# Exploring Housing Cost Data With Conditioned Choropleth Maps

**Brent D. Mast**
U.S. Department of Housing and Urban Development

*The views expressed in this article are those of the author and do not represent the official positions or policies of the Office of Policy Development and Research or the U.S. Department of Housing and Urban Development.*

Micromaps (for example, see Carr and Pickle, 2010) display multiple maps on the same exhibit, with different geographic units highlighted in each map. A conditioned choropleth map (CCM) is a type of micromap with multiple maps arranged in a panel layout conditional on one or two categorical variables, with highlighted regions shaded according to a main variable of interest. CCMs have been used in a wide variety of applications (for example, see Carr and Pickle, 2010; Carr, Wallin, and Carr, 2000; Carr, White, and MacEachren, 2005; Friendly, 2007).

In this article, I demonstrate how to use CCMs to explore variation in housing cost-to-income ratios (HCIRs) in the 50 states and Washington, D.C. For renters, HCIR equals monthly rent plus utility and fuel costs, divided by monthly household income. For homeowners, monthly costs can include mortgage payments, property insurance, property taxes, utility and fuel costs, condominium fees, and mobile home fees, in addition to other costs.

I analyze the association of HCIRs with two other variables: percentage of the population in urban areas (hereafter, percent urban) and married couple households as a percentage of total households (hereafter, percent married). Data on HCIRs and percent married are from the American Community Survey (ACS) Public Use Microdata Sample (PUMS) for 2007 through 2011 (Ruggles et al., 2010);[1] data on percent urban are from the 2010 census.

State HCIR estimates vary from 23.8 percent in West Virginia to 36.9 percent in Washington, D.C., with a median of 29.5 percent in North Carolina and a mean of 29.9 percent. State percent urban estimates vary from 33.0 percent in Vermont to 100 percent in Washington, D.C., with a median of 73.1 percent in Virginia and a mean of 72.5 percent. State estimates of percent married vary from 19.3 percent in Washington, D.C., to 58.5 percent in Utah, with a median of 46.5 percent in Michigan and a mean of 46.1 percent.

---

[1] The ACS PUMS does not identify married same-sex couples. As such, married couple data in this study are for married opposite-sex couples.
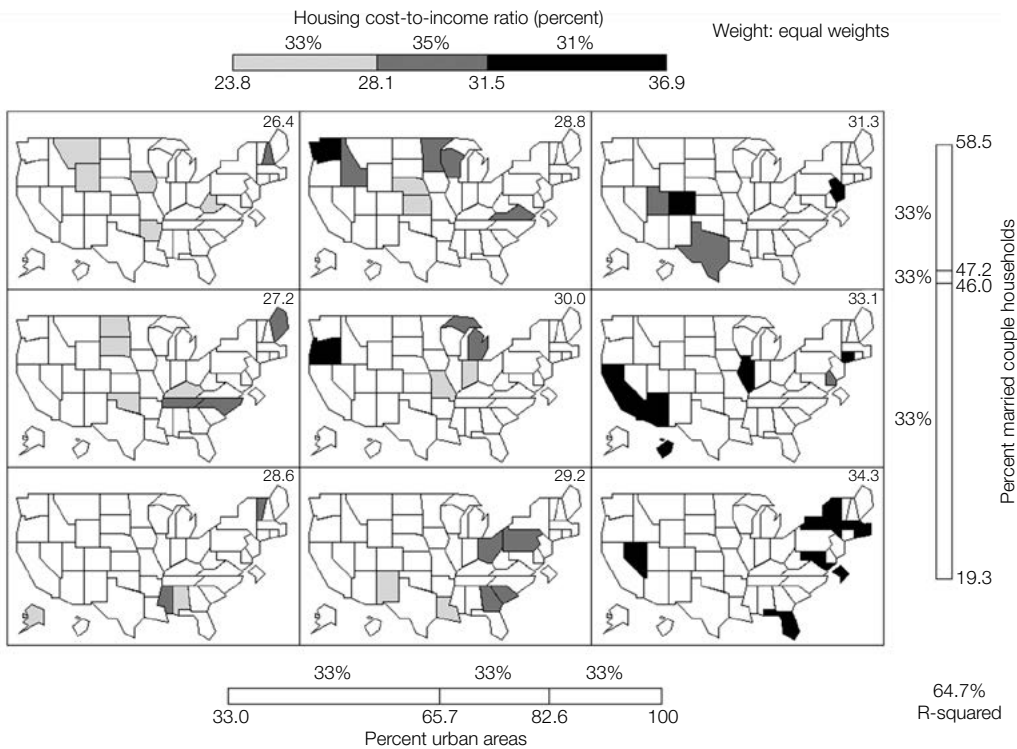
When computing means and percentiles, we could weight geographic units by population or number of households. For this article, I gave all observations equal weight.

