Higher Ground: An Exploratory Multivariate Analysis of Characteristics Affecting Population Displacement in the Wake of Hurricane Katrina

By

Jonathan D. Stringfield

Department of Sociology
DePaul University
Chicago, IL

Statistical Research Department
Nielsen Media Research
Schaumburg, IL

For Presentation at the 2008 Annual Meeting of the Population Association of America
New Orleans, LA

Session 29: "Characteristics of Populations Affected by Forced Displacement"
Chaired by Stan Smith
University of Florida

Please direct all questions and comments to Jonathan Stringfield, 150 N. Martingale Road, Schaumburg IL 60173 or e-mail at jonathan.stringfield@gmail.com. Though I am severely indebted to Black Hawk Hancock and Kiljoong Kim of DePaul University and Christine Pierce of Nielsen Media Research for their patience and insightful comments on my work, I must reserve all criticism and blame for errors for myself.

I. Abstract

This study is an exploratory analysis utilizing multivariate techniques upon Current Population Survey data. The objectives are three-fold: First, to examine the relationship between individual and household variables in an evacuee population. Second, identify population factors which have affects on the geographic spread and propensity of return of Katrina evacuees. Third, extract latent variables in an attempt to further explain these differentiations. While acknowledged as standard practice, correlation between individual and household variables has seldom been directly addressed. In addition, though the spatial components of vulnerable populations are well documented the spatial displacement of vulnerable populations is not. These deficiencies are addressed while evaluating the viability of multivariate techniques for assessing populations after disasters. Initial results indicate individual and household variable assumptions are valid and a limited set of variables can discriminate between returned and displaced populations across geographic groupings while meaningful latent variables have been extracted for future analysis

II. Introduction

The devastation inflicted upon the Gulf Coast by Hurricanes Katrina, Rita, and Wilma through August and September of 2005 are one of the most significant events in recent American history and modern American demography. Not since World War II has population movements been so drastically altered in the United States. The magnitude of the 2005 hurricanes has presented a complex problem by which to study populations and social trends which require creative approaches to address limitedly available data. The uniqueness of the population and social landscape post-impact within the Gulf Coast in terms of the magnitude of damage, polarity in socioecomonic status among affected populations, and size of the evacuee disapora necessitates multiplicative approaches for analysis to fully deconstruct the underlying associations, groupings, and latent relationships. In essence, studying demographic trends in the Gulf Coast post-Katrina is an exercise in "abnormal demography[1]" which requires flexibility and adaptability in interpretation, analysis, and data acquisition.

[1] Used here as an extension of the "Demography of Disaster" (Smith 1996, Smith and McCarthy 1996). Where demography of disaster incorporates traditional techniques (survey application, public record) abnormal demography is proposed as a secondary data analysis framework useful for situations where the application of traditional methods is either impossible or too costly for the population of interest. This method entails the application of flexible statistical techniques to extract as much information as feasibly possible from limitedly available or flawed data.

The necessity in utilizing approaches outside of dominant paradigms in population-based research is based in the inadequacy to which public record can capture the vast diaspora of evacuees[2]. Due to the general temporal lag of public records and limitations in the information contained therein, standard public records based demographic methods do not account for migration in the short term – particularly manifested at the scale and suddenness in the wake of Hurricane Katrina. Further, standard sources of demographic information predominately address the pre- and post-impact environments. As such, standard population methods potentially miss factors manifest during intermediate recovery periods crucial to fully understanding the consequences surrounding the population affects of disaster while filling the proverbial and literal gaps between the pre- and post- impact periods. In addition to methodological problems endemic to standard techniques, the majority of population research surrounding Katrina to date has focused upon measuring the population levels of Orleans Parish (primarily) and the remaining impact area (secondarily)[3]. However, analyses concentrating upon this basis do not address the root causes of the displacement – the unequal affects of the disaster experience by populations of varying characteristics in addition to the connotations surrounding the disproportion.

The basis of this work is to add to the collected empirical knowledge and theoretical discussions around Hurricane Katrina as well as challenge the traditional perceptions and methods to which population based disaster research is based. The primary objectives of this work are three-fold: First, the relationship between person-level and household-level variables in an evacuee population in analyzed. While acknowledged as standard practice in standard population based analysis to assume correlation between individual and household variables, this data relationship has been seldom addressed in disaster, social, or population research beyond recent advances in

[2] Partially capture via U.S. Postal Service Delivery Statistics (Plyer and Bonaguro 2007)
[3] In addition to special data releases from the U.S. Census Bureau, The Louisianan Health and Population Survey has been the primary gauge to measure population movement in many Louisiana Parishes.

multi-level analysis[4]. The results of this component will provide useful insight to individual and household relationships in an evacuee population as well as demonstrate aspects of social networks.

The second objective is to examine population factors which appear to have the greatest affect upon the geographic spread of evacuee populations and propensity to return to residences of origin. While analyzing factors salient to differentiating between returned and displaced populations, this study attempts to deconstruct the post-impact[5] spatiality of displaced populations via differentiations of the geographic spread based upon specific population characteristics. Third, it is assumed that underlying relationships affecting displacement and repopulation go beyond race or class dichotomies – variables will be joined into factors representing combined latent relationships and variables as means to base future inquiry towards the social and demographic forces which may provide useful directions and implications for further study into post-disaster population movements.

The objectives will be analyzed via an "abnormal demographic" approach utilizing an array of multivariate techniques: a series of canonical correlations to assess the person and household level variable connection, a series of discriminant analyses to differentiated between returning and displaced evacuee populations by geographic spread, and a series of common factor and principal component analysis to extract latent relationships within the variable structure. Disasters often manifest as population and social constructs, thus, standard population based methods can be challenged via methods usually reserved for social or behavioral analysis to address the problematic of standard techniques. A unification of population and social analysis is presented herein as an attempt to unravel the reciprocal meaning between disciplines and overall understanding of the

[4] See Raudenbush and Bryk (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*, among others.

[5] Research upon populations and disaster tends to focus upon spatially constructed disparate experiences of the disaster agent (Bolin 2006, Tierny 2006, Girard and Peacock 1997, among others). As such, previous population studies after disaster focus upon the *initial* spatial distributions of populations.

post-Katrina Gulf Coast. The usefulness and applicability of these techniques is particularly salient given the conditions surrounding Hurricane Katrina – multivariate statistical techniques are flexible and carry few restrictions (Hair et. al. 2006) enabling the method as potentially ideal for analyzing limited data endemic to disaster situations[6].

In addressing these objectives, this work adds to the ongoing theoretical discussion of vulnerability science[7] in application to population studies by attempting to determine the extent vulnerable populations groups differentiate returned and displaced populations. The socio-political ecology standpoint (Peacock and Ragsdale 1997) is referenced as grounding to the relative isolation of vulnerable groups in metropolitan New Orleans. However, the principally empirical focus of this work will yield useful insight on post-Katrina evacuees and provide solid directions for ongoing research as an exploratory study; and as an evaluation of the relative success of multivariate statistical techniques in disaster related or general population studies.

The implications of this study are numerous. The socioeconomic structure of the Gulf Coast in reaction to the 2005 disasters (as manifest through the population phenomenon) speaks volumes about stratification by race and class within the United States. The vulnerability of vast populations is escalating via increased spatial and structural disparity between population groupings and the increasing global density of population in hurricane vulnerable locations (Clarke 1989). Moreover, population densities and vulnerability to disaster can not simply be specifically targeted to coastal communities – man-made disasters and land-bound natural disasters threaten major populations sectors. Thus, it is imperative to understand population processes in a disaster situation

[6] The focus of this work is additionally limited by more pragmatic concerns. The vast swath of information after the 2005 hurricanes has concentrated upon the New Orleans metropolitan area. This study must also work in the same vein – the basis herein is population affects, accordingly, the greatest population affects happened within metropolitan New Orleans. Further, this study is principally a secondary data analysis – the majority of Katrina-related data focuses upon New Orleans or the surrounding area.

[7] Used here as labeled by Cutter (2003) encompassing many years of work using a "vulnerability approach" as the broad theoretical approach for grounding environmental inequalities and disproportionate affects of hazards or disasters to a (disruptive) part of the ongoing social order rather than framing disasters as unique and temporally limited events (Hewitt 1983).

along with the potential applicability of new analytic techniques for such study as a means to better understand and react to adverse situations, aid the vulnerable populations, and perhaps gain an increased understanding of the stratified structures within the United States.

