Assessments and Property Tax Variability:

A Quantile Approach

Daniel McMillen

Department of Economics
214 David Kinley Hall
1407 W. Gregory
Urbana, Illinois 61801
217-333-8930
mcmillen@illinois.edu

April 10, 2013

Abstract

The statutory incidence of the property tax typically implies that effective tax rates – the ratio of tax payments to property value – are higher for properties with higher sales prices. The statutory progressivity of the property tax can be reversed if assessment ratios decline with sales prices. Quantile approaches are better suited than standard regression analysis for analyzing the relationship between assessment ratios and sales prices because they provide information on the variability of assessment rates at alternative prices. Using data from Chicago for 2003, 2006, and 2009, I show that assessment rates tend to be both higher and more variable for lower-priced properties. A comparison of the implied distribution of effective rates to statutory rates implies that the statutory progressivity of the property tax is completely reversed, with lower-priced properties having the highest assessment rates.

1. Introduction

The property tax remains by far the most significant source of tax revenue for local governments in the United States. Data from the U.S. Census imply that approximately 75% of local governments' own-source tax revenue was derived from the property tax in 2010 (Barnett and Vidal, 2012). Partly as a result of the property tax's significance – or its "salience" (Cabral and Hoxby, 2013) – various forms of limits have been enacted in nearly every state. According to Anderson (2006), 43 of the 48 contiguous states have some sort of explicit property tax limits. Anderson (2012) suggests that another factor explaining the rise of property tax and assessment limits is the uncertainty associated with assessments. Unlike other common taxes such as income or sales taxes, the property tax's base is not directly observed by either governments or taxpayers. Inaccurate assessments can lead to tax increases for homeowners even if their properties have not actually increased in value.

Although a full analysis of the incidence of the property tax requires knowledge of household incomes and the full set of exemptions and deductions available to a household, a good approximation to the statutory incidence can be obtained by calculating the standard tax payment for a given assessed value. The typical tax schedule is linear: $T = t(A-E)$, where T is the tax payment, t is the tax rate, A is the assessed value, and E is the portion of assessed value that is not taxable – the exemption. Although exemptions may be available for low-income households, the elderly, people with disabilities, and so on, the most common form is the homestead exemption, which typically is a fixed deduction that is available for any owner-occupied property. If the exemption does not vary across households, the average tax burden varies only with the tax rate and assessed value: $T/A = t – E/A$. The statutory incidence within a jurisdiction will be moderately progressive in the sense that average tax payments decline with home value if tax rates and exemptions do not vary by assessed value and assessments are accurate measures of market value.

An extensive literature suggests that assessments tend to decline with sale price. Sirmans, Gatzlaff, and Macpherson (2008) present a review of the literature. Representative examples such as Cheng (1974), Clapp (1990), Cornia and Slade (2006), Ihlanfeldt (1982), McMillen (2013), McMillen and Weber (2008), Plummer (forthcoming), Ross (2010), Smith (2000), and Sunderman et al. (1990) all suggest that average assessed rates are lower for higher-priced homes. Regressive assessment practices can reverse the statutory progressivity of the property tax and produce a high degree of vertical inequity by lowering effective tax rates (i.e., tax payments divided by sales prices) for higher-priced properties.

Less attention has been paid in the literature to the horizontal equity of the property tax. Results in McMillen (2013) suggest that the variation of assessment rates within groups of similarly priced properties is at least as great as the variation across sales prices. Large variations in assessments for properties with similar sales prices can lead to significant variation in tax payments for otherwise identical households. Moreover, the variation in assessment rates in McMillen's (2013) sample of Chicago properties is far higher for low-priced homes than for high-priced properties. In addition to producing horizontal inequities, this high variation in assessment rates for low-priced properties can produce a degree of uncertainty regarding property tax payments that may be a deterrent to the redevelopment of low-income neighborhoods (McMillen and Weber, 2008).

Standard regression procedures are not well equipped to measure assessment *accuracy*. Though a regression can reveal that the expected assessment rate declines with sale price, it does not provide a direct measure of the degree of variability in assessment rates. The core issue in analyzing assessment rates is accuracy: if all properties are assessed accurately, a graph of assessment ratios against sales prices should reveal a tight cluster of points around a horizontal line, which implies an R^2 of zero for the regression. In contrast, quantile regressions are explicitly designed to explain the full distribution of the dependent variable at each value of an explanatory variable. Quantile regressions of assessment ratios on sales prices are capable of simultaneously measuring both horizontal and vertical inequities in assessment practices. Downward sloping quantile regressions imply vertical inequity, while a large difference between the predictions for low and high quantiles is an indication of horizontal inequity.

In this paper, I use quantile regression procedures to analyze the relationship between assessment rates and sales prices in Chicago. Property tax rates are set by multiple jurisdictions within the Chicago metro area. Data from the U.S. Census Bureau show that in 2009-10 the property tax accounted for 83% of local government tax revenue across all of Illinois.[1] Unlike the rest of Illinois, Chicago and the rest of Cook County have a classified tax system that produces significantly lower effective tax rates for residences than for commercial or industrial properties. A homestead exemption produces a moderately progressive statutory property tax system. However, average assessment ratios are much higher for low-priced properties than for high-priced homes in Chicago. The variation in assessment ratios is also much higher for low-priced homes.

I use quantile regressions to determine the extent to which the variation in assessment rates can be explained by structural characteristics of the properties and characteristics of their location. Whereas most studies using quantile regression approaches focus on differences in coefficient estimates across a set of target quantiles, I use kernel density functions to show how

[1] U.S. Census Bureau, State and Local Government Finances by Level of Government and by State: 2009-10, (http://www.census.gov/govs/estimate/).

the full distribution of assessment ratios changes with discrete changes in each of the explanatory variables. The distribution of assessment ratios is estimated to shift to the left with higher sales prices even after controlling for the number of nearby sales, characteristics of the structure, and characteristics of the location. The number of nearby sales has little effect on the distribution of assessment ratios. The assessment ratio distribution tends to shift to the right and become more variable as building area, lot sizes, and the number of bathrooms increase. Other structural characteristics have little effect on the assessment ratio distribution. The location of the home has a very large effect on assessments.

I use these results to compare the statutory and actual incidence of the property tax. Whereas the property tax structure in Cook County implies a modestly progressive structure in which effective tax rates should increase with sale price, the results imply that the frequency of extremely high assessment rates for low-priced homes leads to exactly the opposite outcome: effective tax rates tend to be much higher for low-priced homes. The variability of assessment rates also tends to be much higher for low-priced homes, and this variability translates directly into high variability of effective tax rates for low-priced properties.

2. Property Taxation in Cook County

The 1970 Illinois Constitution permitted counties with a population greater than 200,000 to adopt a classified property tax system. Whereas statutory tax rates must be identical for all properties, assessment rates can vary across property classes. Properties that have higher assessment rates will have higher effective tax rates, other things being equal.

Of the 10 eligible counties in the state (out of 102), only the largest, Cook County, has adopted classification. Prior to 2009, commercial properties in Cook County were supposed to be assessed at 38% of market value, while the statutory assessment rate for industrial properties was 36%. The statutory rate was 16% for the largest property class, Class 2, which comprises residential properties with six units or fewer. Cook County announced a "re-calibration" of the assessments rates in 2009. Since then, commercial and industrial properties are supposed to be assessed at 25% of market value, while Class 2 properties are supposed to be assessed at 10% of market value. To a first approximation, the classification system should produce effective tax rates that are 2.5 times as large for commercial and industrial properties as for Class 2 properties. The Illinois Constitution constrains the ratio of the highest to lowest statutory assessment rate to be no higher than 2.5.

