An analysis of commonality in the twin metropolitan areas of San Diego, California, and Tijuana, Mexico

Nanda K. Viswanathan\textsuperscript{a}, James B. Pick\textsuperscript{b}, W. James Hettrick\textsuperscript{c}, and Elliot Ellsworth\textsuperscript{d}

\textsuperscript{a}School of Management, Delaware State University, Dover, DE 19901, USA
\textsuperscript{b}School of Business, University of Redlands, Redlands, CA 92374, USA
\textsuperscript{c}City of Loma Linda, Loma Linda, CA 92354, USA
\textsuperscript{d}City of Ontario, Ontario, CA 91764, USA

Originally submitted October, 1999
Third revision submitted Feb. 25, 2003

Corresponding author: Dr. James B. Pick
Address: School of Business
University of Redlands
1200 East Colton Avenue
Redlands, California 92373-0999

e-mail: james_pick@redlands.edu
Fax: 909 335-5125
Phone: 909 748-6261
Abstract

A new application of spatial analysis is developed and applied to study the international twin metropolitan region of Tijuana, Mexico, and San Diego, California. Information on sixteen variables from the U.S. and Mexican censuses are merged in a common framework. A two-stage cluster analysis is applied to small area samples to determine metropolitan zones with common sets of characteristics. The results of the cluster analysis are spatially mapped and analyzed. This is done for all sixteen variables, as well as for four variable groups that are determined through literature review and experience, and for an alternative set of variable groups derived from the data. The extent of commonality, defined as the similarity between socioeconomic and demographic profiles of the population in the San Diego-Tijuana region, is examined through the development of a commonality index that is determined by the extent of cluster overlap on both sides of the border. The findings are interpreted with respect to the research literature. The methods can be applied within academic research or in the planning of border cities having consistent data sets on both sides.

Keywords: Spatial analysis, U.S.-Mexico border, cluster analysis, commonality

1. Introduction

This paper studies the urban structure of the U.S.-Mexico binational metropolitan regions of San Diego-Tijuana through spatial and cluster analysis. The objective is to analyze spatial patterns of important socioeconomic attributes within the twin metropolitan area. Overall patterns are examined, as well as those for more specific dimensions, such as population-migration, socioeconomic status, economic strength, and housing. The study
also seeks to measure the amount of commonality between the metropolitan areas on both sides of the border. The unit of analysis is a small area with an average population size of 1,699. It was selected to emphasize small, homogeneous intra-urban areas, consisting of several thousand persons. This allows for more refined comparisons of socioeconomic commonalities and differences than would be possible using larger heterogeneous units such as ZIP codes in the U.S. or municipios in Mexico.

An indicator of the importance of the border is the growing population of the twin metropolises. The population of the eight major U.S.-Mexican border twin cities in 1990 was 4.96 million, representing approximately 1.5% of the combined population of the two nations. This ratio of border population to total country population is much higher for Mexico than for the U.S. i.e. 3.5% versus 0.8%. The twin cities grew rapidly in the 20th century, in particular by 40% between 1980 and 1990 and by 34% between 1990 and 2000. We estimate the twin metropolises' population at 6.78 million people in year 2000, an increase in the 1990s of 37%. Medium-assumption estimates in the year 2020 for the U.S.-Mexico border county and municipio population is 19.46 million [18]. (A municipio in Mexico is equivalent to a U.S. county.)

2. Research Questions

The current research seeks answers to the following questions for the twin metropolises of San Diego and Tijuana.

1. What are the salient patterns of spatial distribution in the twin metropolitan region for 16 key socioeconomic attributes? Does this patterning conform to the concentric ring [4], Latin America city structure [9], or Mexican border city theory by Hoffman [12]?
2. What is the level of commonality of the twin-city for groups of variables classified \textit{a priori} based on theory or groups of variables derived from the data?

3. Do results of the data driven approach differ from the \textit{a priori} approach?

4. How do the commonality findings enhance understanding of the socioeconomic and demographic profiles of the twin cities of San Diego and Tijuana?

To answer these questions, a small area unit of analysis was used rather than the individual. The latter was not possible due to confidentiality rules of the U.S. and Mexican censuses [16]. More specifically, the unit of analysis chosen for this study was a group of blocks (U.S.) or, equivalently, "manzanas" for Mexico [13,14]. In Mexico, a group of blocks is referred to as "Area Geoestadística Basica Urbana" (AGEB) [14]. It is the smallest unit of analysis available to users. In the U.S., the equivalent unit is the block group [16]. The reason for choosing block groups (equivalently AGEBs) as the unit of analysis is that they are comparable in the U.S. and Mexico while other units, such as the tract or the block/manzana, are not available for both sides of the border. Finally, the block group and AGEB are commonly utilized in statistical analyses of small areas [16,17].

3. Literature Review

The U.S.-Mexico border cities have been of interest as research sites for generations of scholars. Most earlier studies have involved narrative [1,15], social history and urban description [10], historical-political-governmental qualitative analysis [29], merging of non-quantitative techniques from numerous disciplines [28], and quantitative analysis [8,20,21]. Such interest derives from these cities' importance as gateways and sources of interchange
between two of the world's largest nations, and from rapid population growth of said cities in the 20th century [2,15,29]. Growth of the maquiladora industry [6], and the North American Free Trade Agreement (NAFTA), which went into effect in 1994, have served to further stimulate the economic and labor force exchanges between the border cities and their hinterlands. It is also noteworthy that the 2000 mile U.S.-Mexico border is the world's longest between an advanced and a developing nation.

In spite of the growing importance of the border, there are relatively few studies that have analyzed the urban structure of the border using detailed quantitative data. Herzog's [10] study of San Diego-Tijuana adds some preliminary theory of urban structure for one of the border twin metropolises on both sides. Herzog thus analyzed the growth of San Diego from a historical perspective. In particular, he examined the impact of different historical events, and the impact of variables such as industrial policy and population growth that have led to the current spatial structures of both San Diego and Tijuana.

