -- that is, whether the actual reward distribution or the just reward distribution is the more unequal -- determines whether the exponential will be positive or negative.

When the actual reward A and the just reward C are lognormal, then the distribution of justice evaluations is normal in five of the six ideal types, the sole exception being the case where A and C are identical lognormals and perfectly positively associated, in which case the outcome distribution is the Equal (Table 1). Of course, the five normal outcome distributions are not all the same member of the normal family. In particular, we know that in the two cases where A and C are identical, the mean of the normal is zero; however, in the three cases where the lognormals are different, the normal's mean may or may not be zero, depending on whether the two lognormals have the same or different degree of inequality. The lower-righthand panel of Figure 3 depicts the graph of the normal for a case where the two lognormals are different yet give rise to a normal with mean zero.[10]

Note that the problem of finding the distribution of justice evaluations, given the distributions of the actual reward and the just reward and the relation between them, exemplifies a particularly interesting problem in social science, namely, the problem of how "social structure" (here the distribution of the actual reward) combines with "personal ideology" (here the distribution of the just reward) to produce a variable which has both behavioral and social consequences (here the justice evaluation).

---

[10] Figure 3 also reports, in the lower lefthand panel, the graph of the probability density function of the distribution of justice evaluations when the actual reward and the just reward are identically and independently distributed exponentials; as shown, in this case the justice evaluation distribution is logistic.

2.2.2.3. The Multivariate Choice Problem

Extension of the choice problem from the bivariate to the multivariate case is similarly straightforward. It can be shown that the number of subpopulations in the multivariate choice problem can still assume the value two. That is, it is theoretically possible for only two candidates to take all the vote in a multi-candidate election. As in the binary-choice version, there can be great variability in the number and size of the subpopulations. Figure 5 presents the multivariate choice problem for a case where there are four competing distributions, two drawn from the Pareto family (with shape parameters set to 1.5 and 2.0) and two drawn from the lognormal family (with shape parameters set to 0.6 and 1.0). As shown, the less unequal of the two Paretos dominates over the leftmost region of dominance, which contains slightly over 28 percent of the population. The next region of dominance is represented by the less unequal of the two lognormals; this subpopulation contains slightly more than half of the population. The third region of dominance extends from about the 78.9th percentile to the 98.3th percentile; the dominant candidate over this region is the more unequal of the two lognormals. Finally, there is a fourth small region of dominance (not visible in Figure 5); the more unequal of the two Paretos dominates over this region, which contains the top 1.6 percent of the population.

-- Figure 5 about here --

The four-candidate case depicted in Figure 5 has interesting real-world implications. In a multiparty contest of this sort, the four parties represent two contrasting political philosophies, one advocating a "safety net" for the poorest members of society and the other willing to tolerate substantial poverty. Each philosophy is espoused by two of the parties -- the safety-net parties represented by the Paretos and the non-safety-net parties represented by the lognormals – the elements within each pair differing somewhat in the extent of inequality they tolerate. The winning party, by majority vote, is the less unequal of the two non-safety-net parties. Meanwhile, the richest members of society, who favor a generous safety net for the poor (though not the more generous of the two parties favoring a safety net) are isolated into a tiny elite.

One can speculate about the next election, indeed, about a sequence of subsequent

elections. For example, the two safety-net parties (the Paretos) may decide to join forces; if they face each other in a primary, the less unequal of the two -- the party which took over 28 percent of the vote in the general election -- will win with more than 90 percent of the vote, as discussed in section 2.2.1.3 on the binary-choice problem. The two parties may, however, decide to compromise, proposing a platform that is less egalitarian than that of the winner. As for the two parties representing the non-safety-net philosophy, there is less incentive for them to join forces, given that one of them won the previous election. If they do decide to join forces, then the less unequal among them -- the winner in the previous election -- will win the primary. Again, the ensuing party platform may represent a compromise between the two parties. Which party wins in the general election in fact depends on the extent of the compromise. If the income distributions associated with the two parties are represented by the two top parties in the previous general election -- which were, respectively, the less unequal from each pair (the Pareto with shape parameter of 2 and the lognormal with shape parameter of 0.6) -- then the non-safety-net party wins. But if the non-safety-net party espouses a more unequal income distribution in the new election, then it may lose to the safety-net party; for example, if the non-safety-net party espouses the income distribution associated with its previous opponent (a lognormal with shape parameter of 1.0), then the safety-net party wins with a coalition consisting of the poorest 61 percent of the population and a tiny elite of about a quarter of one percent.

The preceding example is both substantively and methodologically interesting. We began with a multivariate choice problem in which each individual has a choice between one of four alternative parties. The choice produced a subpopulation structure consisting of four distinct subpopulations. The outcome of the choice led the four parties to combine into two parties. The ensuing binary choice generates a new subpopulation structure consisting of three subpopulations. Notice the dynamism inherent in this type of problem. Each round generates both a subpopulation structure and an outcome that contains the seeds for a future choice that may generate an altogether new subpopulation structure. Of course, associated with each subpopulation structure there are both a set of macro variables and a set of submacro variables.

Additionally, the usual new micro variables -- such as strength of preference for one candidate over another -- and submicro variables are generated.

2.2.2.4. Studying the Effect of a Macro Variable on a Micro Variable

We now describe the case first mentioned in section 2.1. We begin with what appear to be one population and one micro variable. As already noted, the variable must be quantitative and its probability distribution mathematically specified. To fix ideas, let the variable in the problem be income, so that the population is a population of income receivers. Each person has an amount of own income, and, consequently, an income rank (a new micro variable generated from the original micro variable). Moreover, the income distribution can be described by a number of summary measures, including the mean and the inequality; these are macro variables.

A perennial question is, What is the effect of income inequality on own income? More generally, we may investigate the effects of the macro variables on the micro variables, asking, What are the effects of mean income, income inequality, and other macro variables on the two micro variables, income amount and income rank?

In an empirical context, such questions cannot be addressed without altering the fundamental parameters of the problem. That is, given that there is only one population and hence only one observed magnitude for each macro variable, the problem of assessing the effect of a macro variable on a micro variable is not defined.

In the theoretical context, however, such a question is readily addressed. The only thing that is required is that the distribution of the original micro variable be mathematically specified. For example, one can let income be Pareto distributed or lognormally distributed, etc. Suppose that we choose as modeling distributions a large set of continuous univariate two-parameter distributions, where the two parameters are the mean μ and the general inequality parameter c and where the associated cumulative distribution function is continuous and strictly increasing. It then follows that the cumulative distribution function $F(x)$ and the quantile function $Q(\alpha)$ can be expressed in a form in which one of the two micro variables is a function of the second micro

variable and the two macro parameters:[11]

$$\begin{aligned} x &= Q(\alpha) = Q(\alpha; \mu, c) \\ \alpha &= F(x) = F(x; \mu, c) \end{aligned} \quad . \tag{15}$$

Viewed in this way, it is straightforward to achieve our objectives. The effect of inequality on income amount is given by the first partial derivative of $Q(\alpha)$ with respect to c; and the effect of inequality on income rank is given by the first partial derivative of $F(x)$ with respect to c. Similarly, the effect of mean income on income amount is given by the first partial derivative of $Q(\alpha)$ with respect to μ; and the effect of mean income on income rank is given by the first partial derivative of $F(x)$ with respect to μ. For example, consider the effect of mean income on income amount. It can be shown that the corresponding first partial derivative is always positive. This means that, holding constant both own rank in the income hierarchy and income inequality, increasing the mean income increases the individual's income amount. Similarly, consider the effect of mean income on income rank. It can be shown that the relevant first partial derivative is always negative. This means that, holding constant both own income amount and income inequality, increasing the mean income reduces the individual's rank in the income hierarchy.

Note that implicit in the forms given in equations (15) is an infinite number of distributions, defined by different magnitudes of μ and c. That is why we include this procedure in the present section rather than in the one-population/one-variable section. Note also that implicit in this procedure is the idea that in a given society all the feasible income distributions come from the same variate family. Thus, the correct interpretation of the derivatives is context-specific.

The technique of taking the first partial derivative of the quantile function with respect to the inequality parameter was used by Jasso (1989) to address the question, "How does own

[11] For fuller discussion of the general inequality parameter in continuous univariate two-parameter distributions, see Jasso and Kotz (2008).

income vary with income inequality?" If inequality were to decline, whose incomes would rise? and whose would fall? In general, the dependence of own income on income inequality can assume three qualitative forms -- own income can increase, decrease, or change nonmonotonically as income inequality increases.

The analysis in Jasso (1989) was motivated by an idea owed to St. Anselm of Canterbury -- that human volition has a twofold inclination, such that individuals seek both their own good, an inclination known as the affectio commodi, and also the common good, an inclination known as the affectio justitiae.[12]

To the extent that an increase in own income enhances one's own good and that a decline in income inequality enhances the common good, a new micro variable is generated in a natural way. Persons whose own income increases when income inequality declines can simultaneously promote their own good and the common good; thus we may say that they are in the state of Harmony. Persons whose own income decreases when income inequality declines cannot simultaneously promote their own good and the common good; we may say that they are in the state of Conflict. Finally, persons whose own income is a nonmonotonic function of income inequality, perhaps increasing over one subset of the alternative income distributions and decreasing over another subset, may be said to be in a state of Ambiguity. The new micro variable -- which may be called volitional state -- characterizes each person in the population according to which one of the three states he or she is in.

