Waking Up from the American Dream

A Study of Foreclosures in Mesa County, Colorado

For eight consecutive months in 2006, Colorado had a greater number of home foreclosures than any other state in the country. Largely attributable to the Denver housing market, this dubious statewide distinction raises questions as to the foreclosure rates and trends in communities beyond the state’s capitol, where little analytical research has been conducted. This project delves into the local foreclosure data available in Grand Junction and Mesa County to determine its spatial concentrations, and establish demographic trends that may be contributing to the foreclosure rate in the largest city of western Colorado.

Garrett Groves
LBJ School of Public Affairs, University of Texas
Executive Summary

For eight consecutive months in 2006, Colorado had a greater number of home foreclosures than any other state in the country. Largely attributable to the Denver housing market, this dubious statewide distinction raises questions as to the foreclosure rates and trends in communities beyond the state's capitol, where little analytical research has been conducted. This project delves into the local foreclosure data available in Grand Junction and Mesa County to determine its spatial concentrations, and establish demographic trends that may be contributing to the foreclosure rate in the largest city of western Colorado.

The recent home mortgage crisis is unique not only in size, but also in nature. Caused by the rapid spread of subprime housing loans, few experts predicted the subsequent collapse of the nation’s financial and banking sectors. At the time of this writing, the full extent of the crisis is still not well understood. At its heart is the uncertainty of knowing how many more families are in jeopardy of losing their homes, and how the growing number of foreclosures might ripple through the economy.

The following study focuses on three specific demographic trends: income level, racial minority concentrations, and educational attainment. Each has a plausible link to the spread of subprime mortgage loans, and can be used to study foreclosure data in the aggregate.

These demographic characteristics were used in combination with a set of control variables in an effort to identify spatial patterns in the occurrence of home foreclosures. From this analysis a series of models were fit to a set of geocoded addresses on currently foreclosed homes, as compiled by the Mesa County Trustees Office. These were then used to forecast areas with the highest propensity for future home foreclosures.

During the creation and refinement of these models, new data was routinely collected from the Trustees Office over a four-week period, resulting in a second set of testable data. In predicting the location of these foreclosures, the models performed better than anticipated, with two predicting nearly 50 percent of future home foreclosures.
Introduction

A home foreclosure is more than a devastating financial event. It forces an eviction that can be humiliating as well as disruptive to nearly every aspect of that family or individual's life. However, when even a small number of foreclosures occur in near proximity to each other, a series of cascading events can occur that lower adjacent home values, slow or halt otherwise healthy development in neighboring areas, and even depress the community’s overall economy.

Colorado is experiencing an accentuated version of a national trend in rising home foreclosures due to the proliferation of nontraditional loans to subprime or high-interest borrowers. The cause is important. Unlike previous sources of foreclosures, such as unemployment, declining population, or stagnant and falling wages, this particular trend is not the result of poor economic fundamentals. Grand Junction and Mesa County are now experiencing a rise in almost every indicator that previously would have indicated a strong local economy. And yet, a sudden burgeoning of foreclosed homes has appeared, and in patterns that suggest a strong correlation with certain socio-economic data, such as low-income populations. However, the spatial concentration of subprime borrowers and their demographic characteristics are not well understood.

For this reason, the three most plausible demographic characteristics were selected and mapped for the Grand Junction area by census blockgroups. Blockgroups are generally composed of a few thousand individuals, and include less than a thousand homes, making them the ideal size for gathering and displaying aggregate demographic data without losing the level of detail required for the study.

Minority Concentration

The first demographic trend singled out for analysis was racial minority concentration. A growing number of studies have begun to draw attention to the higher foreclosure rates among minorities, including a troubling analysis done as early as 2006 by the National Community Reinvestment Coalition (NCRC). Concerned for those who fall into a “high-cost” mortgage category, they found that this included 35 percent of American Indians, 40 percent of Hispanics, and almost 55 percent of African Americans. By contrast, only 23 percent of Whites fell into the same category. A more recent report released last July confirms that these trends remain entrenched (National Committee Reinvestment Coalition 2008).