3.1 Theories of Urban Structure

A number of theoretical studies of the urban structure have been performed on the urban structure of cities [2,3,8,9,12,19]. They have focused on generalizations of urban structure for particular city typologies including U.S. and Latin American cities, Mexico City, and U.S.-Mexico border cities. Burgess [4], a founding theorist on delineating structure of U.S. cities, proposed a model with a central business district (CBD) and concentric rings around the outside (see Fig. 1). The rings include a low socioeconomic zone of transition near the core, industrial and middle class residential zones, and "better off" suburban zones at the periphery. This model originated from observations and data on Chicago and other
American cities of the early twentieth century. Those cities had structure traceable to the evolution of transport systems as well as outward economic expansion away from the centralized business districts.

The model of the Latin American city largely turns inside-out the traditional Burgess model. In the U.S. city model, poorer zones are more central; for Latin America, on the other hand, they are on the periphery. Indeed, in Latin American cities, there are elite residential zones located near the center, alongside a grand transportation artery [9].

It is important to consider whether the presence of an international border alters the theoretical frameworks of cities on both sides. Key questions thus arise: are there binational groupings which supersede the theoretical arrangements on one side or the other? Is the borderline an important barrier or is it permeable and pliable, serving only as a theoretical marker?

3.2 Influences of the U.S.-Mexican Border on Theories of Urban Structure

While there are theories of urban structure on the Mexican side that are modified to take into account the border, none could be found for the U.S. side. Gildersleeve [8] offers a model of Mexican border city structure. The central business district is right up against the border in the center, and there are surrounding residential zones, which, however, are not concentric. Generally, the better off zones are more central while the poorer zones more peripheral. Scattered in this array are small business and industrial zones. A more elaborate
model of Hoffman [12] again shows the central business district up against the border in the center. There are now concentric half-circles surrounding the central business district, with the inner one of medium quality and deteriorated quality on the periphery. There are also several major transport corridors extending from the center and beyond i.e., internally into Mexico. Towards the periphery, high quality residential areas surround these corridors. There are scattered industrial areas near the center. The complex Hoffman model appears to apply more to larger cities.

--- Fig. 2 about here ---

Arreola and Curtis [2] present an even more elaborate model (see Fig. 2). It builds on the Hoffman model, since it also has a central business district next to the border, middle income areas in an inner semi-circle, and middle/lower income in an outer semi-circle. However, it also has elite areas located centrally near the CBD, a wedge-like inner city slum against the border on one side, scattered maquiladora zones, and several slum areas at the far periphery. In many respects, this conforms more to the traditional Griffin-Ford Latin American city model [9], although modified into a semi-circle arrangement and altered by the presence of the maquiladora districts.

Although research literature on the border [e.g., 2,5,8-12,15,23,28,29] is informative about the history, theory, and qualitative social and economic structure of the border twin cities, there is a relatively little detailed quantitative analysis of urban structure on the Mexican side. One exception is an atlas put together by San Diego Dialogue that shows some matched characteristics for small areas of San Diego and Tijuana [25]. The present authors
have studied the spatial structure of the counties and municipios in the Mexican border states [23].

Although maps are presented in the atlas of [25] for several characteristics, there are no multivariate methods or summary measures that analyze border and binational trends and patterns. For the larger U.S. border cities of San Diego and El Paso, there are also reports from various planning departments that show spatial patterns, but do not include the other side of the border [24]. The present study, on the other hand, provides multivariate data analysis, performs cluster analysis, derives data driven cluster groupings, and presents a new index of commonality, all of which yield deeper and more quantitative findings on the binational urban structure for the border cities of San Diego and Tijuana.

4. Proposed Methodology

Our proposed methodology consists of three steps. The first step involved matching 16 variables in definition for the two national censuses. Second, a cluster analysis was conducted for each of the matched variables, and for each of the variable groups, in order to cluster together areas with common characteristics. Third, a commonality index based on cluster overlap between the twin cities was developed to measure the extent of commonality between San Diego and Tijuana.

4.1. Matching of U.S. and Mexican Census Data

The study data were drawn from the 1990 U.S. Census of Population [30] and the 1990 Mexican Census of Population [13]. The data were for small areas with an average size from the two censuses of 1,699 persons. On the U.S. side, the unit of analysis was the block
group, while on the Mexican side the AGEB (Area Geográfica Estadística Básica) was used. There was a total of 1,381 small area units in San Diego and Tijuana that comprised a total population of 2,346,138 persons. The U.S. Census has hundreds of attributes available at the small area level, about the same number as for larger units of analysis such as the city, county, and state. However, for the Mexican Census, the number of attributes is reduced at the small area versus larger area units of analysis.

Although the rules for determining boundaries for small areas are different between the U.S. and Mexican censuses, this was not considered a problem for the current research since the focus is on characterizing patterns for the entire twin city area. On the U.S. side, a small proportion of block groups on the urban fringe are larger in size. However, the U.S. classifies them as part of the urban area.

The 16 variables that were matched from the U.S. and Mexican censuses are shown in Table 1. A small proportion of all variables available in the two censuses could be matched -- about 20% of the variables on the Mexican side, and less than 5% of those on the U.S. side. The table shows the comparable definitions from the U.S. and Mexican censuses. The variables were classified into four groups. The population variable group consists of population, dependency ratio, gender ratio, nativity, and immigration; the social variable group includes education, marriage, and home ownership; the economic variable group consists of economic activity, underemployment, unemployment, and poverty; and the housing group consists of one bedroom housing, separate kitchen, and public sewer i.e. drainage connected to the street.

--- Table 1 about here ---
4.2. Multivariate Spatial and Cluster Analyses of the Border Twin Cities

It is important not only to examine the spatial distributions of individual characteristics for the San Diego-Tijuana twin city, but also to analyze the patterns for groups of characteristics consolidated together, for instance, economic characteristics. This can be done by utilizing the multivariate technique of cluster analysis, which agglomerates together cases based on how similar they are for a group of attributes.

The two major methods of cluster analysis are hierarchical and nonhierarchical [1,7,26]. In the former, the two most similar cases are agglomerated into a single case. Then, the mean values for the agglomerated cases are added to the full set of cases and the two original cases eliminated. Further agglomeration takes place, continuing by order of the next most similar cases. This continues until some stopping point is reached.