Investigation of the first partial derivative of $Q(\alpha)$ with respect to c, in five modeling distributions, showed that in all five types of societies, there is a subpopulation in Harmony; and this subpopulation, whose size ranges from 37 percent to 63 percent, consists of the leftmost, or poorest, persons. In all five hypothetical societies, there is only one other subpopulation; in two

[12] St. Anselm's (1033-1109) idea that the created will has two inclinations is worked out first in his famous thought experiment on the devil, reported in the dialogue De Casu Diaboli (The Fall of Satan) which dates from about 1085-90, and again in De Concordia Praescientiae et Praedestinationis et Gratiae Dei (The Harmony of God's Foreknowledge, Grace, and Predestination), written in about 1107-08.

of the societies the second subpopulation is in Conflict, whereas in the remaining three societies the second subpopulation is in Ambiguity.

Substantively, this procedure generates a wealth of new questions and highlights many interesting outcomes. In particular, it exemplifies the intersection of several social sciences. For example, it raises the question whether each of the three volitional states -- Harmony, Conflict, Ambiguity -- generates a distinctive sense of self and a distinctive view of society. To the extent that the set of feasible income distributions is constrained -- whether by resource endowment, technological level, or inherited political arrangements -- the "social structure" (here the income distributional type) shapes the individual psyche, and the individual psyche in turn shapes the social and political discourse.

Methodologically, note that the underlying mathematical structure in this problem is the same as that in a special case of the multivariate choice problem, namely, the special case in which all the candidate-distributions are drawn from the same variate family. For example, it can be shown that the limits of the regions of dominance in the multivariate-choice problem correspond to the boundaries between the subpopulation in Harmony and the second subpopulation in the Anselmian problem. Similarly, the feature in the Anselmian problem whether the second subpopulation is in Conflict or in Ambiguity corresponds to the feature in the multivariate choice problem whether the critical point α^* is a constant or ranges within a region.

2.3. Introducing Populations

2.3.1. Introducing a Second Population

To this point, we have focused on a single population, allowing it to have first one variable, then two variables, finally many variables. Now we systematically add populations, beginning with the case of two micro quantitative variables. We consider three models, the hierarchical and multilevel model, the pooling model, and the matching model.

2.3.1.1. Hierarchical and Multilevel Models

In section 2.2.1.1, we saw that introducing a single new micro quantitative variable generated for the population a set of new macro variables describing the association between the

two micro variables -- the intercept and slope of the regression line, the correlation between the two variables, various sums of squares, and so on. Introducing a second population enables a new type of analysis, in which the macro variables obtained from the micro variables are systematically related to other macro variables which may be inherently macro (that is, which cannot be obtained by aggregating over a micro variable), such as a country's distance from another country or a classroom teacher's style of teaching. This type of analysis has been extensively studied in recent years (Goldstein 2003; Raudenbush and Bryk 2002).

2.3.1.2. The Pooling Model

Suppose that there are two populations; a characteristic of interest is observed in both populations. An operation familiar to all readers is the pooling operation, in which the two populations are combined to form a superpopulation (or, equivalently, the two distributions are combined to form a superdistribution). For example, the two sections of a third-grade class may be combined and their scores on a standardized test examined in a pooled superdistribution. The size of the superpopulation is the sum of the sizes of each of the constituent populations. By well-known theorems, the mean of the pooled distribution is equal to the weighted sum of the means from the constituent distributions; and other characteristics of the pooled distribution can also be derived. For example, if the two distributions are mathematically specified, then the probability density function of the superdistribution (a mixed distribution) is found by taking the weighted sum of the probability density functions of the component distributions.

To illustrate, suppose that two social clubs decide to merge. The shape of the income distribution of the new merged super-club can be obtained from the income distributions of the two original clubs. If, say, the two original income distributions are both Pareto, one with parameters (10, 2), the other with parameters (15, 1.5), and if the two original populations (i.e., the social clubs) are the same size, then the PDF of the superdistribution is given by:

$$f(x) = 25x^{3} + \frac{15\sqrt{5}}{4} x^{-5/2}, \qquad x > 5 . \tag{16}$$

In this case the density is smooth, everywhere decreasing, with mode at the lower extreme value

(like the Pareto). Panels A.1 and A.2 of Figure 6 depict graphs of the probability density functions of the two original distributions and the new pooled distribution, respectively.

-- Figure 6 about here --

A unimodal distribution, such as that shown in panel A.2 of Figure 6, is by no means the universal outcome when the two component distributions are Pareto. In fact, a unimodal distribution is obtained only when the two component Pareto distributions have the same lower extreme value. In all other cases the mixed distribution will be bimodal. If, say, the two original distributions have parameters (10, 2) and (8, 2), then the probability density function of the superdistribution is given by:

$$f(x) = \begin{cases} 16x^{-3}, & 4<x<5 \\ 41x^{-3}, & x>5 \end{cases}. \qquad \textbf{(17)}$$

The superdistribution has modes at the lower extreme values of the two original distributions. Figure 6 shows graphs of the probability density functions of the two component distributions, in panel B.1, and of the bimodal pooled superdistribution, in panel B.2.

The pooling operation produces not only a new superpopulation but also, along with it, a new set of macro variables, namely, the summary measures of the new superdistribution; these new macro variables may be called supermacro variables. In the example of the two social clubs, the new supermacro variables are the mean, variance, and other parameters of the new pooled income distribution. Additionally, a new supermacro variable is generated which characterizes the component structure of the superpopulation and which yields the proportion of the superpopulation that originated in each of the original populations. Furthermore, a new micro variable is generated, namely, the rank in the new superpopulation, which we may call a supermicro variable.

2.3.1.3. The Matching Model

Another important kind of operation that arises in the study of two populations and one variable is the "matching" operation. Suppose that we have a population of married women, and

we measure their income. Now suppose that we become interested in their husbands, and construct the population of husbands and measure their income. From these two populations we can derive a new matched population, consisting of couples, and along with it many interesting new supermacro variables and relationships. Two of the simplest new supermacro variables are: (i) the variable "combined income", which is the sum of wife's income and husband's income; and (ii) the variable "spousal income disparity", which is the signed difference between the two spouses' incomes. Note that the size of the superpopulation generated by a matching process is exactly the same as the size of each of the two component populations (in contrast to the size of the superpopulation generated by a pooling process, which is the sum of the sizes of the component populations).

To illustrate the range of questions that can be investigated using a matching framework, consider that the spousal income disparity variable can assume three qualitative conditions -- positive, negative, and zero -- corresponding to the situations where husband outearns wife, wife outearns husband, and both earn equal amounts. Thus, a new qualitative variable is generated describing couples, and from it a new subpopulation structure is generated, with each subpopulation containing one of the three kinds of couples. Note that aspects of the subpopulation structure -- e.g., whether more than half of the couples fall into any one of the three conditions, and, if so, which one -- may turn out to be important macro variables determinative of societal gender-related institutions.

Of course, the income disparity variable is a quantitative variable, and a theoretical analysis of this problem can utilize methods for quantitative distributions. A first objective is to describe the distribution of the disparity variable. This can easily be accomplished by means of the same techniques introduced in section 2.2.2.2 under the rubric of distribution of a function of two random variables. In the present case, the two random variables are wife's income W and husband's income H. As before, six ideal types are generated. Wife's income and husband's income can be identically or differently distributed. As well, wife's income and husband's income may be independently distributed (as in random mating), perfectly positively associated

(as in perfect positive assortative mating), or perfectly negatively associated (as in perfect negative assortative mating).

Note how two distinct situations give rise to the same formal mathematical structure. In the first, discussed in section 2.2.2.2, there is one population and two micro variables; the objective is to find the distribution of a function of the two micro variables, where both micro variables are observed in each unit of the population. In the second, discussed in this section, there are two matched populations of the same size and one micro variable observed in both; the objective is to find the distribution of a function of the two micro variables, one each observed in each of the two populations in the match. Because the formal mathematical structure is the same, work undertaken to advance substantive understanding of one topical domain (say, the distributive justice problem) is directly pertinent to substantive understanding of a completely different topical domain (namely, the spousal disparity problem).

Accordingly, we may import results obtained in the study of the distribution of justice evaluations, and state the following distribution-independent conclusions: First, if wife's income and husband's income are independently and identically distributed, then the distribution of the income disparity variable is symmetric about zero -- that is, the number of couples in which the husband outearns the wife is exactly equal to the number of couples in which the wife outearns the husband, and the disparity subdistribution among one type of couple is a mirror image of the disparity subdistribution among the other type of couple. Second, if wife's income and husband's income are identical and perfectly negatively associated, then the disparity distribution is symmetric about zero -- as in the first result. Third, if wife's income and husband's income are identical and perfectly positively associated, then the disparity distribution is degenerate at zero -- that is, within every couple, the husband and wife are equal earners.

A potentially important general result can be stated, based on the distribution-independent results just given: Restricting attention to societies in which mating is either random, perfectly negatively assortative, or perfectly positively assortative, a necessary condition for one of the two unequal types of couples to become the dominant form is that the wife's income distribution and

the husband's income distribution differ. This result exemplifies some of the most interesting and fundamental problems of social science, namely, how the "social structure" (here the two income distributions) shapes the "couple structure" (here whether the couple is husband-superior, wife-superior, or equal), and the "societal couple structure" (here the distribution of the three couple types) in turn shapes gender-related customs and institutions. Notice the lines of influence, from a macro level, to a couple level, back to a macro level.