III. Theoretical Frame and Hypotheses

Population effects are central to our understanding of disasters. All disasters are population oriented – without people present there would be no disaster – in other words "how their lives and activities are imperiled or changed, how they react to crisis, the attitudes they hold, the adjustments they make, and how they confront the everyday problems of risk and vulnerability" (Clark et. al., ed. 1998). Katrina catalyzed one of the largest population movements in U.S. history, releasing a literal and figurative flood of change pent up and compounded after years of stratification and socioeconomic differentiation. Ultimately, the combination of profound poverty along with geographic and spatial polarity of races and classes wrought as a product of "decades of politics and policies that directly or indirectly confined poor households, especially poor black ones, to economically isolate inner-city locales" set the stage for the forthcoming disaster (Brinkley 2002). The aftermath of Katrina is ultimately less about the actual storm and more about the preponderance and continuation of preexisting social and economic factors at play in the region. Population based research is one of the best methods to understand and analyze said factors[8].

The magnitude of the adverse affects in the aftermath of Katrina have been differentially experienced based upon factors such as race and class (Hartman and Squires 2006, Wellman 2005, powell et. al. 2006, Crowley 2005, Logan 2006a, Aptheker 2005, Brown 2005, among others), sex (Jones-Deweever and Hartmann 2006, Enarson et. al. 2006), age (Gullette 2006, Steiner 2005), and relative tenure (Jolivette 2005, White 2005, Frey 2005, Singer and Donato 2005). The vulnerability

[8] Quantitative research within the field would be the primary and much needed alternative method – however, outside of works in progress based upon observations already collected the opportunity for this type of analysis is now gone. Disaster research occupies a distinct temporal window of availability for field data collection (Quarantelli and Dynes 1977, Stallings 2002)

science approach is well equipped to rationalize the likelihood of specific demographic or population factors to mediate the disaster experience and subsequent recovery.

In a disaster situation, three types of vulnerabilities can be addressed: concentrations of energy, concentration of populations (particularly in "risk prone" areas), and concentrations of economic/political power (Perrow 2006). It can be argued then, that the groupings that follow outline specific characteristics of the population that are most vulnerable to disasters based upon one or more of these criterion. As such, to an extent all studies which focus upon population affects after disasters are exercises in vulnerability science and work to (indirectly or directly) identify vulnerable groups. Vulnerability is generated through root causes embedded in the ideological, social, and economic systems tempered by ecological pressures to create specific sets of unsafe conditions which produce effects well beyond the disaster (Oliver-Smith 2006).

Specifically, the vulnerability science perspective argues that disasters stem from the juxtaposition of three factors: the disaster agent itself, the physical setting affected by the disaster (spatial components of proximity hazards or clustering of sub-standard safety features, including housing materials) and population vulnerability - including, but not limited to, material resources and various forms of social and cultural capital as well as involvement in social networks (Tierney 2006). Framing Katrina from this perspective, the catastrophe was triggered by the agent but effects were magnified by failures in protective environmental systems (i.e. the levees) which caused extended damage to the vulnerable populations, as determined by relative social and class positions as well as access to beneficial networks. Morbidity and mortality associated with disasters seems to concentrate upon specific groups, while those with access to superior economic resources not only recover quicker from disaster but also tend to reduce individual vulnerability (Bourque 2006).

Disaster research focusing on populations has "shifted away from the view of disasters as simply physical phenomena to one which sees them more as social issues based on demographic and socio-economic vulnerability." (Clarke et. al. 1998) In terms of the 2005 hurricanes, this coincides well with the apparent plight of the impoverished populations – both in terms of class and spatial positioning. Clarke summarizes this phenomenon:

> "Disasters are always more prevalent where human populations occupy vulnerable positions … Natural and social environments are inherently neither malevolent nor benevolent but largely neutral to their human populations. Primarily it is people who by the nature of their philosophies, attitudes and behaviors modify or transform this environmental neutrality into either a useful resource or a potentially disastrous scenario."

The combination of the vulnerable spatial positioning (i.e. areas below sea level, particularly where populations with less resources available resided), and social characteristics seem to accentuate particular subsets of the population to a greater intensity of negative effects.

The results of specific vulnerabilities are almost always manifested in differential mortality rates. However, mortality in the case of the 2005 storms is not likely to be the most significant differentiated population aspect – a large proportion of the population affects are migratory. Clarke describes these movements in a manner which represents the population movements in the wake of the 2005 hurricanes quite well:

> "Disaster-impelled population movements are exceptional migrations in that they lie outside the pattern of population movements tied to normal life-cycle/life-style considerations. Such movements are sudden, violent, chaotic, largely involuntary and essentially tragic."

Forced migration is a topic central to the population affects in the wake of Katrina. A single disaster can fragment into a series of conflicting circumstance and interpretations. Depending on the individual experience, these circumstances will persuade some individuals to migrate because of disaster and some to remain and rebuild (Oliver-Smith 2006). While most migration after disaster tends to be short term, the extent and strength of ties to family and friends in local communities mediates between the differing types of migration (i.e. permanent or temporary) (Hulbert, Beggs, Haines 2006; Oliver-Smith 2006; Singer and Donato 2005).

However, it would be faulty to assume that the conditions of vulnerability manifest simply by the application of a specific disaster agent or other large-scale disruptions of the social structure.

Quarantelli and Dynes (1977) argue that behaviors and conditions trans- and post disaster are not exclusive results of factors during the emergency period, but rather the "principle of continuity" states that pre-disaster behavior is probably the best indicator of trans- and post-disaster behavior. The distinction is important, insofar that the relative vulnerable or non-vulnerable position of populations affected by disaster is a pre-cursor and likely root cause to disproportionate affects and subsequent recovery, including recovery via insurance or other means (Peacock and Girard 1997). The disaster in this case merely accentuates individual and societal behaviors manifest far before the onset of the disaster, and continue to perpetuate long after.

The vulnerability science approach can be utilized to rationalize the differential effects of Hurricane Katrina between population groupings. However, the approach does not fully encompass rational for differential rates of recovery outside of initial impact. In terms of access to means of recovery and differential population recovery rates Peacock and Ragsdale's seminal work (1997) in applying the socio-political ecology perspective to Hurricane Andrew lends well in application to Katrina. The position draws upon a calling for a broader ecological approach to understanding disaster which goes beyond simply examining the social system and the physical/biological environment to analyzing the interaction of social systems themselves. As such, the community is set not as a single autonomous system but rather a complex ecological network of interacting social systems (Peacock and Ragsdale 1997). Social systems are framed as having autonomous authority structures, however, this is not to say that all interactions and relations between systems are neutral. Power and resource distribution are key in relationship patterns, which determine long-term survival for specific social entities.

The relation of this perspective to the post-impact environment is based upon the inherent competition and conflict between social units – differential access to network resources determine the viability, survival, and reproduction of units. There is almost never a single recovery pattern given the complex interplay and reaction between specific social units. The access to network resources can at best lead to improved housing conditions/recovery and economic standing or at worst spur a decline in socioeconomic status, failure to recover pre-impact housing conditions, and/or an increase in general vulnerability (Peacock and Ragsdale 1997). As such, the reduced access to network resources experience by the lower class and minority populations of southern Dade County after Hurricane Andrew mediated their ability to return and/or recover from unexpected damage. This aforementioned reduced access is primarily attributed to lowered financial resources, social networks (information, temporary housing, etc.), and material resources. Given the disparate population groupings affected by the 2005 storm, the implications of this approach applied to Hurricane Katrina seem to be clear and relevant while conceptualizing the recovery process of the heavily damaged areas.

These perspectives offer varying and poignant directions to the analysis of population disaster phenomenon. The Gulf Coast is an area with long standing disparities in population groupings, which numerous historical case studies indicate that the larger the spatial and economic disparity is between groups the larger the disparity in adverse affects will be. As such, five hypotheses for the population analysis based upon or partially spurred by the theoretical frame can be constructed:

H1: Those who identify as Black or African American race will be less likely to return to their pre-Katrina residence and likely displaced further than other population groups given their predominately disadvantaged class and spatial positioning in metro New Orleans.
H2: The age of the evacuee along with measures of relative socio-economic status (education, income, marital status, labor force participation) will also strongly influence the rate of return and magnitude of displacement.
H3: Household structure in terms of metropolitan living arrangements and location, number of children, and number of people in the household will mediate the diasporas and propensity to return based upon an increased importance or reliance upon infrastructure and availability to beneficial recovery networks.

Hypothesis three relates directly to the fourth hypothesis concerning the first research objective:

H4: The association between individual or person-level variables and household variables is *retained within an evacuee population, or at least be comparable to non-evacuee populations.*

Finally, in relation to the third research objective:

H5: Given that not all of these complex relationships can be fully captured within the manifest data structure, several latent relationships relating to vulnerability or household solidarity should become manifest via common factor or principal component analysis.

These hypotheses should provide a robust exploratory application of multivariate methods in terms of measuring the relative spread and extent of evacuation disparity by populations in the wake of Hurricane Katrina. Given the appropriate intellectual frame and directions for research, pragmatic concerns in relation to the data for study and method of inquiry must be broached.