The remaining 101 counties in the estate have statutory assessment rates of 1/3 for all types of property. In counties without classification, all properties are assessed at 1/3 of market value. The Illinois Department of Revenue conducts assessment ratio studies to determine whether an assessment district systematically over or under assesses properties. Most assessment districts have equalization factors near 1, which implies that properties tend to be assessed accurately at 1/3 of market value. In contrast, Cook County's equalization factor of 2.7076 in 2006 implies that the average assessment rate was far lower than 1/3 across all the

County's properties. A property's "equalized assessed value" (EAV) is the assessed value multiplied by the equalization factor.

Exemptions are the final major factor determining a property's tax base. Although exemptions are available for senior citizens, the disabled, and some other favored groups, the most important exemption is the homestead exemption, which applies to any homeowner's primary inhabited residence. In 2006, the homestead exemption was the difference between the current equalized assessed value and the 1977 EAV, up to a maximum of $5,000. This general homestead exemption was increased to $5,500 in 2008 and to $6,000 in 2009. Cook County also has special provisions that permit greater exemptions for residential properties with very high assessment growth rates.

Combining these provisions, the typical homeowner's property tax base is a simple linear function of market value. For example, the following table shows how the tax base would be calculated in 2006 for representative homes with market values of $100,000 and $500,000 that are assessed accurately at 16% of market value:

Market Value x Assessment Ratio (0.16)	$100,000	$500,000
= Assessed Value x Equalization Factor (2.7076)	$16,000	$80,000
= Equalized Assess Value (EAV) - Exemptions ($5,000)	$43,323	$216,608
= Adjusted EAV	$38,323	$211,608

The adjusted EAV forms the basis for the property tax base. The adjusted EAV is 38.3% of market value for the accurately assessed $100,000 home; it is 42.3% of market value for the higher-priced home. As tax rates are constrained to be identical for all properties within a taxing jurisdiction, the homestead exemption has produced a progressive tax structure in which higher-price homes pay higher effective tax rates. If the tax rate is 10% in this jurisdiction, the property tax payment for the \$100,000 home is \$3,823 and the effective tax rate is 3.82%. The tax payment for the higher-priced home is \$21,116 and the effective tax rate is 4.23%.

As the equalization factor and tax rate are identical for all properties, exemptions and assessment ratios are the only ways that effective tax rates vary across households with identical market values. Moreover, the relatively low value of the general homestead exemptions implies that most households qualified for the full homestead exemption of \$5,000 in 2006, and no more. The main source of potential departures from the statutory tax structure is the assessment ratio. Assessment ratios averaged less than 10% in 2006 with a standard deviation of 2.55%. Low assessment ratios result in low effective tax rates, while high assessment ratios do the opposite. The common empirical finding that assessment ratios decline with market value can potentially reverse the statutory progressivity of the property tax.

3. Quantile Regression Analysis of Assessment Ratios

Traditional studies of the relationship between assessment ratios and sales prices typically use one of the following two equations as the basis for a regression analysis:

$$E\left(\frac{A}{P}\right) = \beta_0 + \beta_1 P \tag{1}$$

$$Eln(A) = \lambda_0 + \lambda_1 \ln(P) \tag{2}$$

In the first equation, assessments (A) are proportional to sales prices (P) if $\beta_1 = 0$, and they are regressive in the sense that assessment ratios decline with sales prices if $\beta_1 < 0$. The second equation implies the following relationship between assessment ratios and sales prices:

$$\frac{A}{P} = e^{\lambda_0} P^{\lambda_1 - 1} \qquad (3)$$

Thus, the second equation implies that assessments do not vary by sale price if $\gamma_0 = 0$ and $\gamma_1 = 1$. In practice, studies using equation (2) as the basis for the estimating equation find that $\gamma_1 < 1$, which implies a regressive assessment system.

A potential problem with using equation (1) as the basis for the estimating equation is that having *P* on both the left and right hand side of the equations can lead to a downward bias in $\hat{\beta}_1$. Equation (2) is a particularly useful specification for data sets in which sales precede assessment dates. In my Chicago data set, assessments date from January of 2003, 2006, and 2009. Pairing each year's assessments from sales from the preceding two years assures that *P* can be taken as exogenous in equation (2).

I use the expectations operator in equations (1) and (2) to emphasize an important characteristic of standard regression procedures: a regression is designed to estimate the conditional expectation of the dependent variable given a value for the explanatory variable. For each sale price, a regression equation implies a single estimated assessment ratio. In practice, assessment ratios can vary dramatically for homes that have identical sales prices. In McMillen (2013), I show that assessments in Chicago are extremely variable, particularly at low sales prices.

Conditional expectations do not imply anything directly about variability. Yet variability is a prominent feature of assessments. In McMillen (forthcoming), I show that the most salient feature of assessment practices in Chicago is the extraordinary variation in assessments at low sales prices. If the objective of an assessment system is to have assessments that closely match the target rate, then the goal is to have assessment ratios that exhibit little variation within a sale price category as well as the absence of systematic relationship between sales prices and expected assessment ratios. Assessment variability within groups of properties with similar sales prices violates the goal of horizontal equity. By focusing on conditional expectations, traditional regression procedures are better equipped to analyze vertical equity.

Quantile regressions are much better suited to analyzing assessment uniformity than standard regression procedures. Parametric quantile procedures use a linear equation to approximate the conditional quantile function of the dependent variable for a given set of explanatory variables. Supplementing (2) with a set of additional explanatory variables, X, the conditional quantile function for the log of assessed value is:

$$Q_{lnA}\,(\tau|lnP, X) = \theta_{0,\tau} + \theta_{1,\tau}\ln(P) + X\theta_{2,\tau} \tag{4}$$

where the quantile, τ, ranges from 0 to 1. For example, when τ = .50, the estimated quantile function is the linear relationship that fits the data best when half of the actual values of the dependent variable lie above the line and half lie below it, i.e., the conditional median. When τ = .10, the conditional quantile function is the best-fitting linear relationship such that 90% of the actual values of the dependent variable lie above the line, and the remaining 10% are below it. The quantile estimator was proposed by Koenker and Bassett (1978). Koenker (2005) presents a full theoretical treatment of the estimator. Buchinsky (1998) and Koenker and Hallock (2001) present good overviews.

Most empirical researchers present quantile estimates for a set of target quantiles such as τ = 0.10, 0.25, 0.50, 0.75, and 0.90. Although this procedure produces tables that look quite similar to typical tables of regression results, it does not take advantage of the real power of the quantile approach, which is to show how the full distribution of the dependent variable shifts when the value of an explanatory variable changes. In the empirical section of the paper, I estimate equation (4) for quantiles ranging from 0.02 to 0.98 in increments of 0.01. The results of these 97 quantile regression estimates show how the full distribution of $\ln(A)$ varies with $\ln(P)$ and X. The estimated value of $\ln(A_i)$ for quantile τ is simply $\hat{\theta}_{0,\tau} + \hat{\theta}_{1,\tau}\ln(P_i) + X_i\hat{\theta}_{2,\tau}$. Thus, the full set of 97 quantile estimates implies a set of $n \times 97$ predicted values of $\ln(A)$, where n is the number of observations. A kernel density function estimate for this set of predictions will look quite similar to the kernel density function for the actual values of $\ln(A)$.