Methodologically, two points are in order. First, as already noted, certain problems that arise in the one-population/one-variable context have the same mathematical structure as certain problems that arise in the two-population/one-variable context. For example, the problem of finding the distribution of the justice evaluation J and the problem of finding the distribution of spousal disparity D are the same, given that the J and D functions are similar:

$$\begin{aligned} J &= \ln(A) - \ln(C) \\ D &= W - H . \end{aligned} \tag{18}$$

Second, in the special case where the input variables -- say, W and H, or $\ln(A)$ and $\ln(C)$ -- are drawn from the same variate family and are perfectly positively associated, the underlying mathematical structure is similar to that in a special case of the binary choice model, namely, the special case where both distributions are drawn from the same variate family. To see this, it is only necessary to explicitly define a new strength-of-preference variable S as the difference between the A and B candidate-distributions:

$$S = A - B . \tag{19}$$

Note that the point at which S equals zero is the critical point dividing regions of dominance in the binary choice problem.

2.3.2. Introducing Several New Populations

All three models just described – hierarchical and multilevel models, pooling models, and matching models -- are readily extended to more than two populations. Indeed, there is a large literature on hierarchical and multilevel modeling with many populations (Goldstein 2003; Raudenbush and Bryk 2002). Pooling several populations, too, is familiar; and extension of

matching models to several populations is also straightforward. We leave to future work a detailed exposition of the methods for pooling and matching in several populations, noting that, while pooling of k populations requires few new methods, matching of k populations requires definition and analysis of n-tuples.

2.4. The General Case: Several Populations and Several Variables

Table 2 summarizes the operations and outcomes for the nine possible combinations of one/two/several populations and variables. The table omits the simplest result which characterizes all cells, namely, that any micro variable can generate a subpopulation structure and subpopulation macro variables, and that any quantitative micro variable can generate new micro variables. The problem of finding the distribution of a function of several random variables occurs in many specifications of matching problems, but, for simplicity, is omitted from the six cells involving matching. Note also that all the operations that can be performed on a single population can be performed on the superpopulation generated by pooling or matching.

-- Table 2 about here --

The entries in Table 2 suggest that the operations arising in this framework reduce to seven main types: single-variable generation of micro and macro variables; regression; hierarchical and multilevel modeling; distribution of a function of one or several variables; choosing between two or more distributions; pooling; and matching. Thus, a relatively small tool-kit can be applied to a wide range of problems spanning many distinct types, both substantively and methodologically. Of course, as new procedures are developed, they, too, can be incorporated into the framework.

3. THEORETICAL ILLUSTRATION: RACIAL SEGREGATION, INDIVIDUAL PREFERENCES, AND COLLECTIVE OUTCOMES

A perennial and important question in social science pertains to the determinants of segregation by personal qualitative characteristics, such as race, ethnicity, religion, or language. A major line of analysis, pioneered by Schelling (1969, 1971), examines how individual

preferences for segregation yield particular collective outcomes, such as racial residential segregation of specified degree. Schelling developed a general model, applicable to segregation of all types (not only residential but also in school, workplace, club, and so on) and with variation along five dimensions – proportion in each subgroup, segregationist/integrationist preference, neighborhood size, initial configuration, and rules for moving (Schelling 1971:152). Schelling explored a large number of combinations of the five elements, stating numerical constraints and deriving many predictions. His work has inspired, and continues to inspire, many researchers to examine more special cases, to vary more parameters, and to enlarge the scope of the problem, for example, by formulating state-of-the-art agent-based modeling (ABM), introducing state-of-the-art geographic information systems (GIS), and assessing the empirical plausibility of the individuals' (or agents') preferences (see Bruch and Mare 2006, Crooks 2008, and the references cited therein).

This section explores theoretically Schelling's (1971) original idea that persons may have different preferences and that these preferences may differ both between and within subgroups – an idea explored computationally and empirically by Bruch and Mare who report models that permit preference variation across race and mention models that also permit preference variation within race (Bruch and Mare 2006:690). That is, we build on Schelling's model by asking where the preferences come from and addressing the question theoretically. Thus, complementing Bruch and Mare (2006:698) who argue, correctly, that the preferences assigned to artificial agents in agent-based modeling should be based "on a sound empirical footing," we argue here that the preferences should also be based on a sound theoretical footing.

The approach in this section enables generating segregationist/integrationist preferences from a general theory of sociobehavioral forces. The predicted preferences vary with the proportions in each subgroup, so that another of the five dimensions in the Schelling model is captured. Notably, this approach covers two distinct generative operations. First, the theory generates each individual's segregationist/integrationist preference. Second, the theory shows how these preferences generate the emergent segregated and mixed subgroups. Obviously, the

second generative operation provides a quintessential example of how individual preferences lead to special properties of collectivities such as their degree of segregation.

Further, we report two distinct processes suggested by the general theory. The first is characterized by homogeneous preferences both between and within subgroups; this is an extreme process in which every individual thinks and acts exclusively as a member of his or her subgroup. The second is characterized by heterogenous preferences both between and within subgroups; the theory provides the mechanisms driving both homogeneous preferences and heterogeneous preferences

The following subsections provide a brief overview of the elements of the general theory which will be used in the segregation application, summarize the two segregation processes, linger on one special case, and, finally, return to examine how the two processes might fit in the micro-macro framework presented in Section 2.

3.1. Brief Overview of the New Unified Theory of Sociobehavioral Forces

The goal of the recently-proposed new unified theory (NUT) is to integrate theories describing five sociobehavioral processes – comparison (including justice and self-esteem), status, power, identity, and happiness (Jasso 2008). The integration (partial, in the case of happiness) is made possible by the remarkable similarity of the internal core of the theories. The core consists of three elements: personal qualitative characteristics, personal quantitative characteristics, and primordial sociobehavioral outcomes. The theory posits the operation of three sociobehavioral forces (comparison, status, and power), each associated with a distinctive way of generating a primordial sociobehavioral outcome from a personal quantitative characteristic – notably, a distinctive rate of change of the outcome with respect to the quantitative characteristic. Each combination of elements – for example, justice-wealth-city or status-beauty-classroom – generates a distinctive identity and a distinctive magnitude of happiness.

The comparison and power forces notice both ordinal and cardinal quantitative characteristics. Thus, when a person reflects on her knowledge of Greek or his relative rank in

the wealth distribution, the quantitative characteristic has a rectangular (or continuous uniform) distribution, but when a person reflects on her amount of wealth or his amount of land, the quantitative characteristic must be approximated or modeled via a continuous distribution (such as the Pareto or lognormal).

In contrast, the status force notices only relative ranks. Thus, the underlying distribution of the quantitative characteristic is always rectangular.

A long literature discusses desirable properties in the functions which generate comparison, status, and power from the quantitative characteristics. Strong cases can be made for the functional forms associated with comparison,

$$Z = \ln\left(\frac{A}{C}\right), \tag{20}$$

where Z denotes the comparison outcome (such as self-esteem or the justice evaluation J discussed above), A denotes the actual amount or relative rank of the good, and C denotes the expected, desired, or just amount/rank of the good, and with status,

$$S = \ln\left(\frac{1}{1-r}\right), \tag{21}$$

where S denotes status and r denotes the relative rank on the good.

For simplicity, we will often refer to the comparison force as the justice force, but it should always be understood that this is shorthand for "justice and all the other members of the class of comparison processes."

Referring to quantitative characteristics of which more is preferred to less as a good and characteristics of which less is preferred to more as a bad, these functional forms show that the comparison outcome increases at a decreasing rate with the good and that the status outcome increases at an increasing rate with the good.

We then reason that if power is a separate force – distinct from comparison and status, as conjectured, for example by Homans (1974:231) – then it must increase at a constant rate with the good.

Thus, in the NUT each of the three sociobehavioral forces has its own distinctive rate of change.

Every person has a large repertory of combinations of sociobehavioral force, quantitative characteristic, and qualitative characteristic. And at each turn of the sociobehavioral wheel, so to speak, a new identity and a new magnitude of happiness are generated. Thus, each person can be characterized by a distinctive sociobehavioral profile.

Meanwhile, a collectivity can be characterized by the instantaneous identities of all its members. In simple, a priori modeling, it is assumed that all persons are governed by the same force and reacting to the same characteristic. Accordingly, if the good is ordinal, the distribution of identities will be rectangular in the power case, positive exponential in the comparison case, and negative exponential in the status case. If the good is cardinal, the distribution of identities will still be negative exponential in the status case – because status notices only relative ranks – but will assume a wide diversity of shapes in the comparison and power cases.

There are yet further complexities and elaborations. For example, a person may value more than one good at the same time, and the goods may be independent or positively or negatively associated. Such combinations can lead to new distributions of the outcome, such as the distribution called "unnamed" by Jasso (2001) and extended to the mirror-exponential family introduced and analyzed by Jasso and Kotz (2007).

Sometimes the group or population has a subgroup structure generated by a qualitative characteristic, such as race or sex. In such case, each individual has access not only to the identity generated by his or her magnitude or relative rank in the quantitative characteristic – now called the Personal Identity – but also to a new identity called the Subgroup Identity and defined as the average of the personal identities within the subgroup (Jasso 2008).

The two models of segregation processes are among many models that arise from the general theory. Below we summarize the first segregation model, providing illustrations for all three sociobehavioral forces – justice, status, and power – for both ordinal and cardinal goods, and modeling the cardinal goods by three probability distributions. For the second segregation

model, which is more complicated, we outline the simplest case, which is the status case.