IV. Methodological Framework and Data Selection

The data to be disseminated will be Current Population Survey (CPS) Basic Monthly data files from August 2006. The CPS data proved to be most receptive to the basic assumptions of multivariate analysis by exhibiting a preponderance of ordinal and continuous variables (some inherent, others requiring recoding), containing an approximate random sample via rotational cluster sampling frameworks and a general preponderance of linear correlation between many variables and variable groups. The date selection of the data is deliberate – it sets the analysis one year after impact and within the time frame of other Katrina related datasets (e.g. LHPS) to allow for relative temporal agreement in parallel or continuing analyses

Initially, the August 2006 file included 154,149 cases and 274 variables. The number of cases was reduced to 518 by restricting the data set to only persons who had actively evacuated their residences due to Hurricane Katrina. Presumably, this reduction will isolate those who were in areas where storm damage and fallout were in actuality or perceived to be most severe. While this severely limits the amount of data available, the distribution of the cases and ratio of evacuee cases to all cases within the file is proportionate to the population of study – Katrina evacuees. However, because derived sample weights may not be fully representative of population characteristics of interest (in particular Black population characteristics[9]) this analysis will focus upon the unweighted sample. While not ideal, the purity of relationships will not be obfuscated by calculated weights and weighting controls – though as a result only particularly strong relationships can be viewed as significant.

CPS data files from November 2006 to December 2006 included 3 variables to identify Hurricane Katrina evacuees. Manipulation of the hurricane screener questions in combination with sundry geographic variables to measure the extent of the displacement form for the focal dependent variables for the discriminate analysis. For the purposes of the discriminate analysis, two independent variables will be analyzed: DSPCLASS and RETURN. DSPCLASS contains 4 values – the first entails persons who had actively evacuated their residence due to Hurricane Katrina and returned – in essence, they represent individuals with residences in the least affect areas or those with the financial ability to reconstruct their residence within a year of landfall. The other three levels are comprised of individuals who have not returned to their pre-evacuation residence – the variability between the groups is the geographic spread.

The second classification variable is, in affect, a simplified version of DSPCLASS. Only two values are present in RETURN: those who actively evacuated and have returned to their pre-evacuation residence and those who actively evacuated and have not returned to their pre-evacuation residence. The logic of the second classification is to isolate phenomenon salient only to differentiating between returned and displaced populations, regardless of the geographic element. Based upon the diminished complexity of this classification variable and the removal of the

[9] While including other racial groups would be viable given the rich cultural diversity of metro-New Orleans, the differentiation between Black and White populations has been of paramount interest to scholars and vehemently discussed in various forms of media. This study is humbly limited to differentiations between White and Black/African American racial groups for data and scope concerns.

(presumably) fine differences in the magnitude of displacement overall better classification results are expected in the discriminate models.

It should be addressed that there is an interpretive flaw in how I have created and conceptualized these variables – if the individual had the ability to return to their previous residence yet decided to move within the various geography levels for any reason other than forced displacement they would be captured within displaced values of DSPCLASS and RETURN. At best, highly valid research implications are still present within this logical flaw (e.g. capturing and classifying individuals who do not wish to return for personal and/or safety reasons). At worst, these values could perhaps classify spurious population movements (e.g. an individual who actively evacuated, and did in fact have means to return to their pre-evacuation residence but before the August 2006 measurement date decided to move to New York for a new job position). It is assumed that the relative magnitude and frequency of this "worst case" group are small and insignificant in magnitude to greatly affect or skew results.

There are additional limitations in using this data. Principal among them are the relatively small sample of positively identified Katrina evacuees and a general exclusion of persons living in shelters, hotels, or other forms of group quarters which may be detrimental to measuring evacuee populations (Logan 2006b). Further, the BLS acknowledged that non-response was higher in affected areas and note that the estimates does not represent the total number of people who had to evacuate due to Hurricane Katrina but rather only those residing in households that are in the scope for the Current Population Survey (CPS 2006). This does, however, concentrate the analysis on potential "best case" evacuees who have at least found semi-permanent household structure rather than more temporary quarters –it is possible results would be more dramatic should "less than best case" evacuees have been included.

The reduced data and constructed classification variables allows for the series of multivariate analysis techniques to be performed. A canonical correlation will determine if lumping household variables and individual variables of respondents remains valid within an evacuee population (half of which has not returned to their previous household). This will test the relative strength of association between household and individual variables in an evacuee population to insure that this basic assumption (e.g. assuming the household characteristics are relevant to the individual characteristics) holds in an evacuee population. For comparison purposes, this canonical correlation will also be applied to a randomly selected sample of individuals who are not indicated as Katrina evacuees to gauge the general magnitude and validity of the findings.

The next layer of analysis entails discriminate models upon the classification variables DSPCLASS and RETURN. The discriminate analysis will attempt to create divisions and classifications that scores group membership by probability of assignment – effectively highlighting the probability of the group assignments which can be speak to the variables which most differentiate between the population groups created through the classification recodes.

The final analysis will be a series of principal component analysis (PCA) and common factor analyses. PCA and factor analysis are variable reduction techniques that tend to highlight latent relationships between variables via orthogonal composite variables. Principal components derived via PCA analysis are linear combinations of the data to approximate the best combinations to reduce the number of loaded variables. Factors extracted from common factor analysis are estimates of latent relationships that are partially measured via the manifest variables. The manner in which the variance is assumed also differs between the two techniques – PCA assumes 100% of variance is accounted for by the loaded components, while factor analysis only assumes shared variance will be manifest within the variables (i.e. not necessarily 100% of variance). A four-step method utilizing orthogonal rotation of the factors is implemented to maximize potential patterns within the eigenvector patterns.

V. Results and Analyses

A series of recode routines constructed two different classification variables: DSPCLASS and RETURN. The frequency distribution of the DSPCLASS and RETURN class variables are outlined in Tables 1 and 2.

Table 1: Frequency Distribution of the "DSPCLASS" Classification Variable

DSPCLASS	Frequency	Weighted Frequency	Unweighted Percent	Cumulative Percent
1 - Returned to Previous Residence	294	808,344	56.76	56.76
2 - Displaced within New Orleans MSA	35	114,982	6.76	63.51
3- Displaced within Impact States (LA, MS, TX, FL)	131	386,429	25.29	88.80
4 - Displaced Abroad	58	144,654	11.20	100.00

Table 2: Frequency Distribution of the "RETURN" Classification Variable

RETURN	Frequency	Weighted Frequency	Unweighted Percent	Cumulative Percent
0 - Displaced, Not Returned	224	646,065	43.24	43.24
1 - Returned to Pre-Katrina Residence	294	808,344	56.76	56.76

Within the data, almost half of the sampling units indicated that they currently resided within their pre-evacuation residence – approximately representative to estimates than only half of residence in the hardest hit areas have returned to their pre-evacuation residence[10]. Within the displaced populations (categories 2, 3, and 4 of DSPCLASS) the populations are predominately located within the impact classified states and near equal parts split between displaced populations within the New Orleans MSA and abroad. Thus, the relative frequency of displaced populations seems consistent with predicted and observed population patterns – likewise, it is assumed these proportions are

[10] As an example, The United States Census Bureau estimated that as of July 2005 (1 month prior to impact) Orleans Parish had an population of 454,863 persons. Recent Louisiana Health and Population Survey Estimates indicate that as of October 2006 the population of Orleans Parish was only 200,665 (9.6% Margin of Error).

representative of the actual population (indicating a preference of proportional priors within the discriminant analyses).

Demographic indicators, traditionally, are often in categorical form – this present challenges for multivariate analyses which assume continuous or ordinal data structure. As such, several variables were manipulated in "indexes" – essentially retaining the same interpretive elements but in a logical/hierarchical order congruent with ordinal variables. The 12 variables also represent a good proportion of variables to cases (approximately 1 : 42).