The estimated quantile regressions can easily be used to show how the distribution of $\ln(A)$ changes as the value of an explanatory variable changes. For example, suppose we want to see how the distribution of $\ln(A)$ changes as $\ln(P)$ rises from $\ln(P_1)$ to $\ln(P_2)$. At each value of τ, the estimated quantile regression implies:

$$\hat{Q}_{lnA}\,(\tau|lnP = ln(P_1), X) = \hat{\theta}_{0,\tau} + \hat{\theta}_{1,\tau}\ln(P_1) + X\hat{\theta}_{2,\tau} \qquad (5)$$

$$\hat{Q}_{lnA}\,(\tau|lnP = ln(P_2), X) = \hat{\theta}_{0,\tau} + \hat{\theta}_{1,\tau}\ln(P_2) + X\hat{\theta}_{2,\tau} \qquad (6)$$

Evaluating these equations at actual values of X but constant values of $\ln(P_1)$ and $\ln(P_2)$ produces a full set of $n \times 97$ predicted values for $\ln(A)$ because the values of X vary across observations and because the estimated coefficients vary across quantiles. Kernel density estimates can then be applied to the two sets of predicted values to see how both the location and the shape of the distribution of $\ln(A)$ predictions change as $\ln(P)$ changes. In contrast, the distribution of predictions from ordinary regressions simply traces out the distribution of values of X because the coefficients are constant.[2]

Trivial algebra converts the estimated values of $\ln(A)$ into estimated assessment ratios:

$$\hat{Q}_{A/P}\,(\tau|lnP = ln(P_1), X) = exp\left(\hat{\theta}_{0,\tau} + \left(\hat{\theta}_{1,\tau} - 1\right)\ln(P_1) + X\hat{\theta}_{2,\tau}\right) \qquad (7)$$

$$\hat{Q}_{A/P}\,(\tau|lnP = ln(P_2), X) = exp\left(\hat{\theta}_{0,\tau} + \left(\hat{\theta}_{1,\tau} - 1\right)\ln(P_2) + X\hat{\theta}_{2,\tau}\right) \qquad (8)$$

[2] McMillen (2012) discusses this approach in some detail using both actual and simulated data.

A significant advantage of quantile regression is that quantiles of a distribution are not affected by monotonic transformations.

As discussed in McMillen (2012, 2013), these estimates can easily be adapted to account for broad geographic trends in the estimated coefficients. Following the notation in Koenker (2005), the standard quantile regression approach involves finding the values for $\hat{\beta}(\tau)$ that minimize $\sum_i \rho_\tau(y_i - x_i'\beta)$, where $\rho_\tau(u)$ is the piecewise linear function $\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$. Chaudhuri (1991), Koenker (2005), and Yu and Jones (1998) derive a fully nonparametric version of quantile regression by adding a kernel weight function to this expression. For the case of a single explanatory variable, *x*, the locally weighted quantile regression problem is:

$$\min_\beta w_i(x)\, \rho_\tau\big(y_i - \beta_0 - \beta_1(x_i - x)\big) \tag{9}$$

where $w_i(x) = K\big((x_i - x)/h\big)/h$ is the kernel weight function, *x* is the target for the locally weighted quantile regression, and *h* is the bandwidth. This approach could be applied directly to the current study by specifying $y_i = ln(A_i)$ and $x_i = ln(P_i)$. The set of explanatory variables can also be expanded, although a full nonparametric approach suffers from a "curse of dimensionality" that limits it to applications with a small number of explanatory variables.

Researchers in urban economics and geography tend to prefer locally weighted versions of nonparametric estimators that allow the coefficients of a base linear model to vary smoothly over space. This version of the nonparametric estimator for standard regression models is sometimes referred to as "geographically weighted regression," although it is simply a special case of an approach that is referred to as "conditionally parametric regression" by statisticians. Equation (9) is made conditionally parametric by having different variables in the kernel weight function and the base regression:

$$\min_{\beta} w_i(z)\, \rho_\tau\big(y_i - \beta_0 - \beta_1(x_i - x)\big) \tag{10}$$

In the empirical application, z represents straight-line geographic distance between the location of each property and a set of target locations. Following an approach discussed in Loader (1999), the estimates are then interpolated to every location in the data set.[3]

I use the conditionally parametric approach to estimate a set of locally weighted quantile regressions using $\ln(A)$ as the dependent variable. The explanatory variables are $\ln(P)$ and sets of variables representing structural characteristics of the home and various controls for the home's location. The estimates imply a separate set of coefficients for each property and for each quantile. Despite this apparent complexity, the estimates are again easily summarized using kernel density functions. The counterparts to equations (7) and (8) are simply:

$$\hat{Q}_{A/P}\,(\tau | lnP = ln(P_1), X_i) = exp\left(\hat{\theta}_{0,\tau}(z_i) + \left(\hat{\theta}_{1,\tau}(z_i) - 1\right)\ln(P_1) + X_i\hat{\theta}_{2,\tau}(z_i)\right) \tag{11}$$

$$\hat{Q}_{A/P}\,(\tau | lnP = ln(P_2), X_i) = exp\left(\hat{\theta}_{0,\tau}(z_i) + \left(\hat{\theta}_{1,\tau}(z_i) - 1\right)\ln(P_2) + X_i\hat{\theta}_{2,\tau}(z_i)\right) \tag{12}$$

All that differs between equation (11)-(12) and equation (7)-(8) is that the coefficients vary by observation as well as by quantile. There again are $n \times 97$ estimates for each target value of P. Kernel density estimates summarize how the distribution of assessment ratios changes as the value of an explanatory variable changes.

The nonparametric approach is important for this application because assessment practices are not uniform across the entire sample area. Some areas of the city tend to be assessed at higher rates than other areas. Assessment rates are also more variable in some areas than others. The locally weighted approach accounts for geographic variation in assessment levels and variability that varies smoothly over space.

---

[3] In the empirical section of the paper, I use a tri-cube kernel with a window size of 25%, which implies that the nearest 30% of the observations are used to estimate locally weighted quantile regression for a target point. I base weights on the straight-line distance between each observation and use a tri-cube kernel weight function for $w_i(z)$.

4. Data

The Cook County Office of the Assessor provided data on assessed values for all counties in the county. Sales prices were provided by the Illinois Department of Revenue (DOR). The DOR data set also includes variables that help assure that the sales represent impersonal, arm's length transactions. For example, I exclude sales between relatives, court-ordered sales and foreclosures, sales with unusual deeds, and sales that include transfers of personal property. I restrict the analysis to Class 2 properties, i.e., residential properties with six units or fewer.

The sale price and assessment data form the raw the data for a standard assessment study. Assessments are conducted on a rotating three-year basis in Cook County. For example, the earlier year in the sample, 2003, is a year when homes in Chicago were assessed. Properties in the northern suburbs were then assessed in 2004, with south suburban assessments following in 2005. I focus on Chicago to help make the tables and figures more manageable. The three years covered in the resulting sample – 2003, 2006, and 2009 – are interesting because home prices had only recently begun to rise sharply in Chicago in 2003 while 2006 was the height of the housing boom. Prices and sales had dropped dramatically by 2009.

For a standard assessment ratio study, assessments for each year would be matched to recent sales to compare the assessed value to the best measure of market value. For example, sales from 2008-2010 might be used to analyze assessments for 2009. A price index is then used to match the timing of the sales to the statutory assessment date of January 1, 2009. This approach is appropriate for a ratio study because the primary objective is descriptive: what is the mean or median of the assessment ratio distribution in a given assessment year, and how variable are the assessments? As the objective here is to determine whether assessment ratio variations can be explained by characteristics of the home and location, I instead restrict the sample to assessments that can be paired with sales from the two years preceding the statutory assessment date. Thus, the 2003 sample includes properties that sold in 2001-2002, the 2006 sample includes properties that sold in 2004-2005, and the 2009 sample includes properties that sold in 2007-2008. I use the Case-Shiller price index for the Chicago metropolitan area to adjust sales prices to the statutory assessment date. The advantage of this approach for the objective of this study is these sales prices are the ones that the Assessor's Office likely could observe when forming the assessments, and there is no simultaneity problem because the future assessment would not be known at the time of the sale.

Additional sample restrictions are imposed by data availability. Structural characteristics for the properties in the sample are available from the Assessor's Office for a single year, 2003. The data are not available for condominiums. I also restrict the sample to properties that were successfully geocoded and to sales that included only a single parcel identification number, which excludes multi-property sales. Finally, I use a nonparametric trimming procedure to trim outlier ratios from each year's sample, and I discard the lowest 1% and highest 1% of the sales prices. These restrictions ensure that the results are not unduly influenced by a small number of unusual assessments and sales prices. The final sample includes 15,233 sales in 2003, 16,637 in 2006, and 5,497 in 2009.