Nonhierarchical clustering procedures such as K-means clustering do not use a treelike, iterative construction process. Instead, they assign all cases to clusters after specification of the sought-after number of clusters. The assignment is based on the distance between the new case and the cluster center, i.e., the object is assigned to the cluster whose center is the shortest distance from the object. There are many alternative methods to judge "similarity." These include least squares and squared Euclidean distance, which can be augmented with agglomeration rules such as K-Means and Ward's Method [7,27]. The present study measures distance as squared Euclidean; and uses hierarchical clustering and the K-Means agglomeration procedure [27]. The K-Means is chosen for this study because of its capability to handle large groups of cases, which apply in the current study [27]. K-Means most
commonly utilizes Euclidian distance and solely for the statistical software package utilized in this research [27].

Cluster analysis was applied to each of the border twin city samples using a two-stage approach. The first stage utilized hierarchical clustering to determine the number of clusters based on patterns of change in the agglomeration coefficient. The large number of cases involved in the analysis, however, made it difficult to interpret clusters based solely on hierarchical clustering. We thus added a second stage that used K-means since it can better handle large data sets and is less susceptible to outliers in the data and to the inclusion of irrelevant and inappropriate variables.

Prior to clustering, all variables were standardized with Z scores and missing values were imputed, based on a linear fit from adjacent areal units. Upon completion of the K-means procedure, each cluster was profiled using the cluster means for its component variables.

Overall cluster analyses, based on all sixteen matched variables, as well as cluster analyses for groups of variables, were performed. The variable groups were identified in two different ways in an effort to provide more information and enhance our understanding of the pattern of commonality in the twin cities. The first method involved an a priori classification of variables into four different groups based on past geographical clustering research in Mexico City [19]. These groups are: population, social, housing, and economic (see Table 1). A data driven cluster analysis was then separately performed for each of the four categories, based on the 1,381 cases. The second method of cluster analysis we term data driven. It was based, first, on a cluster analysis of the variables, as opposed to cases. That is, it seeks to group together similar variables. For San Diego-Tijuana, the data driven cluster analysis
resulted in four different groupings of variables. A second stage, cluster analysis of 1,381 cases, was then performed separately for each of the four groups of variables.

While there were similarities in the composition of the four "data driven" groups, as compared to the four groups derived from the a priori analysis, there were also differences, as discussed below.

4.3. Commonality Index

The extent of commonality was analyzed using an index of cluster overlap. Since clustering here is based on the similarity of attributes within the geographical areas identified as part of the cluster, the extent of overlap would be an indicator of similarity between the U.S. and Mexico. The aggregation of overlap across all clusters that form part of a twin city thus would be an indicator of the similarity between the U.S. and Mexican parts of that twin city. This can be measured by a commonality index, defined as an index that measures whether clusters tend to be similar by being balanced on both sides of the border or different by falling mostly on one side or the other side of the border. We propose the following commonality index:

\[
1 - \frac{\sum |US(i) - Mexico(i)|}{\sum |US(i) + Mexico(i)|}
\]

(1)

where: US(i) represents the number of areal units for cluster (i) located in the U.S., Mexico(i) represents the number of areal units for cluster (i) located in Mexico, and the summation is from 1 to n, where n represents the number of clusters of at least two areal units.

The approach to calculating differences between two populations is similar to the Index of Dissimilarity [26]. That index is computed as follows:

\[
= 0.5 \sum \left| \frac{P(i)}{\sum P(i)} - \frac{R(i)}{\sum R(i)} \right|
\]

(2)
where: \( P(i) \) represents the population of the \( i \)th member of group \( P \) and \( R(i) \) represents the population of the \( i \)th member of group \( R \).

However, the differences are normalized differently in the two indices. In the Index of Dissimilarity, they are normalized as percentages of the total populations, whereas in the Commonality Index, they are normalized by a denominator equal to the sum of the two groups. Our proposed commonality index emphasizes the differences for the larger groups more than the index of dissimilarity. This seems appropriate for planning purposes, because the larger cluster groups represent more population.

The proposed index \((1)\) can be modified by adding weights that balance the importance of cluster groups differently, e.g. weights could be incorporated that value the differences equally. In the present paper, however, weighting is not considered as we use \((1)\) in its base form.

Our Commonality Index is illustrated with the following example. Assume the cluster analysis of a group of variables for a twin city on the U.S. - Mexico border results in two clusters with a total of 240 areal units distributed as shown in Table 2.

<table>
<thead>
<tr>
<th>Cluster group</th>
<th>Number of small area units of cluster group in U.S. (A)</th>
<th>Number of small area units of cluster group in Mexico (B)</th>
<th>Difference between number of small area units (A) - (B)</th>
<th>Sum of small area units of cluster group in U.S. and Mexico (A) - (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>40</td>
<td>60</td>
<td>140</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>10</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>190</td>
<td>50</td>
<td>140</td>
<td>240</td>
</tr>
</tbody>
</table>
The resulting commonality index is: $1 - \frac{140}{240} = 0.417$.

Thus, for a cluster group, a more even division of small area units between the U.S. and Mexico increases the commonality index, while an uneven division of small area units between the two counties reduces the commonality index.

5. Results

5.1. Descriptive Results

This section summarizes selected descriptive results for the 16 matched variables. Detailed results for all 16 variables are reported elsewhere [22].

a) Population. In San Diego, there is a mosaic-like pattern of population. Some peripheral areas have large populations due to large block urban groups. The downtown, however, has a very different pattern with small areas and moderate population. Concentric rings and wedge-shaped patterns which are typical of many U.S. cities are thus not apparent here.