Table 3: Descriptive Statistics of Demographic Characteristic Variables

Variable	Description	N	Mean	Standard Deviation	Minimum	Maximum	Skew	Kurtosis
LEVEL	Black or AA Racial Identification Index	518	1.477	0.947	1	3	0.802	-0.828
PEEDUCA	Educational Attainment	518	31.840	15.984	-1	46	-1.511	0.441
MI	Marriage Index	518	2.282	1.345	1	4	0.362	-1.680
PEAGE	Age	518	37.629	22.214	0	85	0.223	-0.954
PRNMCHLD	Number of Children	518	-0.187	1.112	-1	4	1.495	1.835
LFP	Labor Force Participation Index	518	3.145	1.646	1	5	0.065	-1.725
HUFAMINC	Household Income	518	7.141	5.992	-3	16	-0.399	-1.065
HRNUMHOU	Number of Persons in Household	518	3.286	1.671	1	9	0.668	0.020
GTCBSAST	Principal City Balance	518	2.349	1.128	1	4	0.281	-1.304
HUP	Housing Unit Permanency Index	518	1.114	0.330	0	2	2.109	3.714
SLCI	Student Level and Commitment Index	518	0.220	0.829	0	4	3.851	13.885
GTCBSASZ	Metropolitan Area (CBSA) Size	518	3.004	2.469	0	7	-0.123	-1.617

The remaining continuous variables were selected via tests of normality and a series of experimental runs, tempered by *a priori* knowledge of potentially significant (practical and statistically) elements. LEVEL, MI, LFP, HUP, and SLCI are all categorical variables recoded into ordinal index-type variables (see Appendix B for details). Each of the index variables represents potentially vital distinguishing population characteristics, and as such, the potential risk of including "un-pure" ordinal variables is outweighed by the explanatory power within the relationships. Most indexes, with the exception of SLCI, seemed to have retained an approximately normal distribution – violations in normality were expected within the SLCI index, which essentially highlights the type and commitment to student activities. Given that many institutions of learning are still closed or running at lowered capability, it is quite feasible to extract a non-normal distribution.

Basic assumptions of normality and data construction seem to be appeased – the correlation matrix was examined to detect colinearity and if a sufficient amount of correlation between the variables is present for principal component and factor analyses. A sizeable amount of significant correlation is seen between the variables – particularly within variables relating to age, number of children, and presence of a spouse (all expected relationships). Little overt-correlation is detected within the correlation structure, thus, it is expected that several of the variables will factor well together with little risk of colinearity.

Table 4: Correlation Matrix of Demographic Characteristic Variables

	LEVEL	PEEDUCA	MI	PEAGE	PRNMCHLD	LFP	HUFAMINC	HRNUMHOU	GTCBSAST	HUP	SLCI	GTCBSASZ
LEVEL	1.000											
PEEDUCA	**-0.126	1.000										
MI	***-0.191	***0.475	1.000									
PEAGE	**-0.147	***0.650	***0.612	1.000								
PRNMCHLD	-0.014	***0.343	***0.593	***0.224	1.000							
LFP	**-0.135	***0.644	***0.326	***0.276	***0.319	1.000						
HUFAMINC	*-0.105	0.082	*0.118	0.015	0.031	**0.158	1.000					
HRNUMHOU	0.046	***-0.358	*-0.117	***-0.467	***0.223	***-0.183	0.046	1.000				
GTCBSAST	*-0.111	-0.037	-0.005	-0.062	0.020	0.041	0.043	0.010	1.000			
HUP	-0.069	-0.053	**-0.129	-0.064	-0.031	-0.027	-0.083	-0.084	***0.241	1.000		
SLCI	0.019	*0.114	***-0.253	***-0.223	***-0.169	0.022	*0.105	0.103	0.062	-0.007	1.000	
GTCBSASZ	0.051	0.030	-0.029	0.026	-0.012	-0.011	0.035	0.006	***-0.762	***-0.228	-0.008	1.000

*Probability > |r| under Ho: Rho=0, ***p<0.0001 **p<0.005 *p<0.05*

Having examined the data for potential violations of multivariate analysis assumptions, the first of the series of analyses is a canonical correlation to test a final assumption within the data.

Several variables indicate household relationships – in particular household income, number of persons in the household, permanency of the housing unit, and of course geographic locality. In demographic analysis, a high amount of correlation and causality is experience between the person level responses and the households; however, the population under analysis here is an evacuee population – approximately half of which is not in fact currently residing within *their original* *household structure*. It is quite likely that a large number of persons are currently residing within a household that would not necessarily "match" their personal demographic characteristics. To test whether it is still as "safe" assumption to include certain household variables with person-level variables in an evacuee population, a canonical correlation is constructed between the "individual" demographic characteristics and the "household" characteristics. It is assumed the linear variates of the canonical correlations will closely approximate the combined statistical relationships between the individual and the household – a strong canonical correlation between the persons and household variables will indicate this assumption is not erroneous.

The canonical variate "IND" is composed of the person-level variables (LEVEL, PEEDUCA, SLCI, MI, LFP, PEAGE), canonical variate "HH" is composed of the household-characteristic variables (HUFAMINC, HRNUMHOU, HUP, PRNMCHLD).

Table 5: Canonical Correlation Measures of Individual Characteristics on Household Characteristics

Canonical Variable	*Canonical Correlation*	*Adjusted Canonical Correlation*	*Approximate Standard Error*	*Squared Canonical Correlation*	*Eigenvalue*	*Proportion*	*F Value*	*Pr > F*
1	0.729	0.724	0.021	0.531	1.134	0.735	27.83	<.0001
2	0.510	0.502	0.033	0.260	0.351	0.228	12.94	<.0001
3	0.209	0.191	0.042	0.044	0.046	0.030	3.63	0.0004
4	0.106	0.095	0.043	0.011	0.011	0.007	1.95	0.1207

While there is not necessarily a directional causality (i.e. an independent vs. dependent relationships between the two sets of variables) this is not required for interpretation of the canonical analysis. Four canonical correlations are derived[11] - results indicate that three canonical correlations can be extracted with high significance (H_0 is the probability level that the current and subsequent canonical correlations have no combined correlation). The high eigenvalues and F Values indicate strong relations – approximately 53.1% of the shared variance is captured by the first canonical correlation, and 26% of the residual variance is captured by the second canonical correlation. The

[11] Only as many canonical correlations can be extracted as number of variables in the smallest grouping – HH contains only 4 variables, and as such, four canonical correlations are constructed.

third canonical correlation explains a practically insignificant level of variation, however, the statistical significance provide adequate proof for further analysis (canonical correlation 4 will not be analyzed given the low practical and statistical significance of the correlation). Correspondingly, multivariate measures of significance indicate sufficient strength within the model (Wilk's Lamda, Pilla's Trace, Roy's Greatest R indicate Pr < F is <0.0001). The linear weights and equations of the extracted canonical correlations provide useful preliminary inference to relationships between the variates.

Table 6: Standardized Weighting Structure of Canonical Correlations

	Description	IND1	IND2	IND3
LEVEL	Black or AA Racial Identification Index	0.141	0.038	**-0.516**
PEEDUCA	Educational Attainment	0.423	-0.099	**-0.626**
SLCI	Student Level and Commitment Index	-0.173	0.105	**0.693**
MI	Marriage Index	**0.727**	**0.956**	-0.058
LFP	Labor Force Participation Index	0.131	-0.026	**0.799**
PEAGE	Age	-0.131	**-1.168**	0.324
	Description	**HH1**	**HH2**	**HH3**
HUFAMINC	Household Income	0.125	0.128	**0.983**
HRNUMHOU	Number of Persons in Household	**-0.520**	**0.885**	-0.084
HUP	Housing Unit Permanency Index	-0.168	0.003	0.210
PRNMCHLD	Number of Children	**0.952**	0.281	-0.156

Standardized coefficients are utilized based upon large discrepancies in scale. For the individual factors, the marriage index seems to weight heaviest for the first canonical correlations, while marriage index and age (in opposite directions) weight heaviest upon the second. The third canonical correlation has nearly every variable weighing heavily, in an apparent preponderance of student level/commitment and labor force participation over the racial index and highest level of school completed. Perhaps the most interesting aspect is the relative lack of importance of the racial index variable until the third canonical correlation – throughout this level of analysis, it becomes apparent that the racial identification of an individual may be of less significant importance within the evacuee populations than the relevant literature of the topic would seem to indicate.

For the household factors, a curious preponderance of having children over number of persons in the household is present within the first canonical correlation. Only number of persons in the household weighs heavily on the second, while household income weighs heaviest upon the last. A redundancy analysis performed upon the canonical correlations indicates little redundant variation within the derived canonical correlations. Further, the cross loadings indicate strong correlations between the marriage index and highest level of school completed to the household variate while number of children and number of persons in the household correlate strongly with the individual variate. In essence, relationships expected to strongly link household to individual factors are validated within the cross-loadings, lending validity to this particular analysis.

For comparison and validation purposes, an equal size (n = 518) sample of the non-evacuee population captured in the August 2006 Basic Monthly Extract was entered into the same model. While slightly stronger results were captured within the non-evacuee population the results are quite comparable. Thus, the analysis continues assuming validity in including household variables with individual variables within the evacuee population.