3.2. First Segregation Model

Suppose that the collectivity has two or more subgroups defined by a qualitative characteristic – for example, race, ethnicity, nativity, or sex. Each individual thus acquires both a Personal Identity and a Subgroup Identity. In this first segregation model, every person thinks and behaves as a member of the subgroup; the Personal Identity does not matter. Relations between persons from different subgroups have a tone that can be characterized as varying with the absolute difference between the two Subgroup Identities. Thus, whenever a black and white meet, the tone is always the same; and whenever a black and Hispanic meet, the tone is always the same; of course, the tone of the white-black encounter may differ from that of the black-Hispanic encounter, depending on the distance between the relevant Subgroup Identities.[13]

Suppose now that an individual's preference for residential segregation-integration with others of a particular subgroup is a monotonic decreasing function of the distance between the two Subgroup Identities, so that the larger the distance, the lower the preference for living with persons of the other subgroup. Similarly, suppose that an individual's preference for the proportion from his/her subgroup in the neighborhood is a monotonic increasing function of the distance between the two Subgroup Identities. Then, analysis of the distance between Subgroup Identities generates the homogeneous preferences assigned to agents in homogeneous-type applications of agent-based modeling to residential segregation.

To the question -- Where do the preferences come from? -- the answer is that they come from the operation of the sociobehavioral forces and that the new unified theory of sociobehavioral forces describes how combinations of valued goods, salient qualitative characteristics, and dominant sociobehavioral forces generate specific preferences for segregation and integration.

---

[13] Mathematically, this is an application of the classic Newtonian idea that mass accumulates at a point.

When studying a real-world population, researchers can approximate empirically all the requisite elements. When analyzing a priori the possible span of results, researchers can use a variety of theoretical methods.

For example, consider the case where there are only two subgroups and the two subgroups are nonoverlapping in the valued quantitative characteristic. Suppose that the valued good is wealth, the subgroups are blacks and whites, and the richest black is poorer than the poorest white. Applying tools from probability theory for analyzing censored subdistributions, we can straightforwardly derive all predictions for the case of ordinal valued goods, then choose a few modeling distributions for the cardinal valued goods, and again derive all predictions.

To illustrate, suppose that the status force is dominant. It is straightforward to derive formulas for the Subgroup Identity in each of the two subgroups, for any magnitude of the proportion black (called the subgroup split and denoted p), then to derive the formula for the absolute difference between the two Subgroup Identities. This formula is given by:

$$-\frac{\ln(1-p)}{p}, \tag{22}$$

where, as shown, the distance between the two Subgroup Identities is a function solely of the proportion black p. Thus, the new theory predicts that in this special case of two nonoverlapping subgroups dominated by status, the distance between the two subgroups increases with the proportion in the bottom subgroup. The preference assigned to the agents in an ABM would thus depend on the proportion black.

What about the distance between the two Subgroup Identities in other special cases? Figure 7 depicts graphs under all three sociobehavioral forces, and, for the cardinal-good cases in the comparison and power scenarios, uses the lognormal, Pareto, and power-functions distributions to model the quantitative characteristic. To take the simple ordinal-good cases first, panel A shows that the formulas under the status and comparison scenarios are symmetric about the point where the two subgroups are evenly split ($p = .5$). This is due to the fact, already discussed, that the two outcome distributions are the negative and positive exponentials. Hence,

in the comparison case, the distance between the two subgroups decreases with the proportion in the bottom subgroup. And in the power case, the distance between the two Subgroup Identities is small and constant, impervious to changes in the proportion black. Thus, the character of the underlying society matters a great deal when thinking about distance between Subgroup Identities and exploring plausible preferences to assign to agents in an ABM.

– Figure 7 about here –

And the other panels of Figure 7 indicate still further ways that the underlying society affects the distance between the Subgroup Identities. The two major results are: (1) the distance can be monotonic or nonmonotonic, increasing or decreasing with the subgroup split; and (2) when the valued good is cardinal, the distance is an increasing function of inequality in the good's distribution.

Thus, in this first model of segregation processes, in which the preferences to live with similars or to have certain ratios of similars to others are homogeneous both between and within subgroup, the NUT predicts a lively and diverse set of preferences, reacting to the dominant sociobehavioral forces and the dominant goods and their distributional shape and inequality. Even within this highly restrictive case, the proportion black matters (in all cases except that combining ordinal goods and the power sociobehavioral force), so that varying subgroup proportions is critically important in an ABM, and, of course, varying the preference parameters to reflect all the rich diversity across the possible societies depicted in Figure 7 is as well critically important.

3.3. Second Segregation Model

The first segregation model may be unrealistically restrictive. Segregationist/ integrationist preferences may differ both between and within the subgroups. Moreover, as Schelling (1971:167) understood, some individuals may be subgroup-blind; they may want to go to "a place . . . where color does not matter." To model this scenario, we return to the NUT and the Personal Identity and Subgroup Identity.

Suppose that individuals seek to enhance their identity – by maximizing their score on the

status or comparison or power distribution. Put differently, suppose that individuals seek to maximize their happiness. Then they will compare the Personal Identity and the Subgroup Identity, as discussed in Jasso (2001, 2008). If Personal Identity is lower than Subgroup Identity, they will attach to the subgroup and ignore their Personal Identity, finding comfort and happiness within the confines of their subgroup. If, however, Personal Identity exceeds the Subgroup Identity, they will ignore their subgroup and attach to the idea of themselves and others as individuals, blind to their own and others' subgroup origins.

Accordingly, there will be three emergent subsets: (1) members of the bottom subgroup who are segregationist; (2) members of the top subgroup who are segregationist; and (3) members of both subgroups who are integrationist. In a residential-segregation problem, these subsets predict three types of neighborhoods: (1) all-black; (2) all-white; and (3) mixed.

As Schelling (1971) intuited, the integrationists may be completely blind to the idea of preference for a particular proportion of their similars. The very idea of "similars" is strange and beside the point to someone blind and deaf to race. It was said of St. John of the Cross that he "saw only souls," a thought echoed by Martin Luther King, Jr., who only noticed "the content of our character."

New quantities become important, adding to the proportions in the three main subsets the following quantities: (1) proportions black and white among the integrationists; (2) proportion integrationist among blacks; and (3) proportion integrationist among whites.

It will be no surprise to learn that the new unified theory gives rise to a dazzling diversity of societies, with intricately differing configurations of all the quantities. The proportion black in the overall society continues to be important in most cases. And there is one major surprise: in scenarios involving cardinal goods, inequality has no effect.

To illustrate this second and more elaborate segregation model, we turn to the simplest case, the case of status, for here the distribution of the good is always the same, namely, the rectangular arising from relative ranks. As above, for this a priori illustration, we model the special case in which the two subgroups are nonoverlapping in the valued good. (All the

pertinent formulas are provided in Jasso 2001:123.) Table 3 reports the proportions in each of four subsets – black segregationist, white segregationist, black integrationist, and white integrationist – by proportion black in the society. These four quantities sum to one and are the basic building-blocks for all other quantities of substantive interest. For example, summing the white and black integrationists yields the proportion preferring a mixed neighborhood; and the ratio of black integrationists to all blacks gives the proportion integrationist among blacks. For ease in such algebraic manipulation, each column in Table 3 is assigned a letter label.

– Table 3 about here –

Table 4 presents the corresponding proportions segregationist and integrationist within each racial subgroup. Several results are immediate. First, the proportions segregationist and integrationist in the bottom subgroup vary with the proportion black, but they are constant in the top subgroup. Second, in the bottom subgroup, the proportion integrationist declines as the proportion black increases (and, of course, the proportion segregationist increases). Third, the proportion integrationist among blacks is always larger than the proportion integrationist among whites.

– Table 4 about here –

Table 5 presents a distillation of the major results, including both the predicted proportions in the all-black, all-white, and mixed neighborhoods and also the proportions white and black in the mixed neighborhood. As shown in Table 5 and in Figure 8, the proportion in all-black neighborhoods increases with the proportion black, and the proportion in all-white neighborhoods decreases with the proportion black, but the proportion in mixed neighborhoods is nonmonotonic by proportion black, increasing until blacks are about 63% of the population, and thereafter declining.

– Table 5 about here –

– Figure 8 about here –

Figure 8 also shows that the proportion of the population living in mixed neighborhoods exceeds the proportions in all-black and all-white neighborhoods when the proportion black in

the population is between approximately .36 and .76.

Focusing on the mixed neighborhood, the proportions black and white vary monotonically with proportion black in the population (Table 5). As shown in Figure 9, the proportion black in mixed neighborhoods increases, and the proportion white decreases. The mixed neighborhood has a fifty-fifty composition of blacks and whites when the proportion black in the population is between .435 and .436.

– Figure 9 about here –

The foregoing brief illustration of the second segregation model shows how the new unified theory of sociobehavioral forces generates preferences for residing in all-black, all-white, and mixed neighborhoods and how the ensuing relative size of these neighborhoods and of the racial composition of mixed neighborhoods vary with the proportion black in the overall population. The exact numerical predictions, however, are only for a status society. The predicted preferences will differ substantially in a justice or power society and, within these, by whether the valued good is cardinal or ordinal and, if cardinal, its distributional form.

Moreover, there is a further way in which this brief illustration provides only a glimpse of the theory's segregation predictions. Notice that we have focused only on the proportions in the emergent segregationist and integrationist subgroups and their composition. Further analysis can make use of the individual's predicted attachment to self or subgroup. This new variable is the signed difference between the Personal Identity and the Subgroup Identity. Perhaps even among those "who see only souls" or who notice only "the content of our character," some do so more fully and completely than others. And, thus, many further behaviors – people in segregated neighborhoods going to mixed libraries, people in mixed neighborhoods eating in each other's homes, etc. – may be more deeply understood.