Grand Junction has a lower instance of minority representation than the State of Colorado as a whole. According to the Census Bureau’s decennial census of 2000, nearly 92 percent of Grand Junction’s population was White, and only 0.6 percent were African American (U.S. Census Bureau 2000). However, over 10 percent claimed Hispanic or Latino heritage, which if the NCRC’s numbers hold true, would place as many as 1,620 Hispanic individuals at an especially susceptible risk for foreclosure.
Income Level

This variable has the most obvious connection to the mortgage crisis, and is also one of the most difficult correlations to specify. It should come as no surprise that the amount of income a householder earns has a direct impact on his or her ability to pay a mortgage. However, the strongest correlations appear at the center of the income distribution. Initially counterintuitive, the causes are relatively easy to explain. Though lower income populations are more susceptible to foreclosures, those making the lowest wages often do not have access to lenders or sufficient funding. As income levels increase, people reach a point when their first mortgage becomes a tantalizing possibility. At the opposite end of the scale, as income increases it becomes easier not only to pay the mortgage, but also to remain employed and financially secure.

One of the key characteristics of the current subprime mortgage crisis is the way that predatory lending and financial instability have encroached on both ends of the spectrum. On the one hand, these low cost/high interest loans were made readily available during the recent economic boom to an increasing number of the previously 'home poor'. On the other hand, as the financial and economic pain has spread, so to has its impact on populations who previously viewed their housing situation as anything but precarious. These combined trends make it difficult to draw a distinction between the foreclosure vulnerable middle, and those on the far ends.
Education Level

This final demographic characteristic holds the frailest connection to the mortgage crisis. It is part of an explanatory theory that assumes predatory lending as the primary culprit, and argues that Americans with a stronger educational background stood the best chance of not only avoiding the nefarious practice, but of recognizing it while it was occurring. A home mortgage is after all a financial contract entered into by two willing parties. Further, mortgages are long-term commitments of incredible cost, which presumably aren’t made often in a person’s lifetime.

Education level also serves as a proxy for the immeasurable effects of social and financial opportunity, as well as a specific kind of worldly experience. At the aggregate, attendance in college introduces people to a different world. The effect such an experience may have on one’s decision to purchase a home however has not been well explored.

Supporting information and structure for the theories included here were partially gleaned from both of Robert Shiller’s works The Subprime Solution, and The New Financial Order, as well as Chris Morris’ writings in The Trillion Dollar Meltdown.

Most useful, however, were a series of reports conducted by the Denver Office of Economic Development. The first, entitled Understanding Mortgage Foreclosures in Denver Colorado, is an in-depth study of the home foreclosure phenomenon in Denver, Colorado. Concluded in March 2008, it contains a wealth of analytical data and research, with a particular focus on housing market trends and spatial concentrations, as well as the impact of foreclosure density on surrounding home prices.

The study includes a cursory analysis of racial data, and provides colluding evidence of higher foreclosure rates for both African and American Indian populations. However, the study fails to follow up with a further analysis, and does not provide specific information on the Hispanic community other than a median level of income. This is a glaring omission for a minority group that accounts for almost 32 percent of the city’s population (Statistics 2008).

Research Question

Are there spatial socio-economic trends to the foreclosure data in Mesa County, Colorado? More specifically, can a careful analysis of current home foreclosures lead to the creation of an accurate and precise tool for predicting where future foreclosures are most likely to occur?
Methodology

At the start of this project, I set out to map foreclosed homes in Grand Junction and Mesa County, and overlay that data with demographic information designed to convey poverty, educational and racial trends within the community.

The Information Required

Before I could begin, a list of recent home foreclosures had to be located, as the project hinged on the accuracy of this information. I also had to locate road shapefiles of the area, which could be used to build an address locator for the project. In order to set a visual context for my maps, I also intended to find data to help create the physical shape of the valley that surrounds the city, and the Colorado River that flows through it.

Demographic information, the second focus point of the study, was housed in the Census Bureau’s website. I needed to familiarize myself with that database, and determine the source of the most recent demographics that were detailed enough so as to provide differentiated data for the city and its surroundings. In addition to income, education and racial status, I was also interested in downloading a number of additional demographic trends, including age and sex, in order to test the robustness of my theory.