Descriptive statistics are presented in Table 1. Average sales prices rose from $202,289 in 2003 to $261,359 in 2006, and then fell to $221,127 in 2009. Although the statutory assessment rate was 16% in 2003 and 2006, the mean assessment ratio is significantly under 10% in both years. The median assessment ratio rose to nearly 12% in 2009, the year in which the recalibration of the official assessment rate to 10% was enacted. A more likely cause for this increase was the recession: assessments did not decline as fast as sales prices during the housing crisis and the relatively few homes that actually sold often did so at extremely low prices, leading to high assessment ratios.

Structural characteristics include standard hedonic price function variables such as building area and lot size; the number of rooms, bedrooms, and bathrooms; and indicators that the home has a basement, attic, central air-conditioning, a fireplace, brick construction, a 1-car garage, a garage with space for two or more cars, and the age of the structure. Characteristics of the location include the proportion Black and Hispanic in the census tract, the percentage of housing units that are vacant, and the log of the median income in the census tract. All census data are drawn from the 2000 U.S. Census. Various distance measures were added using a GIS program: distance to the census business distract (the "CBD", taken to be traditional city center at the intersection of State and Madison Streets), to Lake Michigan, the nearest stop on an elevated train line (the "EL"), distance from the EL line itself, and distance from a fixed rail line.

In addition to sale price, characteristics of the home, and location, it seems likely that the number of nearby sales may help to explain assessment accuracy (McMillen and Weber, 2008). I use a kernel density function to estimate the density of sales in the vicinity of each observation. I base the sales density measure on sales from the 3-year period preceding the assessment date. Thus, the sales density for the 2003 sample measures the number of sales that took place in the vicinity from 2000-2003. I use a bivariate kernel in the geographic coordinates of each sale using a fixed bandwidth based on Silverman's (1986) rule of thumb. The kernel density function provides a simple, continuous measure of the spatial variation in the number of sales across the city.

5. Assessment Ratios and Sales Prices

Figure 1 presents scatter plots of assessment ratios by sales prices for 2003, 2006, and 2009. The patterns are remarkably consistent over time. Assessment ratios appear to decline markedly with sale price at low price levels. In a broad range – from perhaps $200,000 on – assessment ratios appear to hover more or less randomly around a horizontal line. The variance of the distribution also appears to be much higher at low sales prices. The primary difference between the results over time is that the upper end of assessment ratios is much higher in 2009 than in previous years. Although the overall pattern appears similar in 2009 to earlier years and the horizontal portion of the plot is not much different, assessment rates approach 40% at very low sales prices. These extremely high assessment rates indicate that assessments did not fully reflect the sharp drop in prices associated with many sales during the housing crisis.

Table 2 presents the simplest specification of the relationship between the log of assessed value and the log of sale price. Table 2 shows the results for the 10%, 50%, and 90% quantiles, along with a test for differences in coefficients across the 10% and 90% quantiles. As would be expected from the assessment ratio plots in Figure 1, the coefficients on log sale price are significantly less than 1 in every model, which implies that assessment ratios decline with sale price at all quantiles.

Figures 3 and 4 present the implied distributions of assessment ratios for various target prices – $50, $100, $150, $200, $400, $600, $800, and $1000 thousand. The graphs make it clear that the distribution of assessment ratios both shifts to the left and has a much lower variance as sale price increases. Although the left side of the distribution does shift toward the left, the most salient change is the large decline on the right. In other words, extremely high assessment ratios become much less likely as sale price increases. These figures demonstrate clearly the advantages of the quantile approach. Standard regression approaches imply that all assessment ratios decline in expectation as sale price increases: the distribution simply makes a parallel lift to the left. The quantile approach reveals that this shift to the left is much more pronounced at high sales prices. Assessment rates become more uniform as sale price increases, implying a higher level of horizontal equity.

Table 3 shows the results for 2006 when the full set of structural characteristics, location measures, and the sales density variable are added to the quantile regressions. Although 2006 is near the peak of the housing market, I have chosen to focus on it because the results look quite similar to those for other years (apart from the extremely high assessment ratios in 2009). For convenience, the cells are shaded when a coefficient is statistically significant at the 5% level. Every variable save one – the presence of an attic – is statistically significant in at least one equation, and most are significant at the 10%, 50%, and 90% quantiles. The fact that the coefficients differ across the 10% and 90% quantiles suggests that the distribution of assessment ratios does not always shift by the same amount at all points in the distribution when a variable changes in value.[4]

Although Table 3 includes several location variables, it is unlikely that the equations adequately control for all spatial effects in a city as large as Chicago. One strategy for controlling for spatial effects might be to include fixed effects for community areas or census tracts. The fixed effect approach has several disadvantages for this application. First, the large number of coefficients makes quantile estimation extremely slow for relatively large data sets with many districts. Moreover, districts with few observations produce imprecise results across a large number of quantiles. Most importantly, there is no reason to expect the spatial effects

---

[4] I use a bootstrap resampling procedure to construct standard errors for the difference in coefficients across the 10% and 90% quantiles.

being approximated by a set of discrete variables to make discrete changes at district boundaries. It is more reasonable to think that the variation in assessment ratios is continuous, and continuous variation is readily modeled using a locally weighted version of the quantile regression model.

I estimate the models using a tri-cube kernel in distance and a 30% window size. Following Loader (1999), I use an adaptive decision tree approach to choose the set of target points for estimation. I then use a bivariate interpolation procedure to calculate separate coefficient estimates for every location. As the nonparametric estimation procedure controls for smooth variation in assessment ratios over space, I drop the distance variables from the locally weighted versions of the model. With 16,637 observations and 97 target quantiles, the results clearly cannot be summarized simply in a set of tables. Nonetheless, the effect of each explanatory variable on the distribution of assessment ratios is readily displayed using kernel density estimates of the predictions from equations (11) and (12).

The results are shown in the series of panels in Figure 5. The first two panels show again that the distribution of assessment ratios shifts to the right and becomes less variable as sale price increases. The effect is even more pronounced after controlling for additional variables. Part of the reason for this result is evident in the results for building area and land area. As building area increases from 1,000 to 2,500 square feet, the distribution of assessment ratios shifts to the right and becomes less variable. A similar rightward shift is evident as land areas increases from 3,000 to 12,000 square foot lots. Thus, controlling for sale price, larger houses with bigger lots have higher and more uniform assessments. Controlling for square footage and lot size, higher priced homes have lower and more uniform assessments. There also is a tendency for homes with more bathrooms to have higher and more uniform assessment ratios after controlling for the effects of other variables. Other structural variables have very little effect on the distribution of assessment ratios.

Although the sales density variable is statistically significant, the density graphs show that the variable has very little real effect on the assessment ratio distribution.[5] In contrast, characteristics of the census tract have large effects on the assessment ratio distribution. After controlling for other variables, a higher percentage of African-Americans or Hispanics in a census district is associated with lower, more uniform assessments. A greater proportion of vacant units leads to a higher probability of extremely high assessment ratios. Controlling for the sales prices of the individual homes and other variables, a higher median income in the census tract is associated with higher, less uniform assessments.