Among the primary purposes of this multivariate analysis is to determine to what degree, if any, a set of demographic and household characteristics can discriminate between individuals who have returned to their pre-evacuation residence – further, to what degree the geographic displacement can be measured in evacuee populations. A series of discriminant analyses were constructed utilizing the aforementioned variables as potential discriminating elements. All model(s) was tested for assumptions of homogeneous covariance matrixes – all of which were found to be significant via Chi-Square analyses. Thus, the models were demonstrated to be non-parametric and quadratic (rather than Fisher's Linear) discriminant models were utilized.

The first discriminant analysis utilizes all variables with the exception of student level/commitment (removed due to normality assumption violations) and metropolitan area size.

Metropolitan area size seemed to bias classifications towards individuals within the New Orleans MSA – while the principal city balance variable is similar in construction, the inclusion of Principal city components seems to alleviate problems of classifying individuals within a Metro/Non-Metro area. Generalized squared distances between groups were demonstrated to be sufficient for discriminant analysis.

Table 7: Univariate Test Statistics for Discriminating Variables, Model I

Variable	*Description*	*Total Standard Deviation*	*Pooled Standard Deviation*	*Between Standard Deviation*	*R-Square*	*R-Square / (1-RSq)*	*F Value*	*Pr > F*
LEVEL	Black or AA Racial Identification Index	0.9466	0.9348	0.190	0.030	0.031	5.350	0.001
PEEDUCA	Educational Attainment	15.9844	15.9569	1.770	0.009	0.009	1.590	0.190
MI	Marriage Index	1.3447	1.3286	0.266	0.029	0.030	5.190	0.002
PEAGE	Age	22.2138	21.8841	4.801	0.035	0.036	6.230	0.000
PRNMCHLD	Number of Children	1.1116	1.1116	0.097	0.006	0.006	0.990	0.397
LFP	Labor Force Participation Index	1.6457	1.6432	0.178	0.009	0.009	1.510	0.210
HUFAMINC	Household Income	5.9917	5.9617	0.867	0.016	0.016	2.740	0.043
HRNUMHOU	Number of Persons in Household	1.6711	1.653	0.318	0.027	0.028	4.790	0.003
GTCBSAST	Principal City Balance	1.1279	1.0732	0.411	0.100	0.111	19.030	<.0001
HUP	Housing Unit Permanency Index	0.3299	0.3294	0.036	0.009	0.009	1.530	0.207

DF=3, Den DF=514

The Univariate statistics indicate that the racial index, marriage index, age, number of rooms in the household, household income and metropolitan area size have significant discriminating power (H_0 is variables cannot discriminate better than random chance). HUP does not classify as significantly as expected, likely because highly impermanent housing situations are typically not measured within the basic file universe. Multivariate test statistics indicate that the overall model is highly statistically significant (Wilk's Lamda, and others indicate Pr < F is <0.0001) - the overall classification summary indicates successful results.

Table 8: Classification Summary for Full Discriminant Model I

	1 - Returned Previous Residence	2 - Displaced within MSA	3 - Displaced within Impact States	4 - Displaced Abroad	Total
1 - Returned to Previous Residence	226	10	36	22	294
	76.87	3.40	12.24	7.48	100.00
2 - Displaced within MSA	9	18	2	6	35
	25.71	51.43	5.71	17.14	100.00
3- Displaced within Impact States	52	7	58	14	131
	39.69	5.34	44.27	10.69	100.00
4 - Displaced Abroad	18	4	6	30	58
	31.03	6.90	10.34	51.72	100.00
Total	305	39	102	72	518
	58.88	7.53	19.69	13.90	100.00
Error Rate	0.231	0.486	0.557	0.483	**0.359**
Priors	0.568	0.068	0.253	0.112	

Overall error rate for the model is well-acceptable for an exploratory model at 35.9%. Most of the error appears to be spread across the “displaced” groupings (2, 3, and 4) – this is expected given the relatively fine difference between evacuee situations based only upon geographic proximity. However, the error rate for the displaced populations is not so high as to indicate it is unfeasible to classify evacuees in geographic placement based upon a small series of demographic and household indicators.

A comparison sample was drawn from the data using simple random selection[12]. Model I returned overall better results (28.8% error rate and lower error across the displaced categories) that the full sample, though within a threshold acceptable for testing the validity of the model.

A simpler discriminate analysis was constructed to test how well the variables could discriminate strictly between evacuees who had returned to their pre-evacuation residence and those who had

[12] A comparison rather than hold-out was preferred given the relatively small sample size – reducing the sample to fewer than 500 could produce complications in analysis.

not, based upon the relative success of this differentiation in the full discriminate model. Univariate test statistics indicate that the racial index, highest level of school completed, marriage index, age, number of persons in the household, and metropolitan area size discriminate significantly – similar to the full discriminant model with the exception of household income no longer being a significant discriminating factor.

This is potentially rationalized via a potential disconnect of family units due to evacuation-displacement. Regardless, the model demonstrates considerable multivariate significance (Wilk's Lamda, Pilla's Trace, Roy's Greatest R indicate Pr < F is <0.0001) and strong classification results.

Table 9: Classification Summary for Full Discriminant Model II

	0 - Not Returned	**2 - Returned**	**Total**
0 - Not Returned	155 69.20	69 30.80	224 100.00
1 - Returned	72 24.49	222 75.51	294 100.00
Total	227 43.82	291 56.18	518 100.00
Error Rate	0.308	0.245	**0.272**
Priors	0.432	0.568	

Unsurprisingly, the error rate is considerably lower – with error spread approximately equally across the two validation groups. Thus, it is quite feasible to discriminate between Katrina evacuees who have or have not returned to their previous residence based upon a selection of sundry demographic characteristics. This model was run upon the comparison sample for validity purposes, with similar results and a highly comparative error rate (22.9%).

The second model also begins to isolate variables that are most important in determining evacuee populations which have (or have not) returned to their previous residence. In an attempt to further isolate which factors are most important between these population relationships, step-wise variable selection models were constructed for both classification groupings.

Table 10: Stepwise Variable Selection Summary for DSPCLASS and RETURN

DSPCLASS Step	Entered	Label	Partial R-Square	F Value	Pr > F	Wilks' Lambda	Pr < Lambda
1	GTCBSAST	Principal City Balance	0.100	19.03	<.0001	0.900	<.0001
2	PEAGE	Age	0.046	8.30	<.0001	0.858	<.0001
3	HRNUMHOU	Number of Persons in Household	0.026	4.51	0.004	0.836	<.0001
4	LEVEL	Black or AA Racial Identification Index	0.023	3.93	0.009	0.817	<.0001
5	HUFAMINC	Household Income	0.019	3.20	0.023	0.802	<.0001
6	HUP	Housing Unit Permanency Index	0.012	2.07	0.103	0.793	<.0001

RETURN Step	Entered	Label	Partial R-Square	F Value	Pr > F	Wilks' Lambda	Pr < Lambda
1	GTCBSAST	Principal City Balance	0.087	49.14	<.0001	0.913	<.0001
2	PEAGE	Age	0.046	24.83	<.0001	0.871	<.0001
3	HUFAMINC	Household Income	0.007	3.54	0.061	0.865	<.0001
4	LEVEL	Black or AA Racial Identification Index	0.006	2.85	0.092	0.860	<.0001

The stepwise selection reduces the two classification models by approximately half. There is certainly an agreement between the variables selection: metropolitan area size, age, and household income appear in both model selections. Naturally, the majority of selected variables also mimics variables found to have statistically significant discriminatory power in Models I and II.

The step-wise reduced DSPCLASS model returned significant results – as expected, univariate test statistics indicate that most variables discriminate significantly with the exception (once more) of the housing unit permanency (HUP) index – the significance of this variable seems to be high enough for entry within the step-wise selection, but apparently faults in operationalizaiton given the data universe.

The step-wise reduced model returned better overall-classification results based upon a more parsimonious model with an overall error rate of 34.8% . However, the overall results are somewhat deceiving – while the overall error is lower, a disproportionate amount of the error was exhibited in classification grouping 2 at 85.7% (and little improvement in groups 3 and 4 with

55.0% and 55.2% respectively). As such, the step-wise discriminate model appears to be primarily effective only in classifying between evacuees who have and have not returned rather than across the geographic spread. It becomes apparent that a greater amount of variable information is require to produce acceptable classifications across the geographic displacement classifications (the overall model was found to be highly significant, Wilk's Lamda, Pilla's Trace, Roy's Greatest R indicate Pr < F is <0.0001).

Based upon these results, it is expected that the step-wise reduced model for RETURNED will produce acceptable results. Univariate statistics indicate that most variables discriminate significantly – curiously, household income seems to have experienced similar circumstances that HUP fell under within this model insofar that the variable was selected in the step-wise model, but found non-significant in the discriminant model. Likewise, the model exhibited strong overall significance (Wilk's Lamda, Pilla's Trace, Roy's Greatest R indicate Pr < F is <0.0001).