3.4. Incorporating the Underlying Operations of the Two Segregation Models into the Micro-Macro Framework

The goal of a micro-macro framework is to show the links between phenomena at the individual level and phenomena in the social level. Often, the Holy Grail is expressed as

showing how individual characteristics or preferences or behavior generate collective outcomes, in particular, the emergent properties of collectivities. Classically, distinction is made between properties of collectivities which are merely aggregates of individual properties (average age, for example) and the more specifically social properties of collectivities (such as degree of inequality or degree of segregation). It is, of course, the second, more specifically social type of property that we would like to link to phenomena at the individual level.

The two segregation models provide two distinct examples of how phenomena at the individual level generate phenomena at the social level. What is the mysterious alchemy by which this occurs? Let us now dissect the two models.

First Segregation Model. This is a very simple situation. It has one population and begins with one quantitative variable – a personal quantitative characteristic (such as wealth or bravery). By the operation of one of the sociobehavioral forces – justice, status, and power – the personal quantitative characteristic generates a magnitude of the sociobehavioral outcome (say, a magnitude of status) and, concomitantly, a magnitude of the Personal Identity. At this juncture, we are in the case of one population and two quantitative variables, in the change-of-variable operation (Section 2.2.1.2 and Table 2). When we introduce a qualitative characteristic and generate the structure of pre-existing subgroups – the racial subgroups, for example – a new variable is generated, the Subgroup Identity. Finally, for each pair of pre-existing subgroups, a new quantity is generated, namely, the absolute difference between the Subgroup Identities.

The micro-macro framework introduced in Section 2 already accommodates a population with two quantitative variables – which yields the Personal Identity – and a population with one qualitative variable – which yields the structure of pre-existing subgroups. Now, the First Segregation Model illustrates the further operation of combining those two situations – i.e., adding a qualitative variable to a population with two quantitative variables – thus generating the new submacro variable – the Subgroup Identity – and thence the new macro variable – Social Distance between all pairs of pre-existing subgroups.

Thus, the very simple situation represented in the First Segregation Model, while fully

describable by means of the building-blocks in the micro-macro framework, already incorporates combinations and operations beyond the initial set described in Section 2.

Second Segregation Model. This situation continues exactly where the First Segregation Model ends. Now each individual compares the Personal Identity to the Subgroup Identity, generating a new variable – attachment to self or subgroup. This new variable generates a new set of emergent subgroups, those attached to self and those attached to subgroup – called in the theory the Selfistas and the Subgroupistas. And there ensue a new set of macro variables – for example, proportions in all-black, all-white, and mixed neighborhoods – and two new sets of submacro variables – (1) proportions segregationist and integrationist within the pre-existing racial subgroups, and (2) proportions from each pre-existing racial subgroup within the emergent segregationist and integrationist subgroups.

In terms of the micro-macro framework, the Second Segregation Model involves a larger set of operations than does the First Segregation Model. But all the operations use the same set of building-blocks assembled in the framework. As in the sentence diagrams of our childhood, the richest theoretical arguments, with the most intricate operations generating the emergent properties of a collectivity from the characteristics and behavior of individuals, quietly lay bare their underlying mathematical structure.

4. CONCLUDING NOTE

How do individuals shape societies? How do societies shape individuals? The starting idea for this paper was the realization that the underlying mathematical structure of many models and theories which connect micro and macro phenomena can be described in a simple way. Accordingly, we developed a framework which builds on two ingredients widely used in social science -- population and variable. Starting with the simplest case of one population and one variable, we systematically introduced additional variables and additional populations. This approach enables simple and natural introduction and exposition of such operations as pooling, matching, regression, hierarchical and multilevel modeling, calculating summary measures,

finding the distribution of a function of random variables, and choosing between two or more distributions. To illustrate the procedures we drew on problems from a variety of topical domains in social science, including the theories and models which led to the framework.

Use of the framework will no doubt suggest many changes and refinements. For example, it may come to be seen that the distinctions proposed in this paper among different kinds of macro variables (submacro, macro, supermacro) are superfluous; on the other hand, it may come to be seen that yet finer distinctions are necessary, such as between a superpopulation whose units are persons (as in pooling) and a superpopulation whose units are groups (as in matching).

Because intense focus on the framework may obscure the substantive richness and the promise for deeper understanding of micro-macro connections that only theory can provide, we presented an illustration focused on residential racial segregation. We presented two segregation models, both grounded in the recently-proposed new unified theory of sociobehavioral forces and illustrating, first, how the theory generates each individual's segregationist/integrationist preference, and, second, how these preferences generate the emergent segregated and mixed subgroups.

The illustrations showed that many seemingly disparate problems and questions, across substantive domains and levels of aggregation, possess the same or similar underlying mathematical structure, albeit in forms that range from the very simple to elaborate combinations. Thus, new and more fundamental questions about behavioral and social phenomena can be posed; and a single set of tools can be used to address them. The illustrations also showed the ease with which chains of behavior, going back and forth across individual and social levels, can be generated.

Much knowledge will be gained, in all the social sciences, as their essential substantive and methodological unity comes to be appreciated.

REFERENCES

Anselm of Canterbury, Saint. (1070-1109) 1946-1961. Opera Omnia. Edited by F. S. Schmitt, O.S.B. Edinburgh: Thomas Nelson and Sons.

_____. (1080-1090) 1967. Truth, Freedom, and Evil: Three Philosophical Dialogues. Edited and translated by Jasper Hopkins and Herbert Richardson. New York: Harper Torchbooks.

Aristotle. ([384-322 BC] 1952). The Works of Aristotle. 2 volumes. Translated by W. D. Ross. Chicago: Britannica Press.

Brickman, Philip, R. Folger, E. Goode, and Y. Schul. 1981. "Micro and Macro Justice." Pp. 173-202 in M. J. Lerner and S. C. Lerner (eds.), The Justice Motive in Social Behavior. New York: Plenum.

Bruch, Elizabeth E., and Robert D. Mare. 2006. "Neighborhood Choice and Neighborhood Change." American Journal of Sociology 112:667-709.

Crooks, Andrew T. 2008. "Constructing and Implementing an Agent-Based Model of Residential Segregation through Vector GIS." University College London Working Paper # 133.

Gibbons, Jean Dickinson. 1988. "Truncated Data." P. 355 in Samuel Kotz, Norman L. Johnson, and Campbell B. Read (eds.), Encyclopedia of Statistical Sciences, Volume 9. New York: Wiley.

Goldstein, Harvey. 2003. Multilevel Statistical Models. Third edition. London: Edward Arnold. The first and second editions were published in 1987 and 1995, respectively, the first edition under a slightly different title.

Hoel, Paul G. 1971. Introduction to Mathematical Statistics. Fourth Edition. New York: Wiley.

Homans, George Caspar. 1967. The Nature of Social Science. New York: Harcourt, Brace, and World.

_____. 1974. Social Behavior: its Elementary Forms. Rev. ed. New York: Harcourt, Brace,

Jovanovich.

Jasso, Guillermina. 1980. "A New Theory of Distributive Justice." American Sociological Review 45:3-32.

_____. 1983a. "Fairness of Individual Rewards and Fairness of the Reward Distribution: Specifying the Inconsistency Between the Micro and Macro Principles of Justice." Social Psychology Quarterly 46:185-99.

_____. 1983b. "Using the Inverse Distribution Function To Compare Income Distributions and Their Inequality." Research in Social Stratification and Mobility 2:271-306.

_____. 1987. "Choosing a Good: Models Based on the Theory of the Distributive-Justice Force." Advances in Group Processes: Theory and Research 4:67-108.

_____. 1989. "Self-Interest, Distributive Justice, and the Income Distribution: A Theoretical Fragment Based on St. Anselm's Postulate." Social Justice Research 3:251-276.

_____. 1999. "How Much Injustice Is There in the World? Two New Justice Indexes." American Sociological Review 64:133-168.

_____. 2001. "Studying Status: An Integrated Framework." American Sociological Review 66:96-124.

_____. 2008. "A New Unified Theory of Sociobehavioral Forces." European Sociological Review 24:411-434.

_____ and Samuel Kotz. 2007. "A New Continuous Distribution and Two New Families of Distributions Based on the Exponential." Statistica Neerlandica 61:305-328.

_____ and _____. 2008. "Two Types of Inequality: Inequality Between Persons and Inequality Between Subgroups." Sociological Methods and Research 37:31-74.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan. 1994. Continuous Univariate Distributions, Volume 1. New York: Wiley. Revision of Johnson and Kotz' 1970 volume.

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1995. Continuous Univariate Distributions, Volume 2. Second Edition. New York, NY: Wiley. Revision of Johnson and Kotz'

1970 volume.

Kotz, Samuel, Norman L. Johnson, and Campbell B. Read. 1982a. "Censoring." P. 396 in Samuel Kotz, Norman L. Johnson, and Campbell B. Read (eds.), Encyclopedia of Statistical Sciences, Volume 1. New York: Wiley.

Kotz, Samuel, Norman L. Johnson, and Campbell B. Read. 1982b. "Conjugate Ranking." P. 145 in Samuel Kotz, Norman L. Johnson, and Campbell B. Read (eds.). Encyclopedia of Statistical Sciences, Volume 2. New York: Wiley.

Plato. ([c. 428-348/7 BC] 1952). The Dialogues of Plato. Translated by Benjamin Jowett. Chicago: Britannica.

Popper, Karl R. 1963. Conjectures and Refutations: The Growth of Scientific Knowledge. New York: Basic Books.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Second edition. Newbury Park, California: Sage. The first edition, by Bryk and Raudenbush, was published in 1992.