6. Statutory versus Actual Property Tax Incidence

Despite its apparent complexity, the tax structure in Cook County actually simplifies to a linear schedule: $T = t(aPm - e)$, where T is the tax payment, t is the tax rate, a is the assessment ratio, P is sale price, m is the equalization factor (or "multiplier"), and e is the homestead exemption. Sale price is directly observed for our sample of sales. The general homestead exemption for Cook County was $4,500 in 2003, $5,000 in 2006, and $6,000 in 2009. The equalization factor was 2.4598, 2.7076, and 3.3701 for these three years. Although the statutory assessment ratio was 16% in 2003 and 2006, actual rates were much closer to the "recalibration" rate of 10%. Thus, I use a = 10% to evaluate the statutory tax payment for all three years. Tax rates can vary significantly across Chicago because even individual municipalities are covered by multiple tax districts. For the purposes of this exercise, I set the

[5] The values for the sales density variable are the 10%, 25%, 50%, 75%, and 90% percentiles.

tax rate to a single arbitrary value of 10% for all properties. The result is a modestly progressive tax structure: the effective tax rate is $\frac{T}{P} = tam - \frac{te}{P}$.

The quantile regression results imply large variations in actual assessment rates. At quantile τ, the predicted assessment rate for property i is $\hat{a}_{i,\tau}$. Thus, the tax payment associated with quantile τ is $T_{i,\tau} = t(\hat{a}_{i,\tau}P_i m - e)$, and the effective tax rate is $\frac{T_{i,\tau}}{P_i} = t\hat{a}_{i,\tau}m - \frac{te}{P_i}$. As the $\hat{a}_{i,\tau}$ take on many possible values for each property, the estimates imply a distribution of possible values for tax payments and effective tax rates for each property in the sample. I again use kernel density estimates to summary the full set of results.

Figure 6 shows the implied distribution of tax payments, *T*, for each year. The statutory distribution simply reflects the distribution of sales prices. The quantile predictions reflect both the distribution of sales prices and the distribution of the quantile estimates, $\hat{a}_{i,\tau}$. The right panel shows the difference in the kernel density estimates: negative values imply that the density of tax payments implied by the quantile estimates is lower than the statutory density. For all years, inaccuracy in the assessment process reduces the frequency of high and low payments, while increasing the mass in the center of the distribution. Thus, the assessment process leads to less variation in tax bills than is implied by the statutes.

The difference between actual and statutory rates is more dramatic in Figure 7, which presents kernel density estimates for the implied effective tax rates. By statute, there should be little variation in effective tax rates. The assessment process produces a large number of vary high effective rates for low-priced homes, while lowering the rates for many high-priced properties. The result is a much flatter distribution of effective tax rates than is implied by statute, with both very low and very high rates being much more common than they would be if properties were assessed accurately. Extremely high rates are particularly common in 2009, when assessments failed to keep up with the rapid declines in the housing market.[6]

Figure 8 shows that the tendency to over-assess low-priced homes is sufficient to reverse the statutory progressivity of the property tax system in Chicago. The statutory effective tax rate is a simple function of sale price: it is low at very low prices, and then rises sharply before leveling out as the homestead exemption becomes small relative to the sale price. In contrast, the quantile estimates imply a full distribution of possible tax rates at every sale price. Figure 8 presents smoothed versions of the 10%, 50%, and 90% quantiles for effective rates. The salient feature of Figure 8 is the extremely high tax rates for very low-priced homes. Effective tax rates are highly variable at low sales prices, but are clearly much higher than they are supposed to be by statute. Effective rates decline rapidly with sale price and the difference between the 10% and 90% quantile narrows. For relatively high sales prices – above about $200,000 – the distribution of tax rates is less variable, less responsive to increases in sales prices, and somewhat lower than implied by statute.

7. Conclusion

By statute, effective tax rates should be much lower for higher-priced properties in jurisdictions with homestead exemptions. In practice, assessment rates tend to be higher for low-priced properties. The empirical results in this paper suggest that the tendency to over-assess low-priced homes leads to a complete reversal of the statutory progressivity of the property tax system in Chicago. The highest effective tax rates tend to be paid by the owners of the lowest-priced properties. While this result violates principles of vertical equity, variations in assessment

[6] It should be borne in mind that the tax rate is actually endogenous in Cook County. If all assessments are high, the tax rate will fall unless additional revenue is authorized. The distributional effects are clear, however: the effective tax rate will be higher for a home that is over-assessed by a higher amount.

rates also lead to significant degrees of horizontal inequity. The variation in effective tax rates is substantial at any sale price, and it is particularly significant for low-priced properties. Very similar patterns are evident during and after the housing boom. Although assessments did not fall as rapidly as prices between 2006 and 2009, the shape of the assessment ratio distribution and the relationship between assessments rates and sales prices is nearly the same in 2009 as before, the main difference being that average assessment rates were higher and extremely high rates became more common in 2009 than before.

Empirical studies of assessment practices typically use a regression approach to estimate the relationship between assessment ratios and sales prices. Regressions do not account for the critical feature of assessment – accuracy. Assessments in Chicago are much more variable at low sales prices than at high prices. The result is that owners of low-priced homes may face uncertainty over future assessments as well as high tax rates due to their current assessment. A series of kernel density functions illustrate how the full distribution of assessment ratios responds to discrete changes in explanatory variables related to characteristics of the housing structure and location. Although most variables are statistically significant, only sale price, building area, lot size, the number of bathrooms, and characteristics of the location have meaningful effects on the distribution of assessment ratios. The number of nearby sales has little effect on assessment accuracy. Overall, the results suggest that sale price and location have the largest influence on the distribution of assessment rates, which implies in turn that it is these variables that explain variations in effective tax rates across properties.

References

Anderson, Nathan B., "Property Tax Limitations: An Interpretative Review," *National Tax Journal* 59 (2006), 685-694.

Anderson, Nathan B., "Market Value Assessment and Idiosyncratic Tax-Price Risk: Understanding the Consequences of Alternative Definitions of the Property Tax Base," *Regional Science and Urban Economics* 42 (2012), 545-560.

Barnett, Jeffrey L. and Phillip M. Vidal, "State and Local Government Finances Summary: 2010," U.S. Census Bureau Governments Division Briefs (2012), http://www2.census.gov/govs/estimate/summary_report.pdf.

Buchinsky, Moshe, "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *Journal of Human Resources* 33 (1998), 1-30.

Cabral, Maria and Caroline Hoxby, "The Hated Property Tax: Salience, Tax Rates, and Tax Revolts," working paper, University of Texas (2013).

Chaudhuri, Probal, "Nonparametric Estimates of Regression Quantiles and their Local Bahadur Representation," *Annals of Statistics* 19 (1991), 760-777.

Cheng, Pao, "Property Taxation, Assessment Performance and Its Measurement," *Public Finance* 29 (1974), 268-284.

Civic Federation, "The Cook County Property Assessment Process: A Primer on Assessment, Classification and Property Tax Exemptions," 2010.

Clapp, John M., "A New Test for Equitable Real Estate Tax Assessment," *Journal of Real Estate Finance and Economics* 3 (1990), 233-49.

Cornia, Gary C. and Barrett A. Slade, "Horizontal Inequity in the Property Taxation of Apartment, Industrial, Office, and Retail Properties," *National Tax Journal* 59 (2006), 33-55.

Ihlanfeldt, Keith R., "Property Tax Incidence on Owner-Occupied Housing: Evidence from the Annual Housing Survey," *National Tax Journal* 35 (1982), 89-97.

Koenker, Roger, *Quantile Regression*, Cambridge University Press, New York (2005).

Koenker, Roger and Gilbert W. Bassett, Jr., "Regression Quantiles," *Econometrica* 46 (1978), 33-50.

Koenker, Roger and Kevin F. Hallock, "Quantile Regression," *Journal of Economic Perspectives* 15 (2001), 143-156.

Loader, Clive, *Local Regression and Likelihood*, Springer, New York (1999).

McMillen, Daniel P., *Quantile Regression for Spatial Data*, Springer, New York (2012).

McMillen, Daniel P., "Conditionally Parametric Quantile Regression for Spatial Data: An Analysis of Land Values in Early Nineteenth Century Chicago," manuscript, University of Illinois, (2013).