For Tijuana, the population per small area has a much higher concentration in the center than in the periphery. This is typical of Latin American cities [12]. This pattern was also observed by delegation and municipio for Mexico City in 1990 [19]. For Tijuana, some of the areas of largest population are adjacent, or near, to the border.

b) Dependency ratio. Dependency ratio is defined as the ratio of population less that 18 years plus population 65 years and older to the population aged 18 to 65. The spatial distribution for dependency ratio reveals, in Tijuana, a classical concentric ring pattern [2,4]. With values in the 0.5-0.6 range, the Tijuana core resembles many areas in San Diego. Ratios on the periphery, however, often exceeded 1.0, meaning that there are more dependent persons, i.e. children and older people than workers. In San Diego, there is substantial variation in the
dependency ratio. There is, for example, a much higher dependency in an area to the south of the central business district and in the area immediately north of the international border crossing gate. Disregarding this latter area, the spatial pattern for the rest of San Diego resembles concentric rings, with higher dependency ratios in the central and middle rings, and lower on the periphery.

c) Nativity represents the proportion of the population that was born in the state. For Tijuana, nativity is lowest on certain areas of the periphery, where it tends to have values under 1/3. These are fast growing migratory areas. On the other hand, there is a mixed pattern in the center, with nativity varying between one third and 56%. Overall, nativity is low compared to other cities in Mexico. The reason stems from the rapid population growth of Tijuana in the twentieth century. It grew more than 11 fold from 1950 to 1990, and continued to expand rapidly in the 1990s. This has led to the presence of substantial numbers of inmigrants in every part of the city.

For San Diego, the nativity varies from lower levels in the west to higher levels in the east. The low nativity for most of the city’s western sections reflects two aspects: the population growth over the past 50 years in the northwestern areas and the military presence in the western and southwestern parts of the city. Also, there are areas of high Hispanic population in the southwest, which implies high levels of migration from Mexico and thus low nativity. A more traditional zone, which has the city’s highest nativity, is found in the northeast.

5.3. Overall Cluster Analysis Results
The cluster analysis results are presented first for the overall 16-variable coverage of San Diego-Tijuana, followed by the four a priori variable groups. We interpreted our findings terms of the methodology and the theories of urban structure discussed earlier in the paper.

The overall Commonality Index generated an average value of 0.015 based on the pattern of clusters. The pattern reveals little in common between San Diego and Tijuana in terms of the combined effect of all 16 variables. Table 3 summarizes the overall cluster pattern by cluster group for the twin cities. For each cluster, the average values are given for the 16 variables.

--- Table 3 about here ---

The table characterizes each cluster. The most important findings are the following. For Tijuana, there is a central core area (cluster 9) surrounded by a periphery (cluster 7). This core has a mean economic level that is relatively prosperous for Mexico. It has a significantly lower dependency ratio than the periphery (0.80 compared to 1.00) and gender ratio (about six% lower) than does the periphery. In addition, the periphery (cluster 7) has notably low levels of housing quality, i.e. the proportion of one room housing is 18%, compared to 6% in the Tijuana core. Even more pronounced is the issue of sewer connections, with 16% in the periphery versus 75% in the core. The dramatic difference in housing quality between the Tijuana core and periphery is consistent with patterns in other Mexican border cities [2] and reflects their rapid growth and concomitant lack of infrastructure.

The overall cluster analysis suggests a markedly different urban structure for San Diego versus Tijuana. San Diego is characterized by a southern spine cluster and a downtown
5.3. Cluster analysis results for a priori groups of variables

While the overall cluster analysis revealed little in common between San Diego and Tijuana, cluster analyses of the a priori variable sets suggest the existence of similarities in
some areas and differences in others. The Commonality Index, for example, is higher for the population and economic groups (0.102 and 0.126), but lower for the social and housing groups (0.062 and 0.005).

An example of the clustering by variable sets, using the population group, is shown in Figs. 3 and 4. Figure 3 presents a map of clusters 1-7 of the population, and Figure 4 shows clusters 8-14 of that group. Two figures are utilized, rather than one, in order to be possible to better distinguish the clusters using black and white mapping. Figure 3 reveals that population clusters 1-7 are predominantly on eastern part of the San Diego side. Population clusters 8-14 in Figure 4 are on the central and eastern sides of San Diego and in Tijuana. It is evident that the most important binational clusters are 9, 10, 11, and 13.

Figures 3 and 4 about here

Table 4 characterizes the population group clusters. Several important features are as follows. This group reflects a large mixing of clusters, especially near the border, with small pieces of these clusters located throughout the San Diego metropolitan area. The three population clusters containing this mixing pattern are cluster 10, 11, 13, and 14. Clusters 11 and 14 are especially populous, with 281,564 and 155,917 persons, respectively, and differ particularly in Cluster 14 having double the Chapter 12’s migration rate. Cluster 13, which is on the periphery, is somewhat similar to cluster 14, but has an approximately 17% lower gender ratio.
The social and economic cluster analyses have been analyzed by the same techniques and provide useful insights into the urban structure of the twin city. Due to space limitations, the details of these cluster analyses are not covered in this paper.

The cluster analysis for the housing group resembles the 16-variable, overall results in terms of urban structure. The spatial patterns are shown in Figures 5 and 6 and the housing clusters are characterized in Table 5. Figure 5 reveals that housing clusters 1-7 are predominantly on the western spine of the San Diego and next to the border, as well as in the Tijuana periphery. Housing clusters 8-14 in Figure 6 are in northern and western San Diego and throughout Tijuana. However, the maps also show that there are no binational housing clusters except trivially. In other words, housing is a characteristic separated on one side or the other of the border.

DESCRIPT THE MAPS

In Tijuana, there is a distinctive central core, cluster 11, which is surrounded by a larger set of peripheral zones. The Tijuana housing core has a low proportion (4%) of one room housing, a low proportion (8%) of non-detached kitchens, and a high proportion (83%) of connected sewers. By Mexican standards, this housing core area is rather advanced. The
Tijuana periphery has become more differentiated. Thus, clusters 9 and 14 have more advanced housing whereas clusters 3 and 4 have much less favorable housing conditions. More specifically, approximately 1/3 of the Tijuana periphery has better housing, while two thirds is more deteriorated. Interestingly, the zones of advanced and deteriorated housing are interspersed, a pattern that was also noted in Mexico City.

On the San Diego side, there are several large housing clusters, in particular Clusters 6, 7, and 13, with 247, 340, and 528 areal units respectively. Clusters 6 and 7 demonstrate the very high housing standards, with nearly 100% separate kitchens and connected sewers, while cluster 6 has twice the one room housing of cluster 7. Cluster 13 appears to be a "suburban ring," similar to the suburban ring for the overall, 16-variable distribution. It shows a very high level of housing quality and conforms to classical urban theory of concentric rings, stemming from the Chicago school [4]

In summary, the levels of commonality based on cluster analyses of four a priori variable sets suggest greater similarity between San Diego and Tijuana than was identified in the overall cluster analysis. The similarities primarily pertain to nine population and economic variables: population, dependency, gender ratio, nativity, immigration, economic activity, unemployment, underemployment, and poverty.