Beyond demographics, there were a number of control variables that had to be located. Among these were detailed lists of housing stocks, as well as the distinction between urban and rural sub-areas. These would be needed to stabilize the models, and hopefully provide a more accurate prediction of future foreclosures.

Lastly, I needed to make a distinction within my data between the ‘current’ foreclosure data that would be used to build the models, and the ‘future’ data used to test them.

Data Collection

After some searching and several phone calls, I located a list that the Mesa County Trustees Office publishes weekly with the area’s currently foreclosed homes, provided in a general spreadsheet format. Though this didn’t include a comprehensive list of all foreclosures over the past several years, it did provide almost 200 data points for the current study. Additionally, it provided me with an excellent solution to the present/future foreclosure data quandary required to test my solutions. I was able to use the 200 initial data points to build the models, and then upload a further 30 addresses that had all fallen into foreclosure during that time. Unfortunately, I was unable to obtain a complete listing of years past, which has prevented me from conducting a background analysis on foreclosure trends in Mesa County.

All necessary shapefiles were attained through Mesa County’s GIS website, and included information not only on city streets, but the area’s rivers and lakes as well. I was unable to locate any useful data to provide a visual context of the surrounding valley.
Demographic information for Mesa County was available at the Census Bureau, which had to be combined with the Census’ tigerline shapefiles. In order to obtain the level of detailed required, I chose the blockgroup geo setting. Blockgroups are designed to include information on only a few thousand people, or more importantly for this analysis, on fewer than a thousand homes. This was an important distinction, as my analysis required a large enough group to support a statistical study of a population, and yet small enough to create as many separate populations as possible. For these reasons, the block level was determined to be too specific, and the tract level too broad.

The Census Bureau is gradually limiting the in-depth survey data that it collects, and is steadily restricting the geographically detailed information to the decennial censuses. The information I was searching for turned out to be no exception. Though there was a small possibility of updating a few of the more general demographic statistics I required, the vast majority were available only in the 2000 census format. This created a large discrepancy between my immediately recent foreclosure data, and the eight-year-old demographic trends I would use to analyze them.

Within the 2000 Census, I relied exclusively on the Summary File 3 (SF3) survey data. This resource included a far greater amount of information on demographic trends, as well as the control variables I hoped to locate. In each instance, I downloaded all blockgroups for Mesa County.

Income
I decided to base my income demographic on the median household income for each blockgroup as it was presented in the 2000 census. Fortunately, the Census provided table P53 contained the pre-calculated median household incomes.

Educational Attainment
For this demographic trend I condensed a series of Census tables into an aggregate percentage of those who have not attended any college within each blockgroup. The Census table is QT-P20, which is a collection of other tables that collectively breaks each blockgroup into the quantity of people who have completed each grade level of education by sex. For each gender, I compiled the data into two categories: those with a high school diploma or less, and all those who attended at least some college. I then combined the categories for both genders, and divided each by the total population for that respective blockgroup.

Racial Minority Concentration
To determine the concentrations of racial minority status in Grand Junction, I needed to find a data source that included Hispanic, as well as other race categories. Not all of the Census’s tables include information on the Latino population specifically. This was important because I wanted to insure that I could include this significant population in my calculations, as the Hispanic community outnumbers the black community in Grand Junction by nearly twenty to one. The Census table with data on Hispanics is P7. To obtain the percent racial minority, I obtained the
minority population by subtracting the *non-Hispanic White alone* category from the total population, and divided the result by the total population.

**Control Variables**
I used five additional Census data sets during this analysis, four of which were directly worked into the model as independent variables in the models. Though the number of housing units was not used as a separate variable, it was central to a number of the forthcoming calculations and manipulations. This information was located in the H1 table.

**Owner Occupied Homes**
This variable helped control for the high concentration of rental properties in the heart of the city, which would have otherwise adversely impacted the accuracy of the model, since renters are incapable of defaulting on home mortgages that they don’t have. The Census’ *Tenure* table, titled H7, provided this information. To calculate the percent of owner occupied homes, I simply divided the number of owned homes by the total housing units.