Exhibit 1 displays a CCM[2] mapping HCIRs conditional on percent urban and percent married. Exhibit 1 displays nine micromaps in a three-by-three panel layout. Panel columns correspond to three categories of percent urban, indicated by the bottom horizontal slider, and rows correspond to three categories of percent married, indicated by the right vertical slider.

HCIRs in exhibit 1 are shaded according to three categories indicated by the horizontal slider at the top of exhibit. Cut points in exhibit 1 for all three variables are roughly the 33rd and 66th

## Exhibit 1

### Housing Cost-to-Income Ratios Conditional on Percent Urban Areas and Percent Married Couple Households



*Note: Cut points are reported below the top and bottom sliders and to the right of the vertical slider. Percentages of observations in the categories are reported above the top and bottom sliders and to the left of the vertical slider. Mean housing cost-to-income ratios are reported in the upper right-hand corner of each panel. The R-squared reported is for a two-way analysis of variance of housing cost-to-income ratios with the percent urban areas and percent married couple households categorical variables.*

*Sources: American Community Survey 2007–2011 Public Use Microdata Sample (housing cost-to-income ratios and married couple households as a percentage of total households); 2010 census (percentage of the population in urban areas)*

---

[2] The CCMs in the article were produced with Carr et al.'s CCmaps software. http://mason.gmu.edu/~dcarr/. CCMs can also be produced by Bivand et al.'s maptools R package. http://www.cran.r-project.org/web/packages/maptools/index.html.

percentiles. Cut points are reported below the top and bottom sliders, and to the right of the vertical slider. Percentages of observations in the categories are reported above the top and bottom sliders, and to the left of the vertical slider. For each variable, I will refer to the lowest category as "low," the middle category as "medium," and the highest category as "high." In exhibit 1, the 33 percent of observations in the low HCIR category with HCIRs between 23.8 and 28.1 percent are shaded light gray; the 35 percent of observations with medium HCIRs from 28.2 to 31.5 percent are shaded medium gray; and the 31 percent of observations with high HCIRs from 31.6 to 36.9 percent are shaded black.

In exhibit 1, the low percent urban category includes states with percent urban estimates between 33.0 and 65.7 percent; the medium percent urban category is for states with estimates between 65.8 and 82.6 percent; and the high percent urban category includes states with estimates from 82.7 to 100 percent. Likewise, in exhibit 1, the low percent married category includes states with estimates between 19.3 and 46.0 percent; the medium percent married category includes states with estimates between 46.1 and 47.2 percent; and the high percent married category is for states with estimates between 47.3 and 58.5 percent.

Exhibit 1 indicates that HCIRs are positively related with percent urban and negatively related with percent married. In the upper left panel, states with low percent urban and high percent married are highlighted. The upper left panel includes five states with low HCIRs shaded light gray (Arkansas, Iowa, Montana, West Virginia, and Wyoming) and one state with a medium HCIR shaded medium gray (New Hampshire). The mean HCIR for highlighted states is reported in the upper right corner of each panel. The mean HCIR of the six highlighted states in the upper left panel is 26.4 percent, which is the lowest among the nine panel means.