The cluster analysis of the a priori variable sets took us beyond the overall, 16-variable analysis in understanding the similarities and differences between San Diego and Tijuana. However, our a priori sets may not be the only, or best, way to group the variables. They
were created by the researchers, based on prior research carried out in Mexico City [19], and on their experience and logic. Because Mexico City is quite different in size and structure from the twin border city of San Diego-Tijuana, the use of such prior research may have problems. Consequently, a second method of grouping the variables, not based on Mexico City or the researchers' experience, was also used.

5.4. Cluster analysis results for "data driven" groups of variables

The second cluster grouping method, termed "data driven," involves first a cluster analysis of the variables, and second, a cluster analysis of the cases. The former resulted in the creation of four "data driven" variable sets.

The data driven and a priori variable sets are rather similar, and are described in Table 6. While the commonality index for Group 1 is low (0.016), the indices for Groups 2, 3, and 4 are moderate to high (0.300, 0.310, and 0.137 respectively). Groups 2 and 3 have magnitudes of commonality much higher than those generated from the a priori approach. Findings from the two approaches are discussed in the following paragraphs.

--- Table 6 about here ---

Group 1 is close in variable composition to the Population Group selected a priori. The only difference the two is that Group 1 replaces nativity with primary education. The commonality indices are also very similar. Hence, the Group 1 pattern will not be explained further.
Likewise, Group 4 is very close in variable composition to the Housing Group selected a priori. The only difference is that the former replaces one bedroom housing with economic activity. However, because of the strength of the common variables i.e., separate kitchen and sewer, the two cluster patterns are highly similar. The commonality indices are very low for both groups. Group 4 is sufficiently close to the Housing Group that it is not discussed further.

The other data driven groups reveal different patterns. Group 2 consists of the variables: married, native to state, and home ownership. Taken together, this dimension reflects, for the U.S., more stable, "suburban like" patterns. The results show considerable commonality in cluster patterns between the two sides, with sharing in the Mexican and U.S. peripheries, but a separated pattern in the U.S. core. The index of commonality is high at 0.309.

Another data driven group with substantial commonality between San Diego and Tijuana is Group 3, consisting of secondary education, unemployment, underemployment, poverty, and one room housing (see Figs. 7 and 8). This dimension has strong economic elements. There is high commonality at 0.310. Figure 7 reveals especially strong binational clusters 5 and 7. In cluster 5, the urban core for Tijuana extends to common areas in the San Diego periphery in the northwest, north, east, and southeast. Some areas in the Tijuana periphery are similar to a scattering of small areas in the north central and northwest parts of San Diego. These two clusters account for most of the high commonality index. Clusters 8-14 in Figure 8 are located only on the U.S. side.
The data driven cluster analysis has further enhanced our understanding of the patterns of commonality. It can combine new and unanticipated sets of factors, such as marriage and nativity, that amplify commonality to high levels. In particular, Groups 2 and 3 reveal patterns that are more binational than any of the a priori groups. Relatively high similarities in marriage and nativity patterns between San Diego and Tijuana may be a function of the 1980s' growth of both cities with a large number of non-native populations in both places.

Results demonstrate that the replacement of economic activity and home ownership in the a priori economic group with secondary education and one room housing in data driven group 3 increased the commonality index from 0.126 to 0.310. This may point to strong contributions from secondary education and one room housing to levels of commonality.

Based on a comparison of the factors used in the a priori and data driven approaches, a new and unanticipated grouping occurred in the latter approach for four variables: namely, marriage, nativity, secondary education, and one room housing. This elevated the values for commonality indices for the data driven groups. A visual examination of the spatial distribution maps of the twin cities for these four variables supports the relatively higher levels of similarity between San Diego and Tijuana with regard to these four variables. To show this, an example of a spatial distribution map for secondary education is given in Fig. 9.

The maps shows that higher levels of secondary education are scattered the center, center south, and northeast of San Diego. Although Tijuana’s secondary education level is lower than for San Diego, it likewise has a scattered distribution throughout the city.
6. Discussion and Conclusions

The research questions outlined at the beginning of this paper address two broad issues. First, questions 1 and 2 are concerned with the nature of the spatial patterns of the twin metropolis of San Diego and Tijuana and how closely they correspond to the conceptual theories of urban structure. Second, questions 3 and 4 address the issue of commonality, and the use of two different methodological approaches to answer it. This section discusses the major results based on the four research questions.

This paper examined the patterns of spatial distribution and commonality for 16 key population, social, economic, and housing variables for the twin city of San Diego, California and Tijuana, Mexico. Data were drawn from the U.S. and Mexican censuses of 1990. The unit of analysis was the small area averaging 1,699 population that is designated a block group in the U.S. and AGEB in Mexico.

The variable definitions for the two nations were matched as closely as possible. The geographic coverages were aligned on both sides of the border. Geographic patterns were analyzed by cluster analysis, based on a priori and data driven groupings of variables. In order to further test for the extent of commonality, an index was constructed to measure the overlap of cluster groups across the border.

6.1. Spatial Distributions
In relation to question 1, analysis of the spatial patterns of individual socioeconomic attributes indicated three variables to be of particular interest: population, dependency, and nativity. The analysis of population patterns suggested that, while the population pattern for Tijuana is similar to that of other Latin American cities in having higher concentrations of population in the core than in the periphery, the population pattern for San Diego is mosaic-like and differs from many other large U.S. cities that have ring and wedge patterns. The findings for dependency ratios indicated that areas on the southern end of San Diego near the border are very similar to the periphery of Tijuana. The spatial patterns of nativity demonstrate high levels of commonality between San Diego and Tijuana. This reflects the fast growing nature of both cities and the high level of population influx from outside.