**Non-Mortgage Homes**
The Census also provided a aggregate number of homes in each blockgroup which were wholly owned without need of a mortgage. These non-mortgage homes provided a relatively accurate percentage that could be used to control for this population, which like the opposite case for renters, are also incapable of defaulting on mortgages if they don’t have one. The Census provides a table H80, titled *Mortgage Status*, with this information. To obtain the percent of non-mortgage homes, I divided this population by the total housing units.

**Urban Populations**
One of the primary suspected contributors to growing home foreclosure trends is proximity to other foreclosed homes. This is a phenomenon that is not applicable to rural areas, where the distance between homes is often measured in portions of miles, rather than feet. I calculated percent urban homes by dividing the number of urban housing units by the total unites, each provided by the Census’ H5 table.

**Farm Residences**
A variable that differs significantly from rural residences, are the subset of homes specific to farming. Grand Junction is located in the heart of once thriving farming community. Many of these residences remain on the edges of the city, and are distinct from other urban and rural residences in two primary ways. First, they earn their livelihood by working the land in at least some minimal fashion, which places them into a distinctly different category than most residences. How this might be a factor in the current subprime mortgage crisis is as yet unknown. Second, and perhaps more relevant, is the
ability farming families have for selling portions of their land in order to
preserve their homes. Though this opportunity may not be as readily
accessible in other parts of the country, it has been a growing trend among
Mesa County farmers for a number of years. As indicated earlier, Grand
Junction’s economy has continued to thrive up until this point. As the
population expands, farmers have been able to make a sizable profit by
turning in the tractors and selling portions of their imminently developable
lands. I calculated the percent farm residences by dividing the number of
farm residences by the total housing units. The Census provides this data in
table same table that contained the urban and rural information above, H5.

**Analyze the Concentration of Foreclosures and Socio-Economic Indicators**

The first task required geocoding the foreclosure addresses and mapping them to the street
layer for Mesa County. Because home foreclosures are focused almost exclusively in Grand
Junction and the neighboring town of Fruita, the majority of maps for this project focused
on the area surrounding these two cities. Additional data layers will were then created that
displayed the concentration of housing stock, as well as the number of foreclosures per
census blockgroup, normalized by housing units.

A series of maps were then created to establish if Mesa County exhibits high concentrations
in poverty, low educational attainment or minority representation, aggregated to the census
blockgroup. Each of these concentrations was classified into six categories using standard
breaks. Additional maps were also created for each of the four control variables in the exact
same manner.

**Calculate Foreclosure Propensity: The Models**

These geographic trends and their underlying data were then dissolved together into a single
layer to create a series of theoretical models, designed to indicate high propensity areas for
home foreclosures. There were four models in total.

**Basic Socio-Economic Model**

This analysis focused exclusively on the three demographic trends: income,
education and race. Each variable was converted into a raster format, and
reclassified to create scoring system from one to six.

With minority status for example, the areas with the highest concentrations were
rated with a six. The next highest received a five, and so on all the way to one. With
education, the numbering measured the highest concentrations of people without
college experience. Income was graded slightly differently, as per the initial
discussion above. Here, the middle ranks received the highest rating, where the two
extremes received lower ratings.

With each raster completed, the raster calculator was then used to create a single
raster model, one that summed the total ratings for each raster pixel. Each variable
was weighted evenly in the basic model:
HF = 0.33I + 0.33E + 0.33M

The resulting raster was then reclassified into the standard 6 categories using normal breaks. The final model was completed once the raster was reconverted into a shapefile.

**Complete Socio-Economic Model**

This analysis combined the four control variables to the original demographic trends. Repeating the same process as before, each variable was converted into a raster format, and reclassified to create scoring system from one to six.

The raster calculator was then used to create a single raster model using all seven variables, each weighted equally.

HF = 0.15I + 0.15E + 0.15M + 0.15F + 0.15U + 0.15NM + 0.15O

The resulting raster was then reclassified into the standard 6 categories using normal breaks. The final model was completed once the raster was reconverted into a shapefile.