McMillen, Daniel P, "The Effect of Appeals on Assessment Ratio Distributions: Some Nonparametric Approaches" *Real Estate Economics* 41 (2013), 165-191.

McMillen, Daniel P. and Rachel Weber, "Thin Markets and Property Tax Inequities: A Multinomial Logit Approach," *National Tax Journal* 61 (2008), 653–671.

Plummer, Elizabeth, "Evidence on the Incidence of Residential Property Taxes across Households," *National Tax Journal* 56 (2003), 739-753.

Plummer, Elizabeth, "The Effects of Property Tax Protests on the Assessment Uniformity of Residential Properties," *Real Estate Economics*, forthcoming.

Ross, Justin M, " Interjurisdictional Determinants of Property Assessment Regressivity," *Land Economics* 88 (2012), 28-42.

Silverman, A. W., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York (1986).

Sirmans, Stacy G., Dean H. Gatzlaff, and David A. Macpherson, "Horizontal and Vertical Inequity in Real Property Taxation," *Journal of Real Estate Literature* 16 (2008), 167-180.

Smith, Brent C., "Applying Models for Vertical Inequity in the Property Tax to a Non-Market Value State," *Journal of Real Estate Research* 19 (2000), 320-244.

Sunderman, Mark A., John W. Birch, Roger E. Cannaday, and Thomas W. Hamilton, "Testing for Vertical Inequity in Property Tax Systems," *Journal of Real Estate Research* 5 (1990), 319-34.

Yu, Keming and M. C. Jones, "Local Linear Quantile Regression," *Journal of the American Statistical Association* 93 (1998), 228-237.

Table 1

Descriptive Statistics

Variable	2003		2006		2009	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Assessment	18,681	10,887	24,615	13,064	25,672	14,458
Sale Price	202,289	122,171	261,359	143,601	221,127	157,715
Log of Building Area	7.1869	0.3922	7.1956	0.4098	7.2174	0.4187
Log of Land Area	8.2141	0.3322	8.2159	0.3127	8.2341	0.3254
Number of Rooms	6.2561	2.5863	6.3813	2.8420	6.5749	3.4419
Number of Bedrooms	3.2107	1.2389	3.2761	1.3152	3.3402	1.5539
Number of Bathrooms	1.4991	0.6637	1.5131	0.6681	1.5489	0.7739
Basement	0.7534	0.4311	0.7457	0.4355	0.7708	0.4204
Attic	0.4021	0.4903	0.3987	0.4897	0.3924	0.4883
Central Air Conditioning	0.1786	0.3830	0.1431	0.3501	0.1836	0.3872
Fireplace	0.0836	0.2768	0.0732	0.2605	0.0888	0.2844
Brick	0.6113	0.4875	0.5814	0.4934	0.6451	0.4785
1 Car Garage	0.2914	0.4544	0.2930	0.4551	0.2934	0.4554
2+ Car Garage	0.4620	0.4986	0.4433	0.4968	0.4750	0.4994
Age	7.0914	2.7364	7.7481	2.6733	7.6729	2.6801
Age^2/10	5.7775	3.7802	6.7179	3.9769	6.6056	4.0794
Census Tract % Black	37.3703	42.9954	48.2579	44.9161	33.1985	42.0136
Census Tract % Hispanic	24.7890	27.1865	21.4509	26.9014	23.8922	26.0809
Census Tract % Vacant	6.3160	4.8138	7.1504	5.0972	5.8672	4.4660
Census Tract Log Median Income	10.5974	0.3339	10.5350	0.3514	10.6310	0.3213
Distance from CBD	8.5503	2.8091	8.6436	2.8496	8.5322	2.8421
Distance from Lake Michigan	5.1863	2.3812	4.9660	2.3136	5.1246	2.4695
Distance from EL Stop	1.4576	1.0844	1.4379	1.0914	1.4570	1.1084
Distance from EL Line	1.7239	1.3153	1.6428	1.2825	1.6806	1.3381
Distance from Rail Line	0.4487	0.2999	0.4181	0.2762	0.4636	0.3042
Sales Density	1.1735	0.6268	1.3165	0.7836	1.1400	0.7868
Median Assessment Ratio (%)	9.2985		9.4757		11.9271	
Number of Observations	15,233		16,637		5,497	

Table 2

Quantile Regressions for Log Assessed Value

Variable	10%	50%	90%	90% - 10%
2003				
Constant	-1.008 (0.080)	-0.314 (0.043)	0.499 (0.050)	1.827 (0.135)
Log Sale Price	0.866 (0.007)	0.831 (0.004)	0.775 (0.004)	-0.107 (0.011)
Pseudo-R^2	0.452	0.567	0.579	
2006				
Constant	-0.419 (0.072)	-0.226 (0.048)	0.555 (0.054)	1.378 (0.151)
Log Sale Price	0.823 (0.006)	0.829 (0.004)	0.777 (0.004)	-0.068 (0.012)
Pseudo-R^2	0.453	0.553	0.564	
2009				
Constant	3.777 (0.170)	3.915 (0.086)	4.172 (0.087)	0.515 (0.347)
Log Sale Price	0.479 (0.014)	0.509 (0.007)	0.504 (0.007)	0.030 (0.029)
Pseudo-R^2	0.201	0.331	0.371	

Table 3

Log Assessed Value Quantile Regressions, 2006

Variable	10%	50%	90%	90% - 10%
Constant	1.7856 (0.1398)	2.8508 (0.0829)	3.5459 (0.1386)	1.7603 (0.2282)
Log of Sale Price	0.3484 (0.0071)	0.2572 (0.0042)	0.2311 (0.0070)	-0.1173 (0.0111)
Log of Building Area	0.2742 (0.0097)	0.2785 (0.0058)	0.2760 (0.0096)	0.0018 (0.0080)
Log of Land Area	0.0563 (0.0080)	0.1095 (0.0047)	0.1508 (0.0079)	0.0945 (0.0115)
Number of Rooms	0.0036 (0.0015)	0.0021 (0.0009)	0.0015 (0.0015)	-0.0022 (0.0022)
Number of Bedrooms	-0.0114 (0.0034)	-0.0068 (0.0020)	-0.0100 (0.0034)	0.0014 (0.0038)
Number of Bathrooms	0.0244 (0.0051)	0.0371 (0.0030)	0.0420 (0.0051)	0.0177 (0.0078)
Basement	0.0699 (0.0059)	0.0674 (0.0035)	0.0557 (0.0058)	-0.0143 (0.0073)
Attic	0.0022 (0.0046)	-0.0050 (0.0027)	-0.0008 (0.0045)	-0.0030 (0.0065)
Central Air Conditioning	0.0077 (0.0065)	0.0134 (0.0039)	0.0030 (0.0065)	-0.0047 (0.0083)
Fireplace	0.0372 (0.0081)	0.0443 (0.0048)	0.0941 (0.0081)	0.0569 (0.0124)
Brick	0.0001 (0.0051)	-0.0064 (0.0030)	-0.0032 (0.0050)	-0.0033 (0.0057)
1 Car Garage	0.0309 (0.0054)	0.0254 (0.0032)	0.0103 (0.0053)	-0.0206 (0.0082)
2+ Car Garage	0.0430 (0.0051)	0.0351 (0.0030)	0.0166 (0.0051)	-0.0263 (0.0064)
Age	0.0346 (0.0045)	0.0108 (0.0027)	-0.0227 (0.0045)	-0.0573 (0.0066)
Age^2/10	-0.0325 (0.0031)	-0.0134 (0.0018)	0.0123 (0.0031)	0.0448 (0.0043)
Census Tract % Black	-0.0041 (0.0001)	-0.0044 (0.0001)	-0.0044 (0.0001)	-0.0003 (0.0002)
Census Tract % Hispanic	-0.0019 (0.0001)	-0.0021 (0.0001)	-0.0026 (0.0001)	-0.0007 (0.0002)
Census Tract % Vacant	-0.0022 (0.0007)	-0.0016 (0.0004)	-0.0025 (0.0006)	-0.0003 (0.0011)