The second research question concerned the urban spatial patterns of San Diego and Tijuana, based on the overall cluster analysis of the 16 socioeconomic attributes. Findings on the spatial distributions of San Diego-Tijuana offer many new insights. The findings strongly support border urban theory for most variable groupings in Tijuana. Patterns in Tijuana thus appear to resemble the urban structure as outlined in the Hoffman [12] model for Mexican border cities. The spatial maps for most variables reveal a central business district right up against the border in the center, surrounded by concentric half circles. The traditional urban theory of U.S. cities describes a central business district surrounded by concentric rings provides relatively little correspondence to patterns in San Diego. The reason may be partly due to the mixing and change factors caused by nearness to the border, especially in the southern half of San Diego. While it is difficult to generalize the applicability of a theory of urban structure based on the study of a single twin city, a larger study that examines patterns of urban structure along the entire 2000 mile border between the U.S. and Mexico would be
useful. The paper’s researchers are conducting such a study of 16 twin cities along the border to enhance understanding of urban structure patterns along the entire border.

6.2. Commonality

Research question 3 assesses the levels of commonality between the twin cities of San Diego and Tijuana. The findings here reinforce our analysis of individual socioeconomic attributes, as well as add new information. For the a priori groups, those factors with the most binational sharing of characteristics were economic and population.

The results for the commonality index derived from the data driven groups reinforced results obtained from the a priori groups. For the former, the commonality index for group 2 (married, native-to-state, home ownership) and group 3 (secondary education, unemployment, underemployment, poverty, and one room housing) was high. A comparison of the variables here revealed five factors: nativity, home ownership, unemployment, underemployment, and poverty. The robustness of these variables is likely due to the rapid growth of both San Diego and Tijuana. Further, both cities are relatively prosperous compared to other cities in their respective countries. Thus, both cities also have relatively low levels of unemployment, underemployment, and poverty.

The fourth research question asks how the findings enhance demographic and economic understanding of the twin cities. This was accomplished in three ways. First, spatial analysis of the merged census data enabled us to treat both cities as one region. Second, the development of a commonality index helped quantify levels of commonality that can be spatially observed. This quantification allows for comparison of commonality across other sets of twin border cities as well as for longitudinal comparisons into the future. Finally, the
pooling of variables helped identify broad themes of commonality versus those based on individual variables.

Although this study presented a detailed quantitative analysis of the twin cities of San Diego and Tijuana, there were limitations as a result of the relatively few variables that could be definitionally matched from the two census. Further, some attributes that are desirable for studies of commonality were not collected by the censuses, e.g., percent English speaking on the Mexican side (to demonstrate common language capability). Attitudinal attributes, such as self-perception of culture and ethnicity, would also be desirable.

Nevertheless, the analyses presented here offer significant insights for working urban planners on either side of the U.S.-Mexico border. Planners can thus see complex patterns of cross-border similarities and differences needed to better understand educational distributions, housing patterns, migration growth areas, and economic shifts. This, in turn, has implications for better planning in terms of infrastructure needs, environmental mitigation, city services, population growth, and cultural programs. Hopefully, the U.S.-Mexico twin cities will develop and implement shared information systems, utilizing the methods in the paper to plan in a more integrated manner. The proposed methods and techniques, including the use of GIS, can also be adopted by other international twin cities that are willing to take a cooperative planning approach.

**Acknowledgements**

The authors acknowledge the Ford Foundation for support of this research and the research assistance of Swarna Keshavan.
References


Table 1. Variable Groups, matched variables, and their definitions from Mexican and U.S. censuses

<table>
<thead>
<tr>
<th>Twin Cities Group and Variable</th>
<th>Comparable Definition - Mexican Census</th>
<th>Comparable Definition - U.S. Census</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population Variable Group</strong></td>
<td>total population</td>
<td>total population</td>
</tr>
<tr>
<td>Population</td>
<td>(pop less than age 18 + pop age 65 plus)/pop between ages 18 and 65</td>
<td>(pop less than age 18 + pop age 65 plus)/pop between ages 18 and 65</td>
</tr>
<tr>
<td>Dependency Ratio</td>
<td>male population/female population</td>
<td>male population/female population</td>
</tr>
<tr>
<td>Gender Ratio</td>
<td>population born in the state/total population</td>
<td>population born in the state/total population</td>
</tr>
<tr>
<td>Nativity</td>
<td>population 5 plus resident outside the state in 1985/population 5 plus</td>
<td>population age 5 plus that resided outside the state in 1985/population age 5 plus</td>
</tr>
<tr>
<td>Immigration</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social Variable Group</strong></td>
<td>primary education persons age 15 plus/population age 15 plus</td>
<td>persons 15 plus with less than 9th grade education/population 15 plus</td>
</tr>
<tr>
<td>Primary Education</td>
<td>secondary education persons age 15 plus/population age 15 plus</td>
<td>persons who have attended high school 15 plus/population 15 plus</td>
</tr>
<tr>
<td>Secondary Education</td>
<td>married persons age 12 plus/population age 12 plus</td>
<td>presently married population/population 15 plus</td>
</tr>
<tr>
<td>Married</td>
<td>owned homes/occupied housing units</td>
<td>ratio of owner occupied housing units to housing units</td>
</tr>
<tr>
<td>Home Ownership</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Economic Variable Group</strong></td>
<td>economically active pop/population</td>
<td>persons in the labor force/total population</td>
</tr>
<tr>
<td>Economic Activity</td>
<td>unemployed/economically active pop</td>
<td>ratio of unemployed persons to total labor force</td>
</tr>
<tr>
<td>Unemployment</td>
<td>underemployed (one to 32 hrs/week)/economically active pop</td>
<td>persons working one to 35 hours per week/total labor force</td>
</tr>
<tr>
<td>Underemployment</td>
<td>workers earning less than minimum wage/economically active pop</td>
<td>ratio of persons age 18 to 64 below the poverty line, by official U.S. Census definition*</td>
</tr>
<tr>
<td>Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Housing Variable Group</strong></td>
<td>single bedroom housing unit/occupied housing units</td>
<td>one-bedroom housing units/all housing units</td>
</tr>
<tr>
<td>One Bedroom Housing</td>
<td>separate kitchen/occupied housing units</td>
<td>housing units with complete kitchen facilities/all housing units</td>
</tr>
<tr>
<td>Kitchen</td>
<td>drainage connected to street/occupied housing units</td>
<td>housing units with connection to public sewer/all housing units</td>
</tr>
<tr>
<td>Public Sewer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The official U.S. census defines poverty by utilizing a group of monetary income thresholds which are varied by family size and composition. If a family is categorized as poor, then every person in the family is categorized as poor. The income thresholds are updated annually by the consumer price index and they do not vary geographically. The income threshold consists of before tax income that excludes capital gains and noncash benefits, such as medicaid and food stamps.