The overall error structure seems to have increased at 31.1% total error with 39.7% associated with not returned and 24.5% associated with returned; however, a 4% increase in classification error is rendered relatively insignificant in comparison to a much more parsimonious model (4 variables instead of 10). Both step-wise models were run upon the comparison sample for validations purposes and found to have similar results to those found within the full sample (28.8% total error with similar distribution across groupings).

Tentative conclusions are once more possible – the step-wise model seems to indicate that a much smaller grouping of variables is required to discriminate between evacuees who have (or have not) returned to their previous residence than evaluating the relative geographic spread of the evacuees. Further, several variables are becoming increasingly clear as to being powerful elements in predicting and classifying individuals within the groupings. However, before definitive conclusions can be extracted a series of principal component and common factor analyses will extract latent relationships and strongly identify central characteristics.

Relationships between the manifest variables have become apparent through the canonical and discriminate analysis – to detect potential latent relationships a series of principal component and common factor analyses will be staged. A four-stage analysis will be utilized to extract the greatest amount of meaning and interpretability from the extracted factors and potential latent relationships: 1) Principal Component, no rotation; 2) Principal Component with Orthogonal Rotation; 3) Common Factor, no rotation 4) Common Factor with Orthogonal Rotation.

By shifting the manner in which the variance is accounted for in addition to rotation of factors it is anticipated that common patterns within the variables should become apparent. The first principal component analysis extracts 12 total factors[13] – 5 of which are extracted for analysis based on the criterion that the first five principal components have eigenvalues greater than 1 (i.e. accounting for more than one variable). Likewise, five factors accounts for 72% of the total shared variance – while a minimum of 80% is ideal, it would require two additional factors of dubious value to reach this criterion. Further, analyses of the associated scree plot does not provide a clear distinction, however, the components appear to level around the fifth factor. Thus, the weighting structure of five components is extracted for analysis.

[13] Principal Component and Factor Analysis will extract as many factors or principal components as variables entered into the model.

Table 11: Eigenvectors of the Principal Component Analysis, no Rotation

Variable	*Description*	*Communality*	*Factor1*	*Factor2*	*Factor3*	*Factor4*	*Factor5*
LEVEL	Black or AA Racial Identification Index	0.679	-0.248	-0.209	0.025	-0.138	**0.744**
PEEDUCA	Educational Attainment	0.868	**0.838**	0.005	-0.172	0.279	0.241
MI	Marriage Index	0.788	**0.799**	0.011	0.302	-0.218	-0.106
PEAGE	Age	0.774	**0.803**	-0.040	-0.290	-0.201	-0.055
PRNMCHLD	Number of Children	0.822	0.552	0.027	**0.668**	-0.162	0.213
LFP	Labor Force Participation Index	0.624	0.658	0.085	0.031	0.377	0.201
HUFAMINC	Household Income	0.584	0.145	0.012	0.219	**0.551**	-0.460
HRNUMHOU	Number of Persons in Household	0.804	-0.389	-0.037	**0.802**	0.085	0.022
GTCBSAST	Principal City Balance	0.834	-0.037	**0.909**	0.073	0.026	0.021
HUP	Housing Unit Permanency Index	0.311	-0.105	0.479	-0.223	-0.136	0.043
SLCI	Student Level and Commitment Index	0.740	-0.197	0.063	-0.078	**0.781**	0.286
GTCBSASZ	Metropolitan Area (CBSA) Size	0.816	0.026	**-0.893**	-0.048	0.083	-0.088
	Total Communality	**8.644**					

The communality of the variables appears to be appropriate, with the exception of the Housing Unit Permanency Index (HUP) and potentially household income. Likewise, HUP does not appear to attach to any factors. The interpretation of the factors appears to be relative clear – the first factor seems to indicate Maturity (or human capital) and Family Strength Factors. The second factor indicates Urban Identifiers; the third factor indicates Household Size and Structure. The fourth factor seems to indicate a propensity for both Student status and income while the fifth indicates racial identification. Principal components and factors are often ordered in terms of importance to summarizing latent relationships – it is interesting that social ties and status seems to be the most important in the first factor while racial identification and pure economic indicators occupy the last factors. The principal components are rotated via orthogonal rotation to attempt to extract additional meaning.

Table 12: Eigenvectors of the Principal Component Analysis, Orthogonal Rotation

Variable	*Description*	*Communality*	*Factor1*	*Factor2*	*Factor3*	*Factor4*	*Factor5*
LEVEL	Black or AA Racial Identification Index	0.679	-0.048	-0.116	0.143	0.202	**0.776**
PEEDUCA	Educational Attainment	0.868	**0.898**	-0.049	-0.219	0.104	-0.031
MI	Marriage Index	0.788	**0.686**	-0.032	0.178	-0.504	-0.175
PEAGE	Age	0.774	**0.671**	-0.063	-0.397	-0.397	-0.064
PRNMCHLD	Number of Children	0.822	0.583	0.008	**0.612**	-0.318	0.078
LFP	Labor Force Participation Index	0.624	**0.755**	0.021	0.002	0.206	-0.110
HUFAMINC	Household Income	0.584	0.123	-0.099	0.165	0.251	**-0.684**
HRNUMHOU	Number of Persons in Household	0.804	-0.295	-0.044	**0.840**	0.091	-0.037
GTCBSAST	Principal City Balance	0.834	0.031	**0.898**	0.095	0.068	-0.114
HUP	Housing Unit Permanency Index	0.311	-0.099	0.504	-0.201	-0.025	0.076
SLCI	Student Level and Commitment Index	0.740	0.058	0.020	0.027	**0.857**	-0.042
GTCBSASZ	Metropolitan Area (CBSA) Size	0.816	-0.036	**-0.900**	-0.073	0.005	0.006
	Total Communality	**8.644**					

The meaning of the first factor appears to have shifted slightly – the original variables are joined with the labor force participation (LFP) index – potentially expounding upon a measure of economic dependence previously manifest through education, marriage (dual income), and age (labor force longevity). The second and third factors did not change composition from the rotations, while the fourth factor now seems to completely indicate student level and commitment level. The fifth factor is now completely composed of racial identification. The orthogonal rotation seemed to have solidified the meaning of the factors; however, it would be ideal for more variables to attach to each factor. Thus, a sequence of common factor analyses is staged. The differentiation of the variance indicates a different set of eigenvalues of the correlation matrix.

Once more, 12 common factors are extracted – only two adhered to the previous selection criterion with Eigenvalues greater than one. However, there appears to be large difference in proportion from the 4^{th} to 5^{th} factors, thus, it is possible a differentiation in values occurs between these levels. Additional analysis of the scree plot supports the decision to extract the four common factors. The results seem to have been slightly diminished by the common factors. HUP retains low communality and is joined by the racial index, correspondingly, neither attach to any of the four factors. Further, the general communality (and Total Communality) dropped significantly.

Table 13: Eigenvectors of the Common Factor Analysis, no Rotation

Variable	Description	Communality	Factor1	Factor2	Factor3	Factor4
LEVEL	Black or AA Racial Identification Index	0.049	-0.184	-0.120	0.025	-0.014
PEEDUCA	Educational Attainment	0.804	**0.820**	0.002	-0.202	0.301
MI	Marriage Index	0.699	**0.749**	0.027	0.309	-0.205
PEAGE	Age	0.740	**0.773**	-0.040	-0.240	-0.289
PRNMCHLD	Number of Children	0.589	0.500	0.043	**0.581**	0.017
LFP	Labor Force Participation Index	0.497	0.587	0.072	-0.018	**0.382**
HUFAMINC	Household Income	0.059	0.106	0.011	0.063	0.209
HRNUMHOU	Number of Persons in Household	0.464	-0.337	-0.010	**0.567**	0.169
GTCBSAST	Principal City Balance	0.690	-0.037	**0.829**	0.006	0.016
HUP	Housing Unit Permanency Index	0.104	-0.081	0.289	-0.109	-0.049
SLCI	Student Level and Commitment Index	0.276	-0.157	0.043	-0.130	**0.482**
GTCBSASZ	Metropolitan Area (CBSA) Size	0.669	0.025	**-0.814**	0.001	0.070
	Total Communality	**5.639**				

However, all factors through factor 3 retain the same meanings as extrapolated from the initial principal component analysis. Factor 4 seems to entail a preponderance of student status and labor force participation – indicating an involvement to outside-household commitments. Rotation of the factors does little to improve the interpretation and generally obfuscated the results.