**Regression Model**

This analysis utilized the statistical program STATA to run a simple linear regression using all seven variables. After calculating for a number of various combinations (using other control variables which were not included in the final project) a set of preliminary results were chosen as a guide. This is because the vast majority of scenarios returned results that were not statistically significant. The table below includes information from the most conclusive test, though all pertinent variables are insignificant save owner occupied homes. The regression was run using the raw foreclosure count as the dependent variable. This required including the additional control variable for the number of housing units. Also, Income was left out of the test, as the author was unable to re-rank the data so as to reflect the unique rating of income mentioned earlier:

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>T-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing Units</td>
<td>.003</td>
<td>.001</td>
<td>3.28</td>
</tr>
<tr>
<td>Education</td>
<td>1.561</td>
<td>2.257</td>
<td>0.69</td>
</tr>
<tr>
<td>Minority Status</td>
<td>.575</td>
<td>3.192</td>
<td>0.18</td>
</tr>
<tr>
<td>Farm Residence</td>
<td>-6.205</td>
<td>8.557</td>
<td>-0.73</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.356</td>
<td>1.187</td>
<td>-1.14</td>
</tr>
<tr>
<td>Non-Mortgage</td>
<td>-1.854</td>
<td>1.570</td>
<td>-1.18</td>
</tr>
<tr>
<td>Owner Concen.</td>
<td>4.604</td>
<td>1.394</td>
<td>3.30</td>
</tr>
</tbody>
</table>

There are a number of problems with the above statistical analysis. For instance, several of the coefficient signs are different than had been predicted. Also, the exclusion of income makes it difficult to predict what its impact may have been.

---

1 HF = High Propensity Areas for Home Foreclosures
I = Income
E = Low Educational Attainment
M = Minority Status
F = Farm Residence Concentration
U = Urban Concentration
NM = Non-Mortgage Concentration
O = Owner Concentration
Finally, no set of answers was consistent over several trials with differing variables. Further research and analysis may be able to establish a more robust set of answers. However, for this analysis, the sample coefficients gathered in the table above have been used in order to demonstrate the possible future uses of this model design.

Using the variable rasters created for the Complete model, the raster calculator was then used to form a single raster model using the coefficients gathered from the statistical test. Each coefficient was divided by the sum total of all, and its percent used in setting the formula below. For simplicities sake, the income variable was given a value equal to education:

\[
HF = 0.1I + 0.1E + 0.04M + 0.35F + 0.06U + 0.1NM + 0.27O
\]

The resulting raster was then reclassified into the standard 6 categories using normal breaks. The final model was completed once the raster was reconverted into a shapefile.

**Spatial Derived Socio-Economic Model**

This final model used ArcMap’s spatial recognition tools to construct relative weights based on the percentage of foreclosures located in each variable’s highest concentration. Per the example below, those variables with the greatest concentrations in their highest field (by percent) received the strongest weights.

![Home Foreclosure Concentration by Control Variables](image)

To calculate the weights, the percent of foreclosures for each variable in its highest field was divided by the total for all variables result in the same field. The results can be seen in the equation used below:

\[
HF = 0.21I + 0.06E + 0.05M + 0.25F + 0.22U + 0.13NM + 0.07O
\]
The resulting raster was then reclassified into the standard 6 categories using normal breaks. The final model was completed once the raster was reconverted into a shapefile.

**Comparative Analysis**

The final step is to contrast the four models based on their ability to predict the location of the 30 new home foreclosures. This required geocoding the new addresses, and joining them to the model shapefiles. Once completed, it was possible to export that data into an excel file, calculate the percentage of foreclosed homes that fell into each model’s highest range, and chart the discrepancies.

A follow-up comparative method was also pursued, one that required calculating the total the total area for each ranked field for every model. This provided a method to test the precision as well as the accuracy of each model, simply by dividing the number of foreclosed homes in the highest category by the area (in square miles) that the model used to predict those foreclosures.
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Home Foreclosure Locations and Density

Foreclosed Homes
Rivers
Lakes
Major Highways
Roads

Foreclosure Concentrations

Lowest
Low
Low Moderate
Moderate
High
Highest

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee
Projection: UTM Zone 12N (NAD83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Home Foreclosure Locations and Density

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Minority Racial Concentration

Minority Concentrations

Foreclosed Homes
Rivers
Lakes
Major Highways
Roads

Lowest
Low
Low Moderate
Moderate
High
Highest
High Density Foreclosures

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee
Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Income Level: Median Income