Table 3.
Average Values of 16 Matched Variables for 14 Overall Cluster Groups for San Diego-Tijuana Twin City, 1990

<table>
<thead>
<tr>
<th>Cluster Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Area Units in San Diego</td>
<td>1</td>
<td>16</td>
<td>12</td>
<td>3</td>
<td>22</td>
<td>265</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>567</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>272</td>
</tr>
<tr>
<td>Small Area Units in Tijuana</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>97</td>
<td>0</td>
<td>112</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Small Area Units - Total</td>
<td>1</td>
<td>16</td>
<td>12</td>
<td>3</td>
<td>22</td>
<td>265</td>
<td>97</td>
<td>1</td>
<td>117</td>
<td>572</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>272</td>
</tr>
<tr>
<td>Population of Cluster</td>
<td>21</td>
<td>34,556</td>
<td>7,330</td>
<td>24</td>
<td>53,746</td>
<td>299,885</td>
<td>177,983</td>
<td>6,141</td>
<td>592,651</td>
<td>750,441</td>
<td>9</td>
<td>13</td>
<td>32</td>
<td>423,325</td>
</tr>
</tbody>
</table>

**Population Variable Group**
- Population: 21,2160, 611, 8, 2443, 1132, 1835, 6141, 5065, 1312, 9, 13, 32, 1556
- Dependency Ratio: 0.000, 0.204, 0.305, 0.997, 0.091, 0.340, 1.000, 0.002, 0.797, 0.561, 0.000, 0.854, 0.000, 0.778
- Gender Ratio: 0.931, 3.704, 3.039, 3.018, 9.818, 1.124, 1.060, 617.400, 0.995, 0.989, 1.121, 1.172, 0.927, 0.974
- Nativity: 0.000, 0.488, 0.214, 0.751, 0.154, 0.337, 0.372, 0.093, 0.431, 0.421, 0.000, 0.000, 0.000, 0.361
- Immigration: 1.000, 0.119, 0.221, 0.000, 0.655, 0.186, 0.304, 0.744, 0.201, 0.077, 1.000, 0.000, 0.000, 0.093

**Social Variable Group**
- Primary Education: 0.000, 0.026, 0.110, 0.285, 0.002, 0.036, 0.249, 0.000, 0.188, 0.039, 0.000, 0.997, 0.000, 0.167
- Secondary Education: 0.000, 0.128, 0.357, 0.000, 0.160, 0.219, 0.166, 0.101, 0.182, 0.284, 1.000, 0.000, 1.000, 0.332
- Married: 0.000, 0.280, 0.265, 0.285, 0.246, 0.373, 0.588, 0.204, 0.527, 0.620, 0.000, 0.000, 0.000, 0.497
- Home Ownership: 0.207, 0.415, 0.042, 0.000, 0.257, 0.236, 0.744, 0.223, 0.613, 0.720, 0.000, 1.000, 0.000, 0.289

**Economic Variable Group**
- Economic Activity: 1.000, 0.470, 0.550, 0.375, 0.883, 0.654, 0.322, 1.000, 0.370, 0.548, 1.000, 0.000, 0.000, 0.415
- Unemployment: 0.000, 0.065, 0.172, 1.000, 0.002, 0.051, 0.031, 0.000, 0.023, 0.048, 1.000, 0.065, 0.124, 0.107
- Underemployment: 1.000, 0.819, 0.210, 1.673, 0.111, 0.217, 0.172, 0.082, 0.186, 0.250, 0.000, 0.384, 0.232, 0.235
- Poverty: 1.000, 0.597, 0.305, 1.000, 0.861, 0.122, 0.091, 1.000, 0.066, 0.056, 0.000, 0.000, 1.004, 0.234

**Housing Variable Group**
- One Bedroom Housing: 0.302, 0.247, 0.297, 0.000, 0.112, 0.382, 0.182, 0.318, 0.057, 0.074, 0.504, 0.000, 0.011, 0.321
- Separate Kitchen: 0.937, 0.848, 0.546, 1.000, 0.997, 0.994, 0.631, 1.000, 0.790, 0.996, 1.000, 1.000, 1.000, 0.988
- Public Sewer: 1.006, 0.990, 1.000, 1.000, 1.000, 1.000, 1.000, 1.000, 0.754, 0.968, 1.000, 1.000, 1.000, 0.989
Table 4.
Average Values of 5 Matched Variables for Population Cluster Groups for San Diego-Tijuana Twin City, 1990

<table>
<thead>
<tr>
<th>Cluster Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Area Units in San Diego</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>233</td>
<td>19</td>
<td>32</td>
<td>1</td>
<td>31</td>
<td>13</td>
<td>234</td>
<td>235</td>
<td>68</td>
<td>15</td>
</tr>
<tr>
<td>Small Area Units in Tijuana</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>62</td>
<td>16</td>
<td>0</td>
<td>80</td>
<td>39</td>
</tr>
<tr>
<td>Small Area Units - Total</td>
<td>1</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td>234</td>
<td>20</td>
<td>33</td>
<td>1</td>
<td>36</td>
<td>75</td>
<td>250</td>
<td>235</td>
<td>148</td>
<td>54</td>
</tr>
<tr>
<td>Population of Cluster</td>
<td>5,958</td>
<td>110,174</td>
<td>14,240</td>
<td>30</td>
<td>255,790</td>
<td>30,984</td>
<td>24,836</td>
<td>6,141</td>
<td>31,748</td>
<td>86,839</td>
<td>281,564</td>
<td>235,254</td>
<td>37,442</td>
<td>155,917</td>
</tr>
</tbody>
</table>