Clearest results and highest total communalities were extracted via the principal component method. With the exception of the HUP variable and a relatively low attachment of variables to factors, it appears that most latent relationships can be obtained via this method. To partially compensate for this problem, the un-rotated principal component analysis is run again with the HUP variable dropped in an attempt to better clarify the results. The eigenvalues of the correlation matrix changed very little – the results and associated scree plot once more indicate five components to be extracted. Ultimately, the rotated factors of the cleansed principal component analysis provided the best results. The final principal component analysis outlines the strongest weights to each factor and the clearest interpretations. The first factor seems to clearly indicate a relative strength of the family unit – income, spousal proximity, number of children, age, and participation in the labor force.

Table 14: Eigenvectors of the Cleansed Principal Component Analysis, Orthogonal Rotation

Variable	*Description*	*Communality*	*Factor1*	*Factor2*	*Factor3*	*Factor4*	*Factor5*
LEVEL	Black or AA Racial Identification Index	0.689	-0.035	0.105	0.139	0.197	**0.787**
PEEDUCA	Educational Attainment	0.868	**0.881**	0.037	-0.278	0.111	-0.037
MI	Marriage Index	0.782	**0.695**	-0.016	0.105	-0.510	-0.168
PEAGE	Age	0.780	**0.638**	0.052	-0.464	-0.389	-0.059
PRNMCHLD	Number of Children	0.847	**0.634**	-0.029	0.578	-0.326	0.069
LFP	Labor Force Participation Index	0.630	**0.755**	-0.033	-0.039	0.210	-0.119
HUFAMINC	Household Income	0.568	0.138	0.060	0.162	0.245	**-0.678**
HRNUMHOU	Number of Persons in Household	0.813	-0.228	0.005	**0.869**	0.072	-0.039
GTCBSAST	Principal City Balance	0.881	-0.001	**-0.934**	0.021	0.054	-0.080
SLCI	Student Level and Commitment Index	0.740	0.059	-0.033	0.041	**0.856**	-0.044
GTCBSASZ	Metropolitan Area (CBSA) Size	0.882	-0.001	**0.938**	0.009	0.020	-0.030
	Total Communality	**8.482**					

The second factor remains and indicator of urban geography – once more exhibiting a preponderance of principal city balance over raw MSA/CBSA size. The final three factors are disappointing insofar that only a single variable attaches to each component, however, the interpretation and implications are relatively clear. Factor 3 indicates household size, while 4 indicate student concerns and 5 indicates racial identification over income. The last factor is interesting insofar there seems to now be preponderance of racial identification over income – once more the general interplay of these often relied upon phenomenon in exploring post-Katrina population patterns is clouded.

The outcome of the principal component and common factor analyses is relatively clear: within evacuee populations, the relative strength (both social and economic) of the family unit appears to be the strongest latent relationship. Further, many of these "family strength" variables were significant in the previous canonical correlations and discriminant analysis. Likewise, racial identification seems to have limited importance as a latent relationships through the principal component and factor analysis (along with canonical correlation, though LEVEL did have significant discriminatory power).

To test the relative explanatory power of the latent relationships, the cleansed and orthogonally rotated principal components were entered into a discriminant model against DSPCLASS and RETURN. The results were not as expected – neither model produced particularly strong results, or improved upon previous discriminant models. However, the significance of the factors in the discriminant analysis displayed an interesting pattern. With only one exception all factors were found to have significant discriminating power – the exception was factor 5 in discriminating between return and displaced populations. Essentially, this seems to indicate that the latent relationships between racial identification and household/family income have no ability to differentiate between evacuees who had or had not returned to their pre-evacuation residence. This is an unexpected and interesting result which warrants future analysis.

VI. Conclusions

As an exploratory analysis, this study exhibits the usefulness of multivariate methods in assessing population displacement. Specifically, these methods tested essential assumptions and differentiated between spatial dimensions of forced migration patterns. Several clear and potentially meaningful population patterns have emerged and presented strong directions for future analysis which may further validate initial findings within this work. First, it was determined that the household to characteristic variable link, often taken for granted, was still valid in the evacuee population (validating hypothesis 4). The discriminant analysis indicated that a small set of population characteristics (as few as four) could successfully distinguish between evacuees who had returned to their pre-evacuation residence and those who had not; in addition to relatively limited success in differentiating non-returned populations through three levels of geography.

A step-wise variable selection technique found that a reduced and parsimonious model could still be effective in discriminating between returned and displaced populations – but tended to obfuscate the distinction between geographic levels. Further, this particular analysis seemed to indicate that the age, household income, racial identification, and urban geographic identification of the respondent were the most important aspects in differentiating between the returned populations - validating hypotheses 2, 3, and to a lesser extent 1. While this seems to agree assumptions outlined in an analysis of relevant literature, the limitations of some of these variables require further testing.

A series of principal component and factor analysis produced several significant latent relationships within the evacuee population – principal amongst them appeared to be a gauge of relative social and economic strength of the family unit, followed by urban geographic identification. Clearly, the relative unity of the family unit is a strong latent relationship in the evacuee population requiring further analysis (though containing relatively little discriminating power). The relative importance of racial identification was once again called into question – where the variable attached to factors, if at all, was always on the last (and relatively least important). In addition, the latent relationship of racial identification over family income proved to have no discriminating power – an interesting result that requires additional insight. Overall, hypothesis 5 was also realized given the success of this layer of analysis.

It becomes clear that some central relationships and focal points are apparent from the multivariate analyses. Further testing with more robust and inferential techniques (e.g. logistic regression) might yield additional and useful information as to the relative population spread of Katrina evacuees. Further, some data issues might be eliminated via a merge of several months of CPS data. The limitations of this analysis have mostly been addressed – the index type variables may be suspect in relative validity, however, it was deemed necessary given the nature of the data and the nature of the relationships to be explored. As with all Katrina-related data, it is difficult to determine exactly how well the CPS data represents the relative evacuee populations. Thus, future analysis utilizing different data is warranted to insure that actual population trends are being captured rather than abnormalities within the data. However, this proposed work of “abnormal demography” indicates that non-traditional methods can yield useful results in population research.

Appendix A: Works Cited and Consulted

Annual Population Estimates, Vintage 2005 and 2006 Estimates, United States Bureau of the Census. Retrieved from the Annual Population Estimates website: http://www.census.gov/popest/estimates.php

Aptheker, Bettina. 2005. "Katrina and Social Justice." Pg. 48-56 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.

Bolin, Bob. 2006. "Race, Class, Ethnicity, and Disaster Vulnerability" Pg. 113-129 in *Handbook of Disaster Research*, edited by Havidán Rodriguez, Enrico L. Quarantelli, and Russell R. Dynes. New York, NY: Springer Science + Business Media, LLC.

Bourque, Linda B., Judith M. Seigel, Megumi Kano, and Michele M. Wood. 2006. "Morbidity and Morality Associated with Disaster" Pg. 97-112 in *Handbook of Disaster Research*, edited by Havidán Rodriguez, Enrico L. Quarantelli, and Russell R. Dynes. New York, NY: Springer Science + Business Media, LLC.

Brinkley, Douglas. 2006. *The Great Deluge: Hurricane Katrina, New Orleans, and the Mississippi Gulf Coast*. New York, NY: HarperCollins.

Brookings Institution. 2005. "New Orleans After the Strom: Lessons from the Past, a Plan for the Future." Metropolitan Policy Program, The Brookings Institution. Washington, DC.

Brown, Michael K. 2005. "Katrina's Aftermath." Pg. 95-99 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.

Clarke, John I., Peter Curson, S.L. Kayastha, and Prithvish Nag, Editors. 1989. *Population and Disaster*. Oxford, UK: Basil Blackwell Ltd.

Crowley, Sheila. 2006. "Where is Home? Housing for Low-Income People After the 2005 Hurricanes." Pg. 121-166 in *There is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina,* edited by Chester Hartman and Gregory D. Squires. New York, NY: Routledge.

Current Population Survey, August 2006 Basic Monthly Extract. Retrieved from Current Population Survey FTP site: http://www.bls.census.gov/ferretftp.htm

Cutter, S.L. 2003. "The Science of Vulnerability and the Vulnerability of Science" *Annals of the Association of American Geographers* 93.1, pp. 1-12.

Enarson, Elaine, Alice Fothergill, and Lori Peek. 2006. "Gender and Disaster: Foundations and Directions." Pg. 130-146 in *Handbook of Disaster Research*, edited by Havidán Rodriguez, Enrico L. Quarantelli, and Russell R. Dynes. New York, NY: Springer Science + Business Media, LLC.

Frey, William H. 2005. "Planning for New Orleans' Demographic Future." Discussion Paper and Lecture. Metropolitan Policy Program, The Brookings Institution, Washington DC.

Girard, Chris, and Walter Gillis Peacock. 1997. "Ethnicity and Segregation: Post-Hurricane Relocation" Pg. 191-203 in *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*. New York, NY: Routledge.