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Educational Attainment Level: High School or Less

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)

© Home Foreclosures in Mesa County, Colorado
Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Farm Concentration
Farm Residence Concentration

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee
Projection: UTM Zone 12N (NAD83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Urban Concentration

Sources: Mesa County, Mesa County Public Trustee
Projection: UTM Zone 12N (NAD83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Home Owner Concentration

Sources: Mesa County
Mesa County Public Trustee
Projection: UTM Zone 12N (NAD83 Feet)

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

© Mesa County
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD83 Feet)
Waking Up from the American Dream

Home Foreclosures in Mesa County, Colorado

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Complete Model

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Regression Model

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
Waking Up from the American Dream
Home Foreclosures in Mesa County, Colorado

Spatially Derived Model

Newly Foreclosed Homes
Rivers
Lakes
Major Highways
Roads
Foreclosure Propensity
- Lowest
- Low
- Low Moderate
- Moderate
- High
- Highest

Garrett C. Groves
GIS Specialist & Policy Analyst
Graduate Student: LBJ School of Public Affairs

Sources: Mesa County
Mesa County Public Trustee

Projection: UTM Zone 12N (NAD 83 Feet)
The first intriguing conclusions resulted from the comparison of each variable’s correlation to home foreclosures independently:

The adjusted income variable had the highest correlation with foreclosures. The minority variable showed some promising results in the high category, but dropped off to a level below the education variable for the highest field. Alternately, the education variable was rendered almost meaningless, showing a relatively flat correlation throughout its fields. This may lend evidence to the theory that predatory subprime mortgage lending makes relatively little distinction between individuals with a high school or college education. The exact percentages for each variable is provided in the table below:

<table>
<thead>
<tr>
<th>Category</th>
<th>Income</th>
<th>Minority</th>
<th>Education</th>
<th>Farm</th>
<th>Urban</th>
<th>Mortgage</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest</td>
<td>72</td>
<td>17</td>
<td>21</td>
<td>84</td>
<td>75</td>
<td>44</td>
<td>23</td>
</tr>
<tr>
<td>High</td>
<td>38</td>
<td>41</td>
<td>23</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Moderate</td>
<td>10</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>8</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Low Moderate</td>
<td>1</td>
<td>9</td>
<td>26</td>
<td>5</td>
<td>1</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>Low</td>
<td>7</td>
<td>17</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Lowest</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>8</td>
<td>16</td>
</tr>
</tbody>
</table>

*All Values in Percents

Among the control statistics, both the farm and urban variables performed surprisingly well, with strong correlations between their highest field and home foreclosures. It is important to note that the farm variable was inverted during the analysis. The results should be
understood as indicating that farm residences have a drastically lower incidence of foreclosure.\(^2\) This data lends credible support to the explanations that were laid out above in the methodology section were all the variables were introduced. The mortgage variable also performed well, but not as convincingly as one might have predicted. Surprisingly, the owner concentration variable performed especially poor. This could be most easily explained by the strong correlation that occurs between income and ownership, decreasing the likelihood that areas with high ownership will also be areas where foreclosure occurs.

**Model Performance**

All four models performed not only well, but remarkably similar given their seemingly varied composition. The most impressive result was the prediction results of both the Complete and the Regression models, which predicted the blockgroups where nearly 50 percent of the new foreclosures occurred, as indicated in the adjacent graph. Both models also began strong with perfect predictions in the lowest categories as well. The other two were not far behind, with the basic model scoring the highest at only six percent. The Spatial model’s performance was relatively disappointing to the author, as the precise and near scientific method in which it was calculated was convincing at the outset.

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Basic</th>
<th>Complete</th>
<th>Spatial</th>
<th>Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest</td>
<td>38</td>
<td>45</td>
<td>38</td>
<td>48</td>
</tr>
<tr>
<td>High</td>
<td>25</td>
<td>21</td>
<td>31</td>
<td>24</td>
</tr>
<tr>
<td>Moderate</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Low Mod.</td>
<td>6</td>
<td>10</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Low</td>
<td>13</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Lowest</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^{*}\)All Values in Percents

Ultimately, the relative strength of each reveals the need for further in-depth analysis. In addition to testing with a larger group of future foreclosures, the four models would also benefit from some extended experimentation. It could prove useful to another slate of variable combinations.