Schelling, Thomas C. 1969. "Models of Segregation." American Economic Review 59:488-493.

_____. 1971. "Dynamic Models of Segregation." Journal of Mathematical Sociology 1:143-186.

Stuart, Alan, and J. Keith Ord. 1987. Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory. Fifth Edition. Originally by Sir Maurice Kendall. New York: Oxford.

Table 1. Distribution of $J = \ln\left(\frac{A}{C}\right)$

A, *C* Variate	Association between *A* and *C*		
	Perfect Positive	**Independent**	**Perfect Negative**
A. Distribution-Independent Results			
Identical	Degenerate at Zero	Symmetric about Zero	Symmetric about Zero
Different	Symmetric/Asymmetric about Any Number		
B. Distribution-Specific Results – *A* and *C* from the same Variate Family			
Rectangular			
Identical	Equal	Laplace	Logistic
Different	---	---	---
Lognormal			
Identical	Equal	Normal	Normal
Different	Normal	Normal	Normal
Pareto			
Identical	Equal	Laplace	Logistic
Different	Positive/Negative Exponential	Asymmetrical Laplace	Quasi-Logistic
Power-Function			
Identical	Equal	Laplace	Logistic
Different	Positive/Negative Exponential	Asymmetrical Laplace	Quasi-Logistic

Note: *J* denotes the justice evaluation, *A* the actual reward, and *C* the just reward.

Table 2. Summary of Micro-Macro Operations, Classified by Number of Populations and Number of Quantitative Micro Variables

Number of Populations	Number of Quantitative Micro Variables		
	One	**Two**	**Several**
One	Summary Measures of Distribution of X	Simple Regression Distribution of a Function of One Random Variable Binary Choice	Multiple Regression Distribution of a Function of Several Random Variables Multivariate Choice
Two	Pooling Matching	Hierarchical and Multilevel Models Generalized Pooling Generalized Matching	Hierarchical and Multilevel Models Generalized Pooling Generalized Matching
Several	Pooling Matching	Hierarchical and Multilevel Models Generalized Pooling Generalized Matching	Hierarchical and Multilevel Models Generalized Pooling Generalized Matching

Notes: In all cells, each micro variable generates macro variables and a subpopulation structure; additionally, each quantitative micro variable generates new rank variables. The problem of finding the distribution of a function of several random variables occurs in many specifications of matching problems, but, for simplicity, is omitted from the six cells involving matching. Moreover, all operations that can be performed on a single population can be performed on the superpopulation generated by pooling or matching.

Table 3. Proportions Black Segregationist, Black Integrationist, White Segregationist, and White Integrationist, by Proportion Black in the Population: Status Force

Proportion Black	Blacks		Whites	
	Segregationist	Integrationist	Segregationist	Integrationist
A	B	C	D	E
.05	.0251	.0249	.600	.349
.10	.0504	.0496	.569	.331
.15	.0760	.0740	.537	.313
.20	.102	.0981	.506	.294
.25	.128	.122	.474	.276
.30	.154	.146	.442	.258
.35	.181	.169	.411	.239
.40	.208	.192	.379	.221
.45	.236	.214	.348	.202
.50	.264	.236	.316	.184
.55	.293	.257	.284	.166
.60	.322	.278	.253	.147
.632	.342	.290	.233	.135
.65	.353	.298	.221	.129
.70	.384	.316	.190	.110
.75	.416	.334	.158	.0920
.80	.450	.350	.126	.0736
.85	.486	.364	.0948	.0552
.90	.525	.375	.0632	.0368
.95	.569	.381	.0316	.0184

Notes: The proportions in each row sum to 1.

Table 4. Proportions of Black and White Subgroups Who Are Segregationist and Integrationist, by Proportion Black in the Population: Status Force

Proportion Black	Blacks		Whites	
	Segregationist	Integrationist	Segregationist	Integrationist
A	B/A	C/A	D/(1-A)	E/(1-A)
.05	.502	.498	.632	.368
.10	.504	.496	.632	.368
.15	.507	.493	.632	.368
.20	.509	.491	.632	.368
.25	.512	.488	.632	.368
.30	.515	.485	.632	.368
.35	.518	.482	.632	.368
.40	.521	.479	.632	.368
.45	.525	.475	.632	.368
.50	.528	.472	.632	.368
.55	.533	.467	.632	.368
.60	.537	.463	.632	.368
.632	.540	.460	.632	.368
.65	.542	.458	.632	.368
.70	.548	.452	.632	.368
.75	.555	.445	.632	.368
.80	.562	.438	.632	.368
.85	.572	.428	.632	.368
.90	.583	.417	.632	.368
.95	.599	.401	.632	.368

Notes: The proportions within each subgroup sum to 1. Within each racial subgroup, the proportion integrationist is always less than half. The proportion integrationist in the black subgroup always exceeds that in the white subgroup.

Table 5. Summary of Predicted Residential Segregation – Proportions Segregated Blacks, Segregated Whites, and Integrated, with Proportion Black and White among the Integrated, by Proportion Black in the Population: Status Force

Prop Black	Segregationist		Integrationist		
	Blacks	Whites	Total	Prop Black	Prop White
A	B	D	C+E	C/(C+E)	E/(C+E)
.05	.0251	.600	.374	.0665	.934
.10	.0504	.569	.381	.130	.870
.15	.0760	.537	.387	.191	.809
.20	.102	.506	.392	.250	.750
.25	.128	.474	.398	.307	.693
.30	.154	.442	.403	.361	.639
.35	.181	.411	.408	.414	.586
.40	.208	.379	.412	.465	.535
.45	.236	.348	.416	.514	.486
.50	.264	.316	.420	.562	.438
.55	.293	.284	.423	.608	.392
.60	.322	.253	.425	.654	.346
.632	.342	.233	.425	.680	.320
.65	.353	.221	.426	.698	.302
.70	.384	.190	.427	.741	.259
.75	.416	.158	.426	.784	.216
.80	.450	.126	.424	.826	.174
.85	.486	.0948	.419	.868	.132
.90	.525	.0632	.412	.911	.0894
.95	.569	.0316	.399	.954	.0461

Notes: The proportions segregationist blacks, segregationist whites, and integrationist total in each row sum to 1. Within the integrationists, the proportions black and white in each row sum to 1.