Gullette, Margaret M. 2006. "Katrina and the Politics of Later Life." Pg. 103-120 in *There is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina,* edited by Chester Hartman and Gregory D. Squires. New York, NY: Routledge.

Hair, Joseph F. Jr., William C. Black, Barry J. Babin, Rolph E. Anderson, and Ronald L. Tatham, Editors. 2006. *Multivariate Data Analysis*, Sixth Edition. Upper Saddle River, NJ: Pearson Prentice Hall.

Hartman, Chester and Gregory D. Squires, Editors. 2006. *There is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina*. New York, NY: Routledge.

Hewitt, K. Editor. 1983. *Interpretations of Calamity from the Viewpoint of Human Ecology.* Winchester, MA: Allen & Unwin.

Hulbert, Jeanne S., John J. Beggs, and Valerie A. Haines. 2006. "Bridges Over Trouble Water: What are the Optimal Networks For Katrina's Victims?" SSRC, Retrieved July 12th 2006 (https://understandingkatrina.ssrc.org/Hulbert_Beggs_Haines/printable.html).

Jones-DeWeever, Avis A. and Heidi Hartmann. 2006. "Abandoned Before the Storms: The Glaring Disaster of Gender, Race, and Class Disparities in the Gulf." Pg. 85-102 in *There is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina,* edited by Chester Hartman and Gregory D. Squires. New York, NY: Routledge.

Jolivette, Andrew. 2005. "Displacement, Gentrification, and the Politics of Exclusion." Pg. 30-32 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.

Logan, John R. 2006. "The Impact of Katrina: Race and Class in Storm Damaged Neighborhoods." Discussion Paper. Department of Spatial Structures in the Social Sciences, Brown University, Providence, RI.

Logan, John R. 2006. "Population Displacement and Post-Katrina Politics: The New Orleans Primary" Discussion Paper. American Communities Project, Brown University, Providence, RI.
Oliver-Smith, Anthony. 2006. "Disasters and Forced Migration in the 21st Century" SSRC, Retrieved July 12th 2006 (https://understandingkatrina.ssrc.org/Oliver-Smith/printable.html).
Peacock, Walter Gillis, and Chris Girard. 1997. "Ethnic and Racial Inequalities in Hurricane Damage and Insurance Settlements" Pg. 171-189 in *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*. New York, NY: Routledge.
Peacock, Walter Gillis, and A. Kathleen Ragsdale. 1997. "Social Systems, Ecological Networks and Disasters: Toward as Socio-Political Ecology of Disasters" Pg. 20-34 in *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*. New York, NY: Routledge.
Perrow, Charles. 2006. "Disasters Ever More? Reducing U.S. Vulnerabilities" Pg. 521-533 in *Handbook of Disaster Research*, edited by Havidán Rodriguez, Enrico L. Quarantelli, and Russell R. Dynes. New York, NY: Springer Science + Business Media, LLC.
Plyer, Allison and Joy Bonaguro. 2007. "Using U.S. Postal Service Delivery Statistics to Track the Repopulation of New Orleans & the Metropolitan Area" Research Note. Greater New Orleans Community Data Center.
powell, john a., Hasan Kwame Jeffries, Daniel W. Newhart, Eric Steins. 2006. "Towards a Transformative View of Race: The Crisis and Opportunity of Katrina." Pg. 59-84 in *There is No Such Thing as a Natural Disaster: Race, Class, and Hurricane Katrina,* edited by Chester Hartman and Gregory D. Squires. New York, NY: Routledge.
Quarantelli, E. L. and Russell R. Dynes. 1977. "Response to Social Crisis and Disaster" *Annual Review of Sociology*, Vol. 3, pp. 23-49.
Rodriquez, Havidán, Enrico L. Quarantelli, and Russell R. Dynes, Editors. 2006. *Handbook of Disaster Research.* New York, NY: Springer Science + Business Media, LLC.
Raudenbush, Stephen W., and Anthony S. Burk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications.
Singer, Audrey and Katharine M. Donato. 2005. "In Katrina's Wake, Who Will Return?" Discussion Paper. Metropolitan Policy Program, The Brookings Institution, Washington DC.
Smith, Stanley K. 1996. "Demography of Disaster: Population Estimates After Hurricane Andrew" *Population Research and Policy Review* Vol 15, pp.459-477.
Smith, Stanley and Christopher McCarty. 1996. "Demographic Effects of Natural Disasters: A Case Study of Hurricane Andrew" *Demography* 33.2
Stallings, Robert A., Editor. 2002. *Methods of Disaster Research*. International Research Committee on Disasters.
Steiner, Andrea. 2005. "Loss of Heritage, Discovery of Justice: Elders and Premature Babies." Pg. 39-43 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.
Tierney, Kathleen. 2006. "Foreshadowing Katrina: Recent Sociological Contributions to Vulnerability Science" *Contemporary Sociology* 35.3
U.S. Department of Labor, Bureau of Labor Statistics. 2005. "Effects of Hurricane Katrina on BLS Employment and Unemployment Data Collection and Estimation" Methodology Paper. Washington, DC. Retrieved from Bureau of Labor Statistics website: https://stats.bls.gov/katrina/cpscesquestions.htm
Wellman, David. 2005. "Katrina: Giving Eyesight to the Blind." Pg. 82-86 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.
White, William Sakamoto. 2005. "Will the Circle be Unbroken?: Reflections on Place, Identity, and New Orleans Culture in the Aftermath of Hurricane Katrina." Pg. 23-29 in *Hurricane Katrina: Response and Responsibilities*, edited by John Brown Childs. Santa Cruz, CA: New Pacific Press.

Appendix B: Master Variable List

Variable Name	Description	Values	
LEVEL (Ordinal)	Black or AA Racial Identification Index	1 - White Alone 2 - Black/White Identification	3 - Black Alone Identification
PEEDUCA (Ordinal)	Highest Level of School Completed	-1 - Not in Universe 31 - Less than First Grade 32 - 1st, 2nd, 3rd, or 4th Grade 33 - 5th or 6th Grade 34 - 7th or 8th Grade 35 - 9th Grade 36 - 10th Grade 37 - 11th Grade 38 - 12 Grade, No Diploma	39 - High School Graduate - Diploma or Equivalent 40 - Some College but no Degree 41 - Associate Degree - Occupational/Vocational 42 - Associate Degree - Academic Program 43 - Bachelor's Degree (BA, AB, BS) 44 - Master's Degree (MA, MS, Meng, etc.) 45 - Professional School (MD, DDS, DVM, JD, etc.) 46 - Doctorate Degree (PhD, EbD, etc.)
MI (Ordinal)	Marriage Index	1 - Never Married 2 - Divorced, Separated, Widowed	3 - Married, Spouse Absent 4 - Married, Spouse Present
PEAGE (Cont.)	Age	Range 0:90	
PRNMCHLD (Cont.)	Number of Own Children < 18	-1 - Not in Universe (Not a parent)	Range 0:99
LFP (Ordinal)	Labor Force Participation Index	1 - Not in Labor Force, Child 2 - Not in Labor Force, Adult 3 - In Labor Force, Unemployed	4 - Employed, Absent from Work 5 - Employed
HUFAMINC (Ordinal)	Household - Total Family Income in Past 12 Months	-3 - Refused -2 - Don't Know -1 - Blank 1 - Less than $5,000 2 - $5,000 to $7,499 3 - $10,000 to $12,499 5 - $12,500 to $14,999 6 - $15,000 to $19,999 7 - $20,000 to $24,999	8 - $25,000 to $29,999 9 - $30,000 to $34,999 10 - $35,000 to $39,999 11 - $40,000 to $49,999 12 - $50,000 to $59,999 13 - $60,000 to $74,999 14 - $75,000 to $99,999 15 - $100,000 to $149,999 16 - $150,000 and over
HRNUMHOU (Cont.)	Number of Persons in Household	0:16	
GTCBSAST (Ordinal)	Principal City Balance	1 - Principal City 2 - Balance (Metro, non-Principal)	3 – Non-metropolitan 4 - Not Identified
HUP (Ordinal)	Housing Unit Permanency Index	0 - Other 1 - House, Apartment, Flat	2 - Mobile Home, Trailer 3 - Transient Hotel, Tent Site (Not in Universe)
SLCI (Ordinal)	Student Level and Commitment Index	0 - Not a Student 1 - High School, Part Time 2 - High School, Full Time	3 - College, Part Time 4 - College, Full Time
GTCBSASZ (Ordinal)	Metropolitan Area (CBSA) Size	0 - Not Identified or Non-metropolitan 2 - 100,000 to 249,999 3 - 250,000 to 499,999 4 - 500,000 to 999,999	5 - 1,000,000 to 2,499,999 6 - 2,500,000 to 4,999,999 7 - 5,000,000+