\(^2\) All of the above variables were normalized by population or total housing units accordingly.
A second method for comparing the performance of each model may help provide a direction for their improvement. Though each model performed relatively equal in predicting foreclosure locations, each specified different quantities of land as having the highest propensity for mortgage defaults. If we take the number of home foreclosures each model accurately predicted, and divide that by the square miles required to make the prediction, we arrive at measure of precision (which has been multiplied by ten in this case, in order to arrive at a number comparable with the other data below). This comparative measure will create a more useful tool for policy makers and community groups who have limited resources at their disposal, and would welcome smaller areas in which to focus their efforts.

As indicated in the graph above, the Spatial and Complete models come closest to matching their precision scores with the area they designate as having the highest propensity for foreclosure. It is not surprising that these models outperformed the others. The Basic model lacks the explanatory power of the others because it only has three variables instead of the full seven. On the other hand, the Regression model strongly favors both the Farm and Owner control variables at the expense of the other five, presumably weakening its precision. Consequently, only the Spatial and Complete models represent all seven variables fairly.

<table>
<thead>
<tr>
<th>Model</th>
<th>Area (sq mi)</th>
<th>Foreclosures</th>
<th>Precision Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>30.2</td>
<td>13</td>
<td>4.64</td>
</tr>
<tr>
<td>Spatial</td>
<td>20.4</td>
<td>14</td>
<td>6.88</td>
</tr>
<tr>
<td>Complete</td>
<td>20.3</td>
<td>11</td>
<td>5.43</td>
</tr>
<tr>
<td>Basic</td>
<td>31.3</td>
<td>12</td>
<td>3.83</td>
</tr>
</tbody>
</table>

It is also encouraging that the Spatial model performed well, given the precision with which it was crafted. It used roughly the same area as the complete model, and predicted three more foreclosures, making it the best model in this first study of home foreclosures in Grand Junction, Colorado.
Before concluding this study, it is important to draw attention not only to the positive results, but also the areas in need of further study and refinement. Perhaps the most important addition to the work presented here would be the inclusion of a more extensive home foreclosure listing. Though there are certainly countless ways that future models can be constructed, they will remain relatively useless without a more data to test. This point is made clear in the table above, where the difference between the best and worst models is only three successful foreclosure predictions.

The other major adjustment required is a significant upgrade in the demographic trends that are used to project current foreclosures. Much can change in a few year’s time, and it will be impossible to know how accurate the models truly are before the next decennial census is made available in 2010.

Finally, it is also possible that Grand Junction and the surrounding area is simply too small a region to provide the optimum conditions to distinguish demographic trends between aggregate blockgroup populations. The study area included a total of only 42 parcels. An urban center with a much greater set of data should help to minimize outliers and stabilize the models.

However, even the above concerns cannot obscure completely the relevance of predicting nearly 50 percent of the region's future foreclosures over an area of nearly 150 square miles. The next step should include a thorough analysis of the concerns, and work to refine a set of models that have already shown some promise.

**Primary References**

*Denver Office of Economic Development:*

**Neighborhood Stabilization Plan:** Link to the Office’s current plan to allocate $6.1 million dollars to acquire and redevelopment foreclosed properties, with a focus on the most highly concentrated areas of home foreclosures.

**Understanding Mortgage Foreclosures in Denver Colorado:** An in-depth study of the home foreclosure phenomenon in Denver Colorado, concluded in March 2008:


All shapefiles have been projected into: **UTM, UNIT: Meter, ZONE: 12, DATUM: NAD83 (1992)**