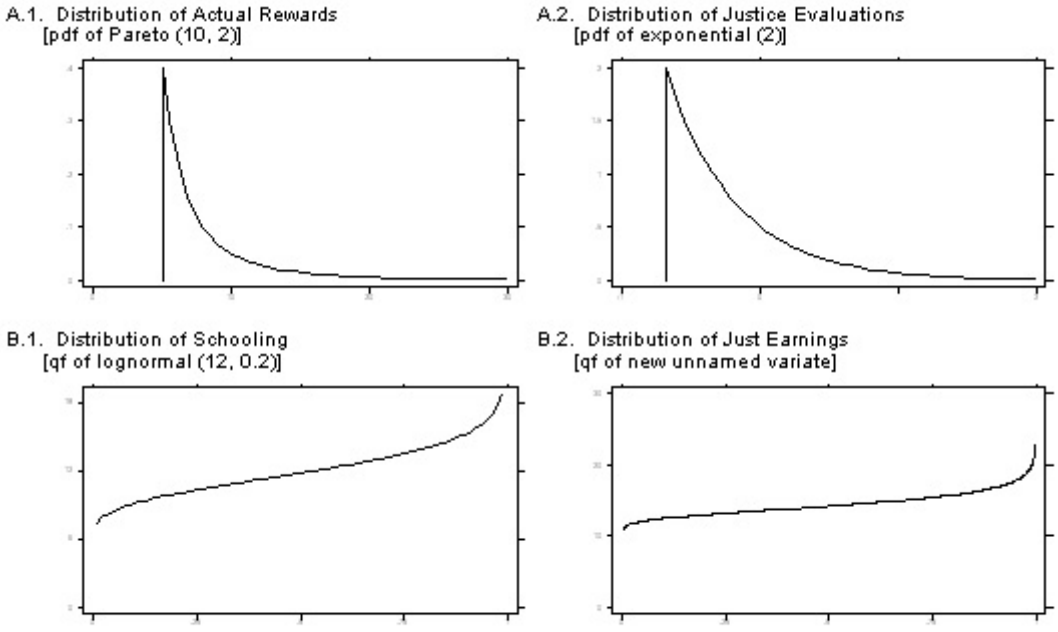

Figure 1. Two Illustrations of Change-of-Variable Problem

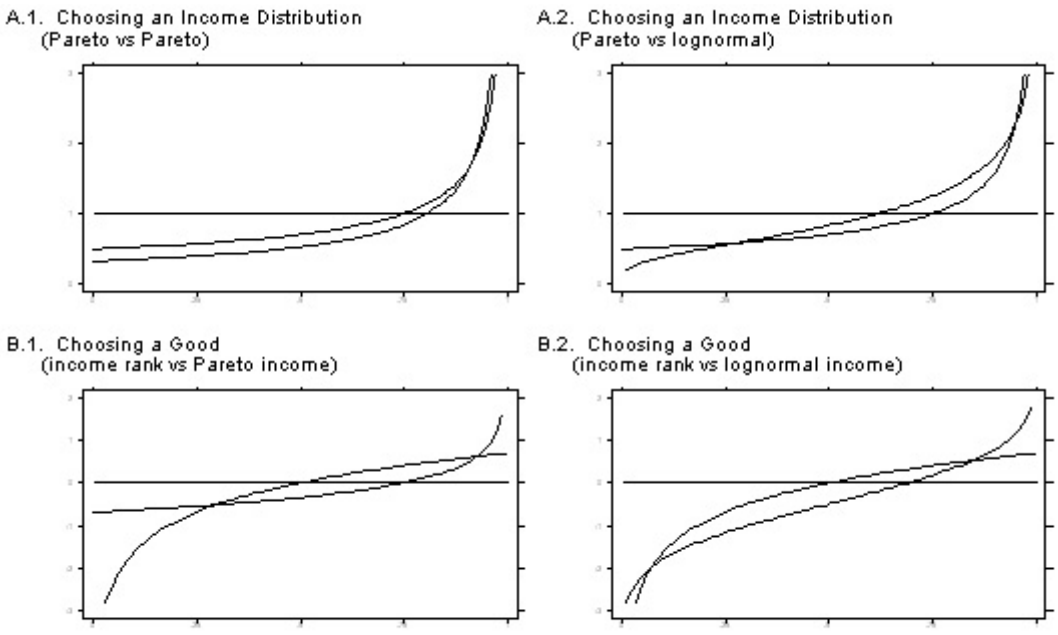

Figure 2. Two Illustrations of Binary Choice Problem

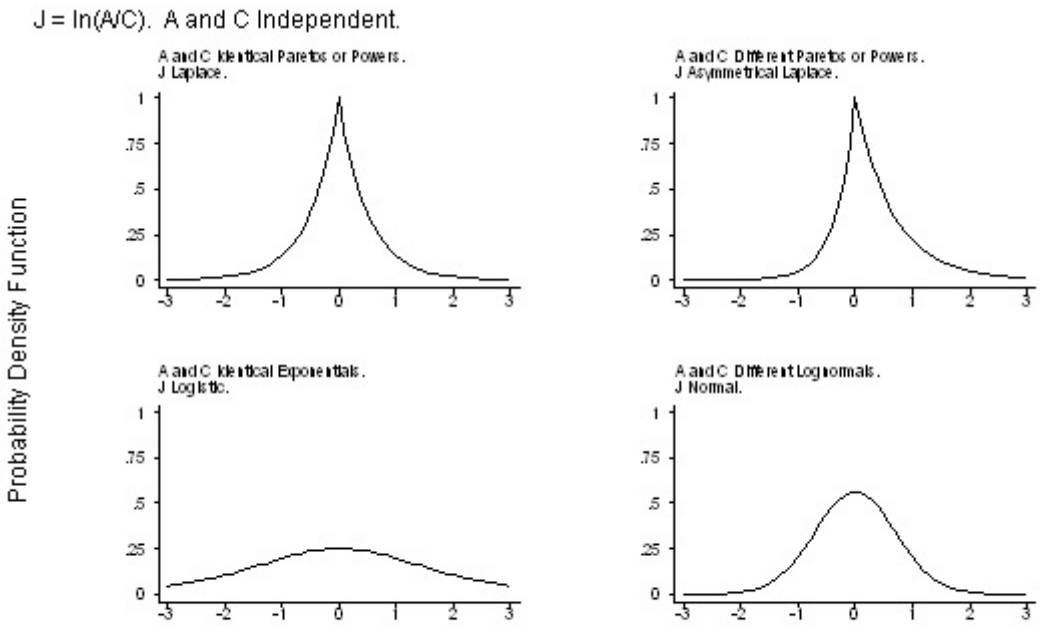

Figure 3. Distributions of Justice Evaluations

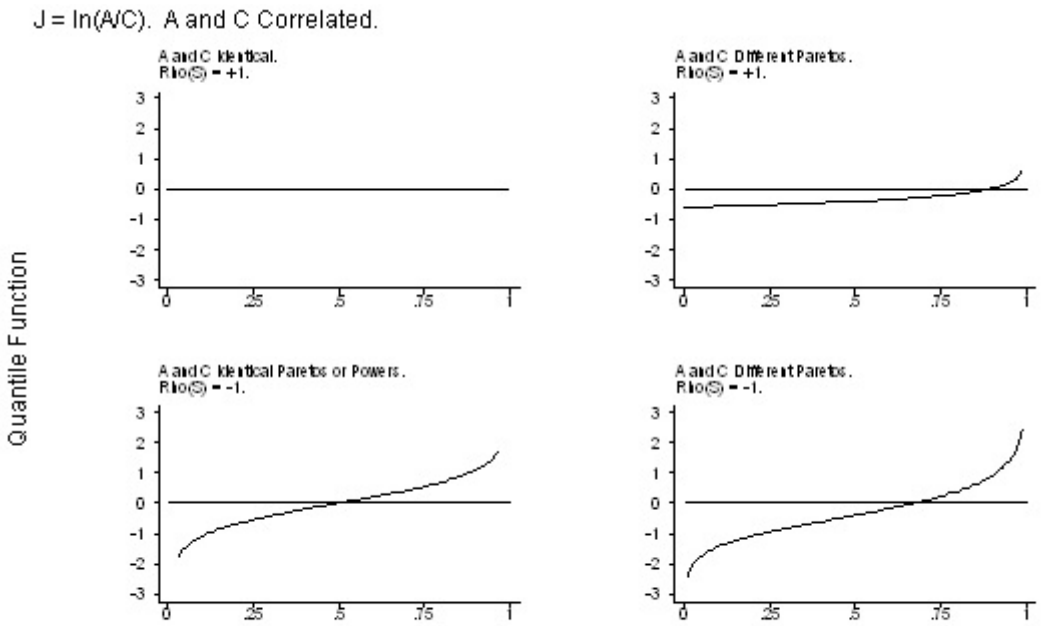

Figure 4. Distributions of Justice Evaluations

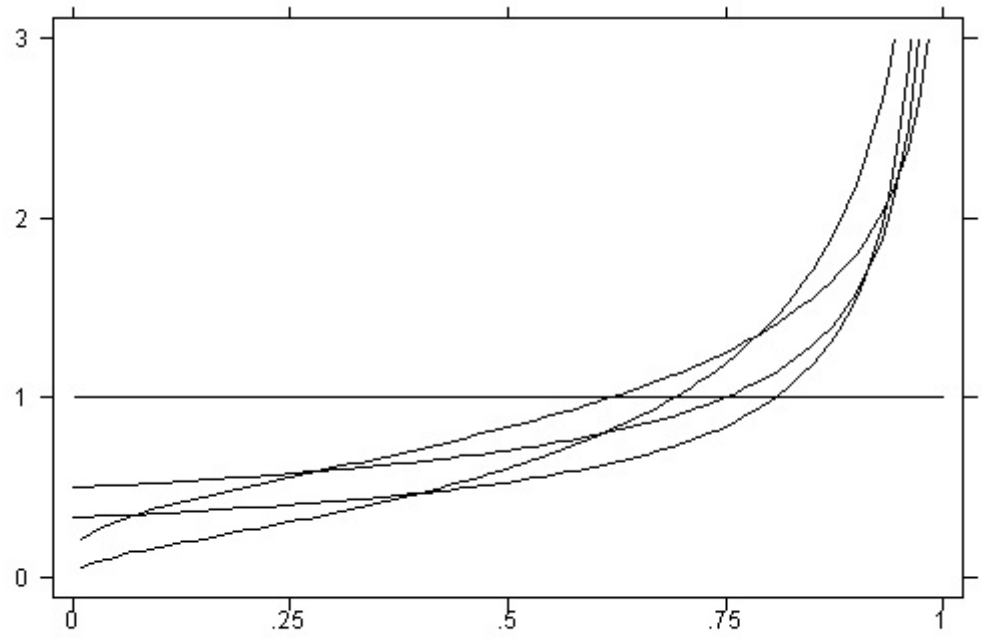

Figure 5. Multivariate Choice Problem

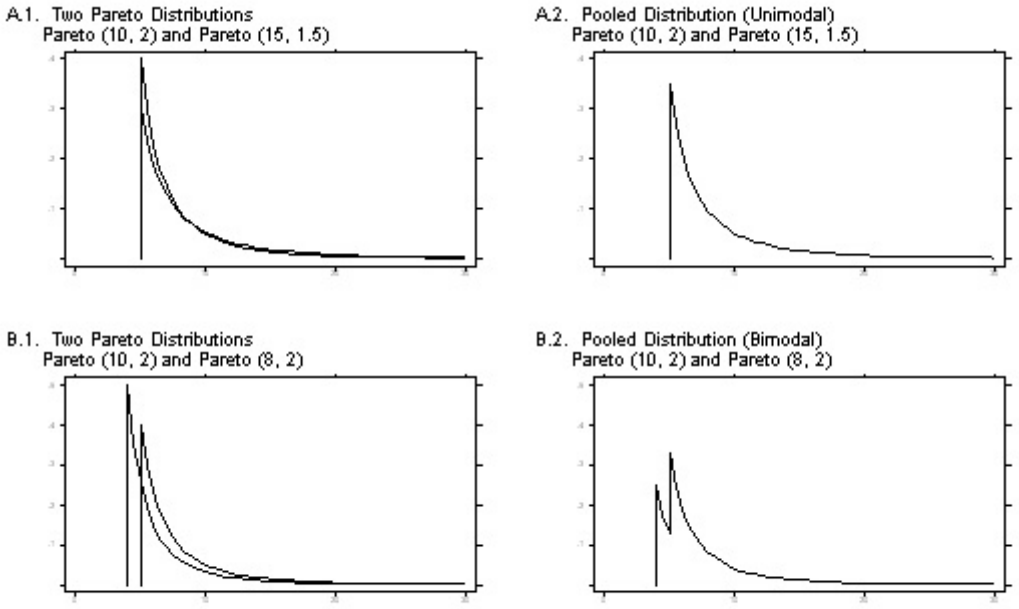

Figure 6. Two Illustrations of Pooling Problem

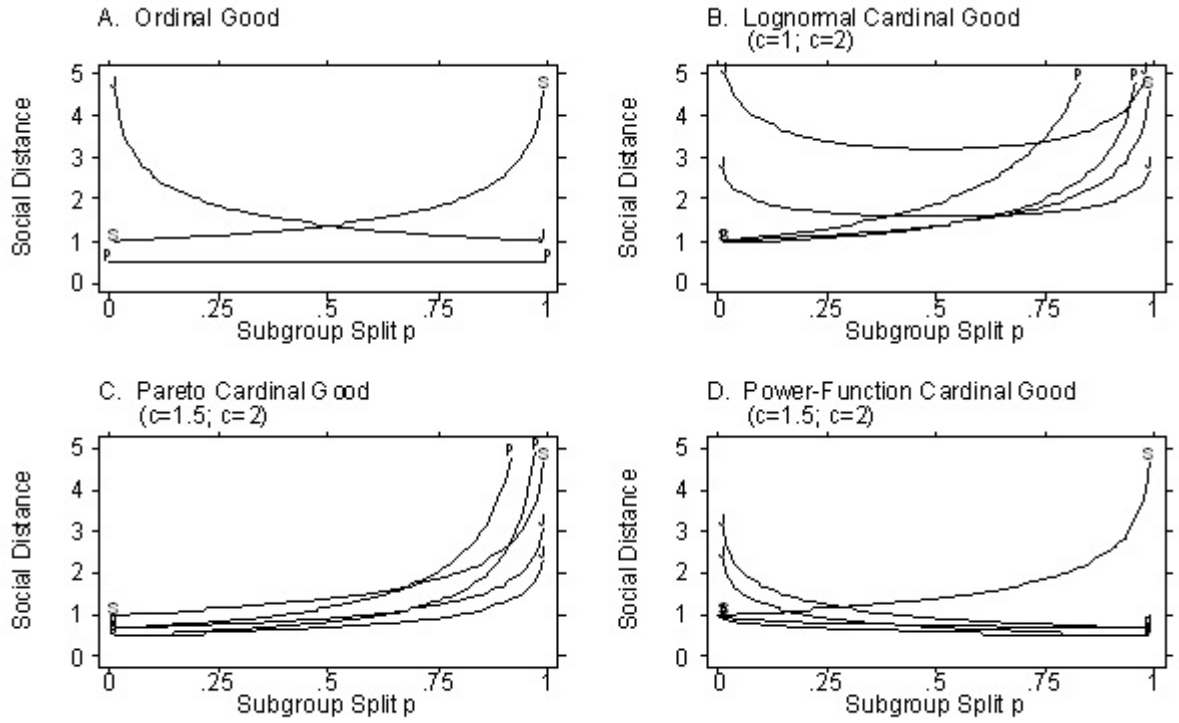

Figure 7. First Segregation Model: Social Distance in Justice, Status, and Power Societies, with Differing Configurations of Valued Goods and Goods' Distributional Forms, by