In the lower right panel in exhibit 1, states with high percent urban and low percent married are highlighted. In the lower right panel, Washington, D.C. and all five states (Florida, Maryland, Massachusetts, Nevada, and New York) are shaded black, indicating high HCIRs. The mean HCIR of the six highlighted observations in the lower right panel is 34.3 percent, which is the highest among all panel means.
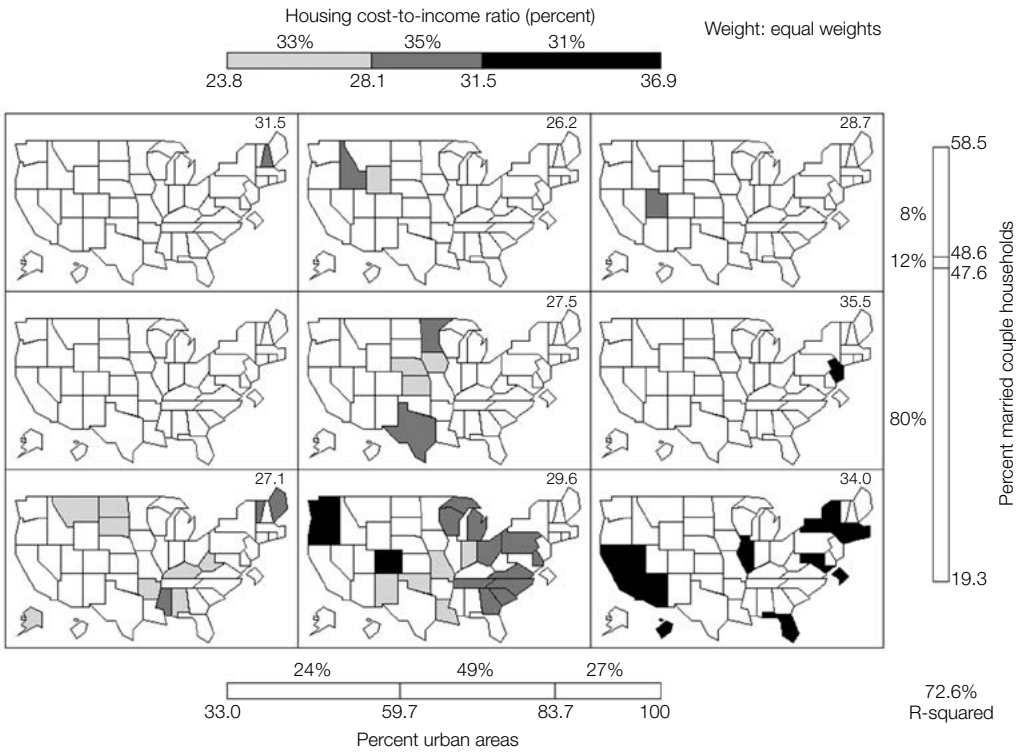
The R-squared from a two-way analysis of variance is reported in the lower right corner of exhibit 1. The R-squared equals 64.7 percent, indicating that the percent urban and percent married categorical variables explain 64.7 percent of the variation in HCIRs.

If we change the cut points used to define the percent urban and percent married categories, we will obtain a different R-squared. Exhibit 2 reports a CCM using 59.7 and 83.7 percent as cut points for the percent urban categories and 47.6 and 48.6 percent as cut points for the percent married categories; these cut points were chosen to approximately maximize the R-squared. The R-squared reported in exhibit 2 is 72.6 percent, which is considerably higher than the R-squared of 64.7 percent reported in exhibit 1.

The CCM is a powerful tool for visualizing and analyzing geographic data. In this article, CCMs clearly demonstrate that much of the variation in state HCIRs can be explained by variation in the percentage of the population living in urban areas and the percentage of households that are married couple households.

**Exhibit 2**

Housing Cost-to-Income Ratios Conditional on Percent Urban Areas and Percent Married Couple Households: Maximum R-Squared



*Note: Cut points are reported below the top and bottom sliders and to the right of the vertical slider. Percentages of observations in the categories are reported above the top and bottom sliders and to the left of the vertical slider. Mean housing cost-to-income ratios are reported in the upper right-hand corner of each panel. The R-squared reported is for a two-way analysis of variance of housing cost-to-income ratios with the percent urban areas and percent married couple households categorical variables.*

*Sources: American Community Survey 2007–2011 Public Use Microdata Sample (housing cost-to-income ratios and married couple households as a percentage of total households); 2010 census (percentage of the population in urban areas)*

## Acknowledgments

## Author

Brent D. Mast is a social science analyst at the U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Program Monitoring and Research Division.

# References

Carr, Daniel B., and Linda Williams Pickle. 2010. *Visualizing Data Patterns With Micromaps.* New York: Chapman and Hall/CRC.

Carr, Daniel B., John F. Wallin, and D. Andrew Carr. 2000. "Two New Templates for Epidemiology Applications: Linked Micromap Plots and Conditioned Choropleth Maps," *Statistics in Medicine* 19 (17–8): 2521–2538.

Carr, Daniel B., Denis White, and Alan M. MacEachren. 2005. "Conditioned Choropleth Maps and Hypothesis Generation," *Annals of the Association of American Geographers* 95 (1): 32–53. Also available at http://www.geovista.psu.edu/publications/2004/Carr_CCmaps_Annals_04.pdf.

Friendly, Michael. 2007. "A.-M. Guerry's Moral Statistics of France: Challenges for Multivariable Spatial Analysis," *Statistical Science* 22 (3): 368–399.

Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota.

# Additional Reading

Wardrip, Keith E., and Danilo Pelletiere. 2008. "Fully Utilizing Housing Cost Data in the American Community Survey PUMS Data: Identifying Issues and Proposing Solutions," *Cityscape: A Journal of Policy Development and Research* 10 (2): 331–339. Also available at http://www.huduser.org/portal/periodicals/cityscpe/vol10num2/ch8.html.