Population Variable Group

<table>
<thead>
<tr>
<th>Population</th>
<th>5,958</th>
<th>12,242</th>
<th>7,120</th>
<th>15</th>
<th>1,103</th>
<th>1,721</th>
<th>801</th>
<th>8,141</th>
<th>1,221</th>
<th>1,772</th>
<th>1,292</th>
<th>1,001</th>
<th>3,120</th>
<th>6,497</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependency Ratio</td>
<td>0.001</td>
<td>0.913</td>
<td>0.010</td>
<td>0.000</td>
<td>0.345</td>
<td>0.147</td>
<td>1.130</td>
<td>0.002</td>
<td>0.696</td>
<td>1.011</td>
<td>0.661</td>
<td>0.347</td>
<td>0.773</td>
<td>0.719</td>
</tr>
<tr>
<td>Gender Ratio</td>
<td>87.217</td>
<td>1.009</td>
<td>15.535</td>
<td>1.071</td>
<td>1.238</td>
<td>4.925</td>
<td>1.013</td>
<td>617.356</td>
<td>0.963</td>
<td>1.052</td>
<td>1.008</td>
<td>1.190</td>
<td>0.974</td>
<td>1.145</td>
</tr>
<tr>
<td>Nativity</td>
<td>0.114</td>
<td>0.441</td>
<td>0.100</td>
<td>0.000</td>
<td>0.287</td>
<td>0.170</td>
<td>0.190</td>
<td>0.093</td>
<td>0.353</td>
<td>0.326</td>
<td>0.515</td>
<td>0.420</td>
<td>0.426</td>
<td>0.424</td>
</tr>
<tr>
<td>Inmigration</td>
<td>0.737</td>
<td>0.185</td>
<td>0.777</td>
<td>1.001</td>
<td>0.203</td>
<td>0.642</td>
<td>0.110</td>
<td>0.744</td>
<td>0.068</td>
<td>0.371</td>
<td>0.073</td>
<td>0.092</td>
<td>0.155</td>
<td>0.185</td>
</tr>
</tbody>
</table>
Table 5. Average Values of 3 Matched Variables for Housing Cluster Groups for San Diego-Tijuana Twin City, 1990

<table>
<thead>
<tr>
<th>Cluster Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Area Units in San Diego</td>
<td>18</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>247</td>
<td>340</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>7</td>
<td>530</td>
<td>0</td>
</tr>
<tr>
<td>Small Area Units in Tijuana</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>27</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Small Area Units - Total</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>27</td>
<td>5</td>
<td>247</td>
<td>340</td>
<td>2</td>
<td>37</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>532</td>
<td>38</td>
</tr>
</tbody>
</table>

**Housing Variable Group**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>One Bedroom Housing</td>
<td>0.035</td>
<td>0.422</td>
<td>0.426</td>
<td>0.231</td>
<td>0.342</td>
<td>0.454</td>
<td>0.249</td>
<td>0.097</td>
<td>0.145</td>
<td>0.079</td>
<td>0.042</td>
<td>0.789</td>
<td>0.051</td>
<td>0.080</td>
</tr>
<tr>
<td>Separate Kitchen</td>
<td>0.994</td>
<td>0.882</td>
<td>0.358</td>
<td>0.548</td>
<td>0.609</td>
<td>0.994</td>
<td>0.994</td>
<td>0.526</td>
<td>0.659</td>
<td>0.237</td>
<td>0.797</td>
<td>0.970</td>
<td>0.998</td>
<td>0.772</td>
</tr>
<tr>
<td>Public Sewer</td>
<td>0.313</td>
<td>0.993</td>
<td>0.079</td>
<td>0.089</td>
<td>0.976</td>
<td>0.998</td>
<td>0.994</td>
<td>1.000</td>
<td>0.234</td>
<td>1.000</td>
<td>0.827</td>
<td>1.000</td>
<td>0.990</td>
<td>0.216</td>
</tr>
</tbody>
</table>
Table 6.
Indices of commonality for the San Diego-Tijuana twin city, 1990

<table>
<thead>
<tr>
<th>Binationality Indices -- A Priori Variable Groups</th>
<th>Overall</th>
<th>Population</th>
<th>Social</th>
<th>Economic</th>
<th>Housing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.015</td>
<td>0.102</td>
<td>0.062</td>
<td>0.126</td>
<td>0.005</td>
<td></td>
</tr>
</tbody>
</table>

A priori population group consists of the variables: population, dependency, gender nativity, and immigration.
A priori social group consists of the variables: primary education, secondary education, and married.
A priori economic group consists of the variables: economic activity, unemployment, underemployment, poverty, and home ownership.
A priori housing variable group consists of the variables: one bedroom housing, separate kitchen, and sewer.

<table>
<thead>
<tr>
<th>Binationality Indices -- Data Driven Variable Groups</th>
<th>Overall</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.015</td>
<td>0.137</td>
<td>0.309</td>
<td>0.310</td>
<td>0.016</td>
<td></td>
</tr>
</tbody>
</table>

Data driven group 1 consists of the variables: population, dependency, gender, immigration, and primary education.
Data driven group 2 consists of the variables: married, native to state, and home ownership.
Data driven group 3 consists of the variables: secondary education, unemployment, underemployment, poverty, and one bedroom housing.
Data driven group 4 consists of the variables: economic activity, separate kitchen, and sewer.
Fig. 1. Burgess's concentric ring model

Source: based on Burgess [4].
Fig. 2. Hoffman's Mexican border urban structure model

Source: Based on Arreola and Curtis [2]
Figure 3. A Priori Cluster Analysis, Population Group, Clusters 1-7, San Diego and Tijuana, 1990
Figure 4. A Priori Cluster Analysis, Population Group, Clusters 8-14, San Diego and Tijuana, 1990
Figure 5. A Priori Cluster Analysis, Housing Group, Clusters 1-7, San Diego and Tijuana, 1990
Figure 6. A Priori Cluster Analysis, Housing Group, Clusters 8-14, San Diego and Tijuana, 1990
Figure 7. Data Driven Cluster Analysis, Variable Group 3, Clusters 1-7, San Diego and Tijuana, 1990
Figure 8. Data Driven Cluster Analysis, Variable Group 3, Clusters 8-14, San Diego and Tijuana, 1990
Figure 9. Secondary Education, San Diego and Tijuana, 1990