- **Mesa County Trustee’s Office:**

- **Mesa County GIS website:**

- **U.S. Census Bureau**
Appendix

Template Creation

- Uploaded all shapefiles into ArcCatalog, defined each accordingly, and then projected all into the format used by Mesa County: UTM, UNIT: Meter, ZONE: 12, DATUM: NAD83 (1992)
- Uploaded the Mesa County road, lake and river shapefiles into ArcMap.
- Centered the extent window on both the Grand Junction and Fruita townships, so as to include all of the most densely populated areas.
- Selected portions of the two major highways that run through the area and created a separate layer in order to create labels for both I-70 and Hwy 6&50, which were added in the attribute table using the editor tool.
- Followed the same process to create separate labels for the two rivers, the Colorado and the Gunnison.
- Uploaded the Census tigerline shapefile for the regions block groups, which would serve as the platform for the majority of my ensuing work.
- Inserted a new data frame, into which I placed the same blockgroup shapefile. The frame was reduced in size, and used as a reference map. I also uploaded the rivers shapefile, and created a separate roads layer from the original, which included only the county’s interstates. This file was also uploaded to the reference map. I then included an extent rectangle to outline the area of the larger map.
- The final additional elements include the Mesa County logo, north arrow, scale, legend, sources, projection information and contact information.
- I also added the geocoded home foreclosure shapefile as detailed below.

Address Locator Creation Geocoding results

- In ArcCatalog I created a new address locator, using the US Streets with Zone option.
- I then specified the Mesa County road shapefile, and left the rest of the fields as they appeared within the ensuing primary table.
- Next I uploaded the new address locator to the template file, and geocoded the foreclosure list with the new address locator.
- Once crucial step for this particular list of addresses was to leave the zone field in the address locator set as “none”.
- I also needed to do a manual merge in excel in order to put the house number in the same cell as the street name.
- A number of foreclosures did not match. I was able to link almost all of them manually.
- I then spatially joined the resulting geocoded data to the Mesa County blockgroups shapefile.

Home Foreclosure Locations and Density Map
• I began with the template file above, and uploaded the census date for housing units by blockgroup and joined it to the Mesa County shapefile.
• In the symbology, I set the value to the foreclosure cnt, and normalized with the housing unit data.
• I then set the colors, and inserted the north arrow, legend, scale, source & author information.

Education, Race, Income and Control Variable Map Creation
• Uploaded the appropriate demographic spreadsheet into the ArcMap template file, and joined it with blockgroup shapefile.
• On the symbology tab I classified the field to 6 categories with natural breaks.
• I then joined the blockgroup file with the foreclosure shape file.
• Next step to prepare for model creation was to transform the newly created demographic shape file into a raster.
• I then reclassified this raster in 6 categories by natural breaks, numbered 1-6. It was most important to make sure that the new classifications were correct, and not backward.
• I then reconverted the raster into a shapefile.
• Next I needed to rejoin the foreclosure file to this new shapefile, in order to have an attribute table prepared for the spatially derived model procedure.
• This procedure was repeated for each of the variables in the title.

Basic Model
• The first task required uploading each of the variables’ attribute tables into excel, and specifically the tables that contained the final raster conversion data as well as the original foreclosures counts.
• The spreadsheets were then combined, and totals calculated for the number of foreclosures in the highest propensity classification for each variable.
• This was then divided by the total number of “points” accumulated by all the variables added together to obtain the final weights used below to create the model.
• The raster calculator was then used to create the adjusted raster model, per the following formula:  \( HF = 0.33I + 0.33E + 0.33M^3 \)
• This raster was then converted to a vector shapefile, and new future foreclosure data joined with it in order to repeat the process above through the second bullet point. This information was then used to compare the relative accuracy of each of the models.

Other Models
• I followed the same steps above, respective to the appropriate model.
• The equations for each of the models are included below:

\( HF = \) High Propensity Areas for Home Foreclosures \( F = \) Farm Residence Concentration
\( I = \) Income \( U = \) Urban Concentration
\( E = \) Low Educational Attainment \( NM = \) Non-Mortgage Concentration
\( M = \) Minority Status \( O = \) Owner Concentration
Complete: \[ HF = 0.15I + 0.15E + 0.15M + 0.15F + 0.15U + 0.15NM + 0.15O \]
Regression: \[ HF = 0.1I + 0.1E + 0.04M + 0.35F + 0.06U + 0.1NM + 0.27O \]
Spatially Derived: \[ HF = 0.21I + 0.06E + 0.05M + 0.25F + 0.22U + 0.13NM + 0.07O \]

Analyses
- All analysis was conducted using excel spreadsheets per the information included in the methodology.

Additional Works Cited