Subgroup Split *p*. The letters J, S, and P denote the justice, status, and power societies, respectively. In the three cardinal-good societies, the greater the inequality in the good's distribution, the greater the social distance (and the higher the vertical placement of the curve).

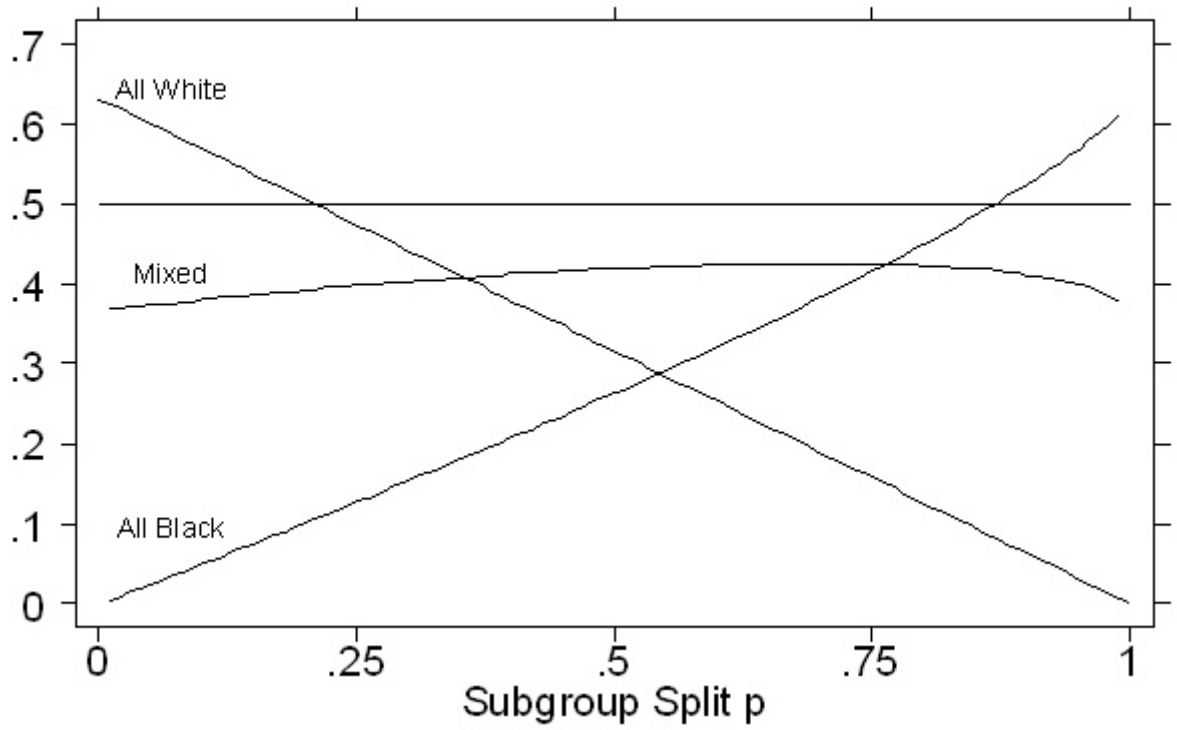

Figure 8. Second Segregation Model: All-Black, All-White, and Mixed Neighborhoods in a Status Society, by Subgroup Split *p*. For each magnitude of the subgroup split (or proportion black), the proportions in all-black, all-white, and mixed neighborhoods sum to one.

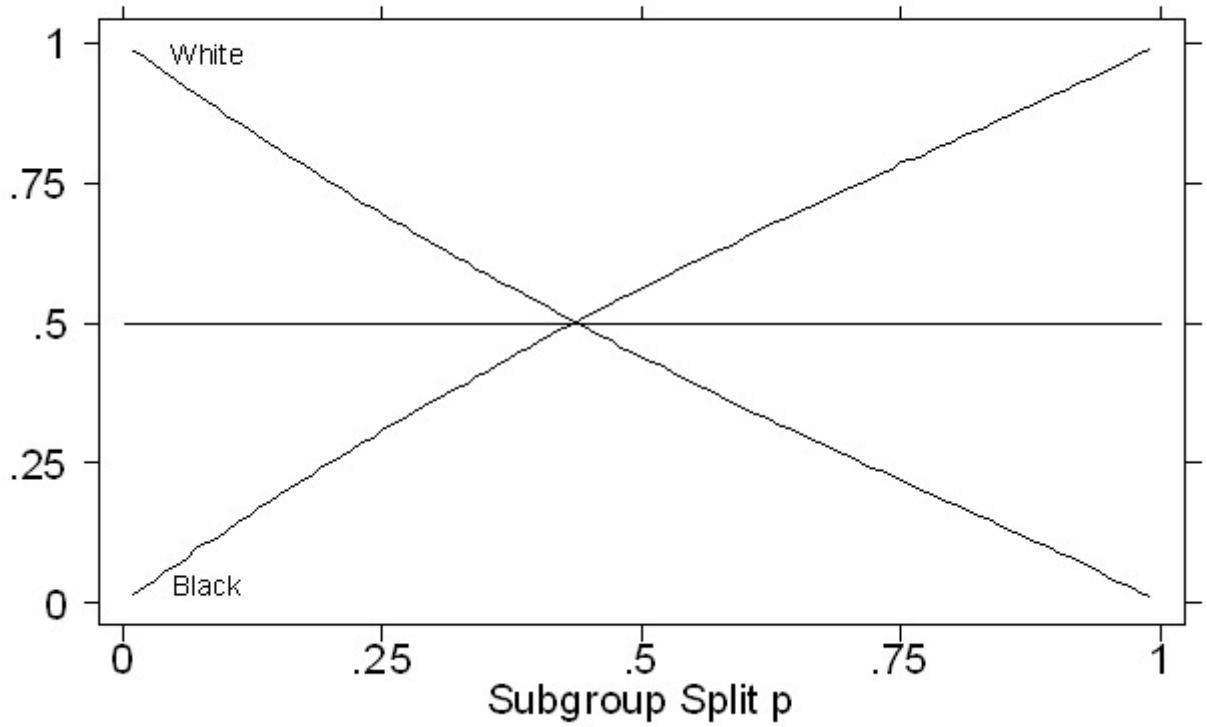

Figure 9. Second Segregation Model: Proportions Black and White in Mixed Neighborhood in Status Society, by Subgroup Split *p*.