Census Tract Log Median Income	0.1718 (0.0094)	0.1718 (0.0055)	0.1439 (0.0093)	-0.0279 (0.0138)
Distance from CBD	-0.0370 (0.0014)	-0.0517 (0.0008)	-0.0602 (0.0014)	-0.0232 (0.0018)
Distance from Lake Michigan	0.0177 (0.0012)	0.0140 (0.0007)	0.0059 (0.0012)	-0.0119 (0.0016)
Distance from EL Stop	0.0078 (0.0039)	0.0288 (0.0023)	0.0384 (0.0039)	0.0306 (0.0039)
Distance from EL Line	-0.0603 (0.0034)	-0.0752 (0.0020)	-0.0765 (0.0033)	-0.0162 (0.0036)
Distance from Rail Line	0.0225 (0.0083)	0.0027 (0.0049)	-0.0052 (0.0083)	-0.0278 (0.0094)
Sales Density	-0.0677 (0.0040)	-0.0898 (0.0024)	-0.1012 (0.0040)	-0.0335 (0.0045)
Pseudo-R^2	0.6348	0.7242	0.7429	

Notes. Standard errors are in parentheses. The number of observation is 16,637.

Figure 1

Data Plots

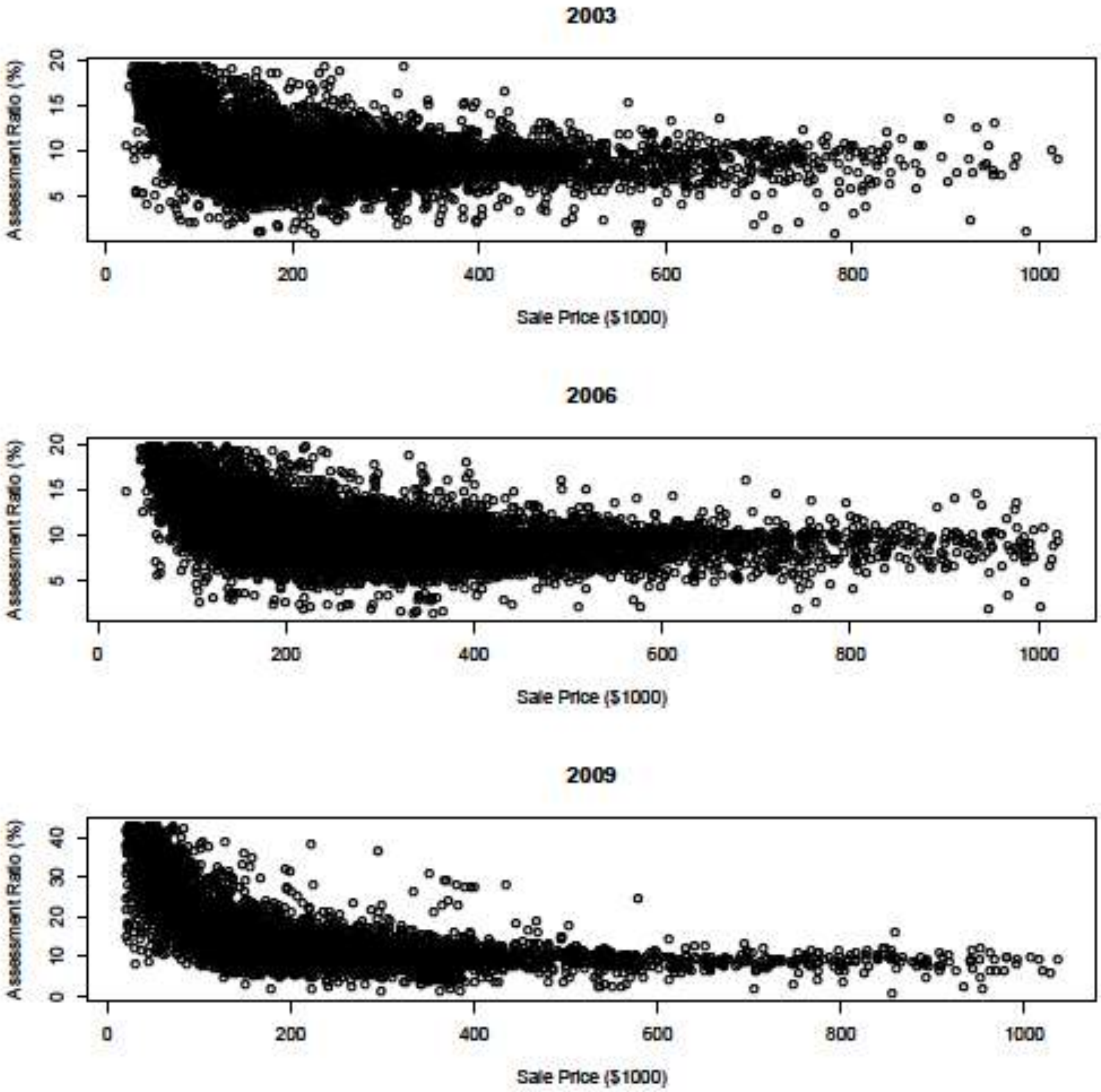


Figure 2

Estimated Assessment Ratio Densities by Sale Price, 2003

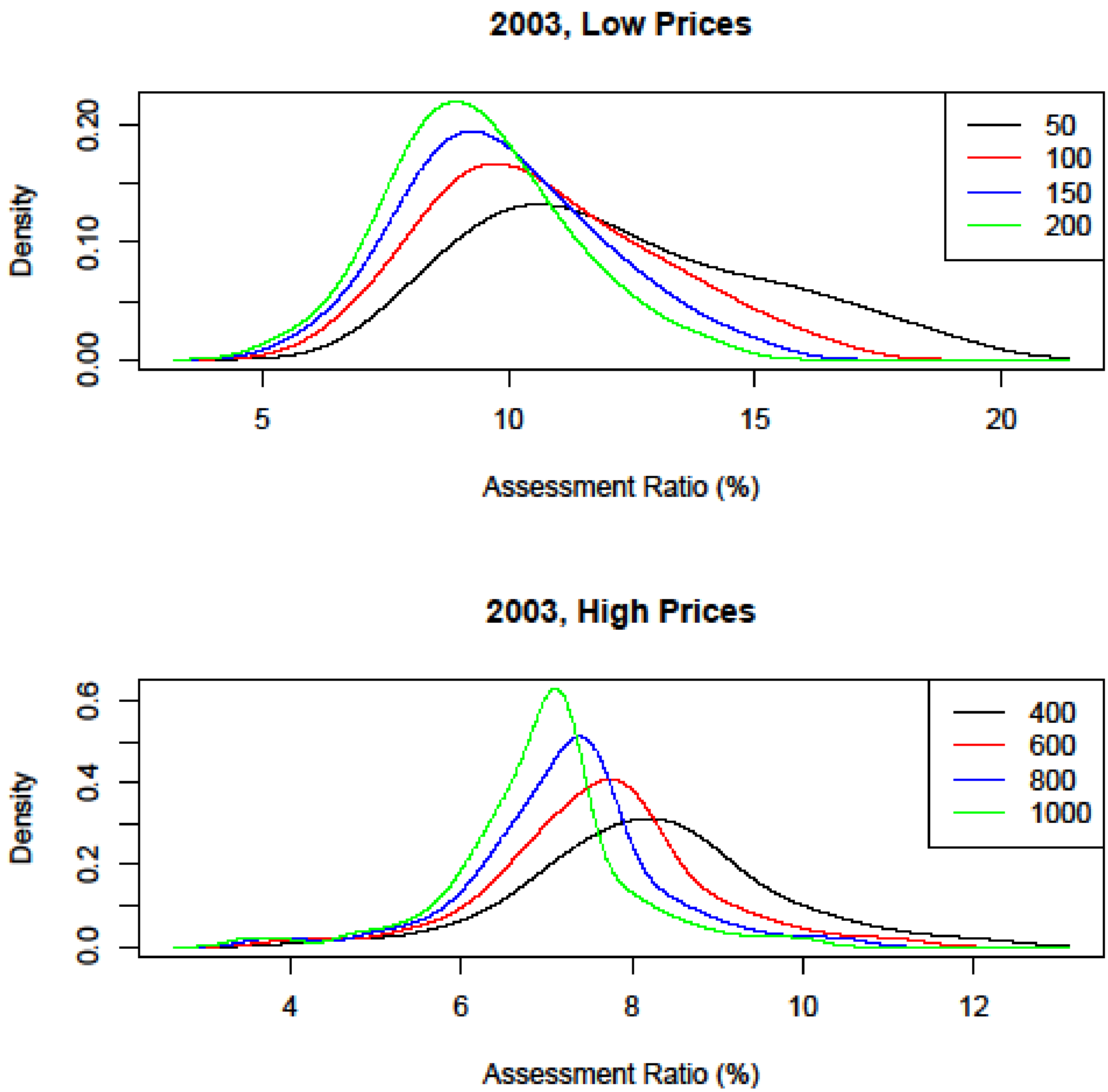


Figure 3

Estimated Assessment Ratio Densities by Sale Price, 2006

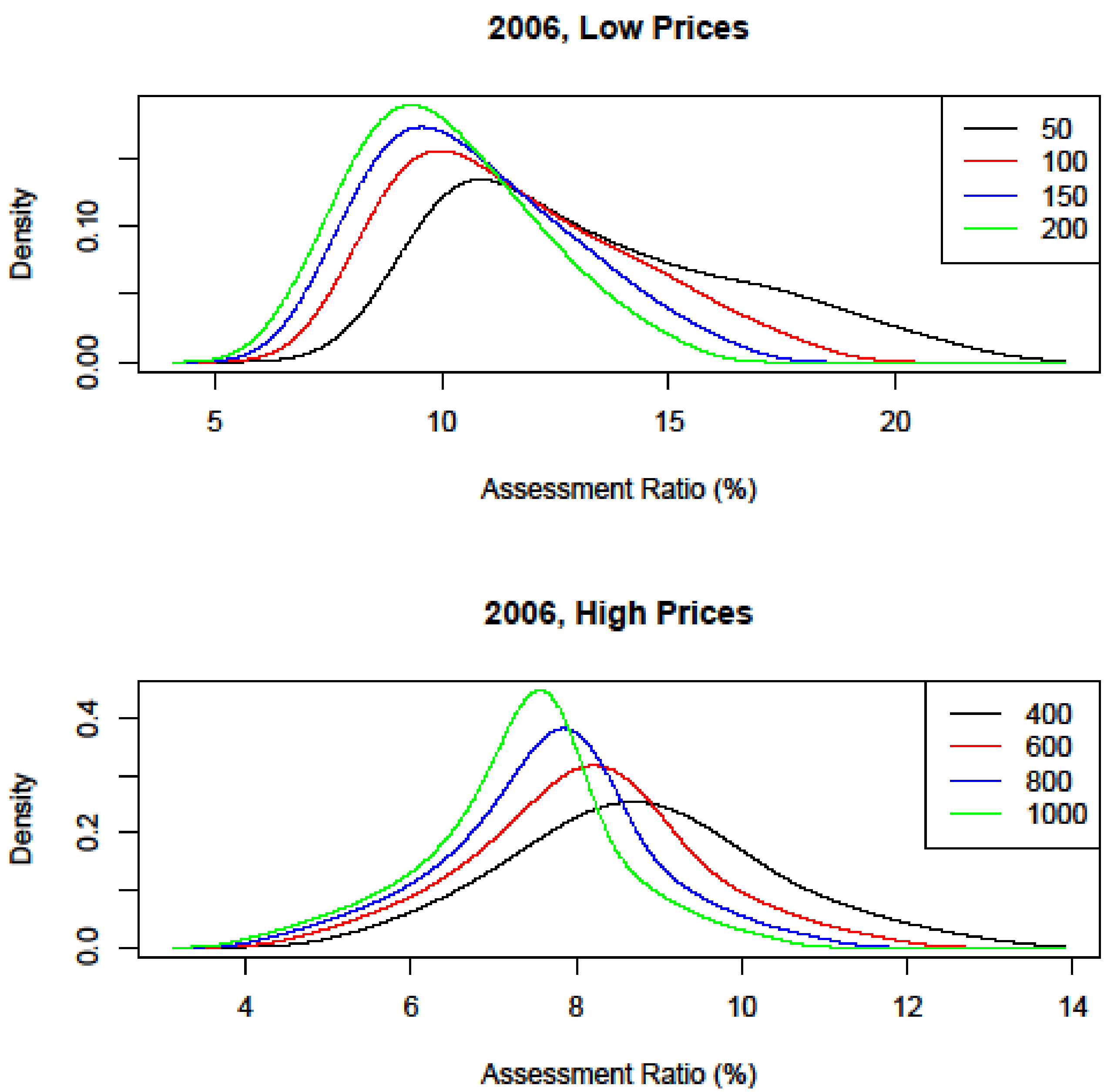
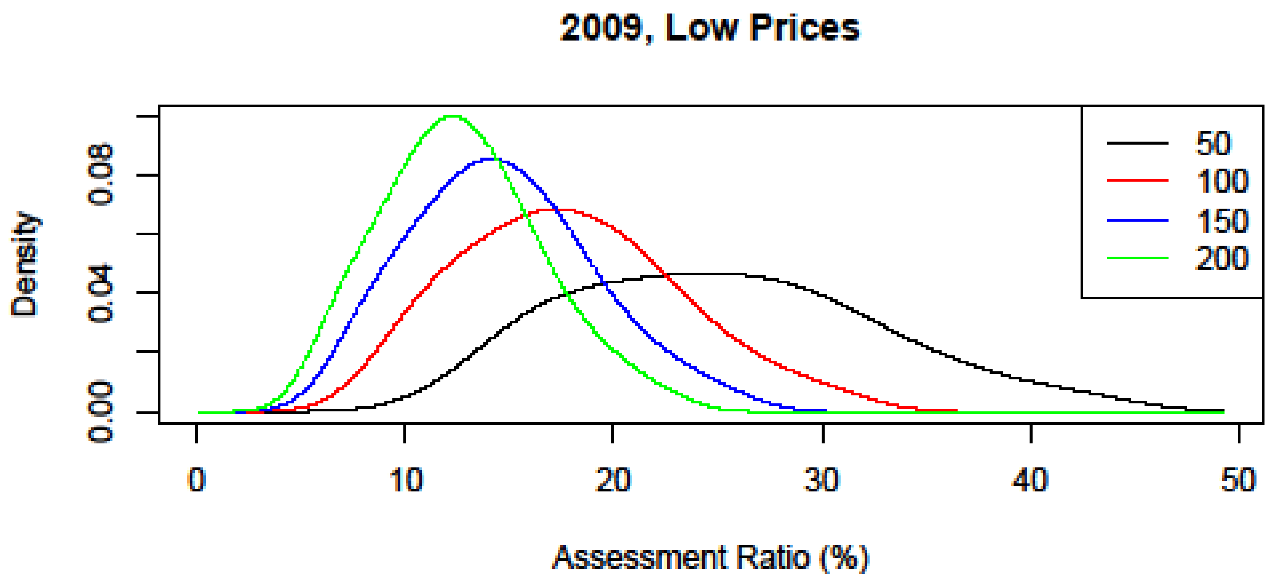


Figure 4

Estimated Assessment Ratio Densities by Sale Price, 2009



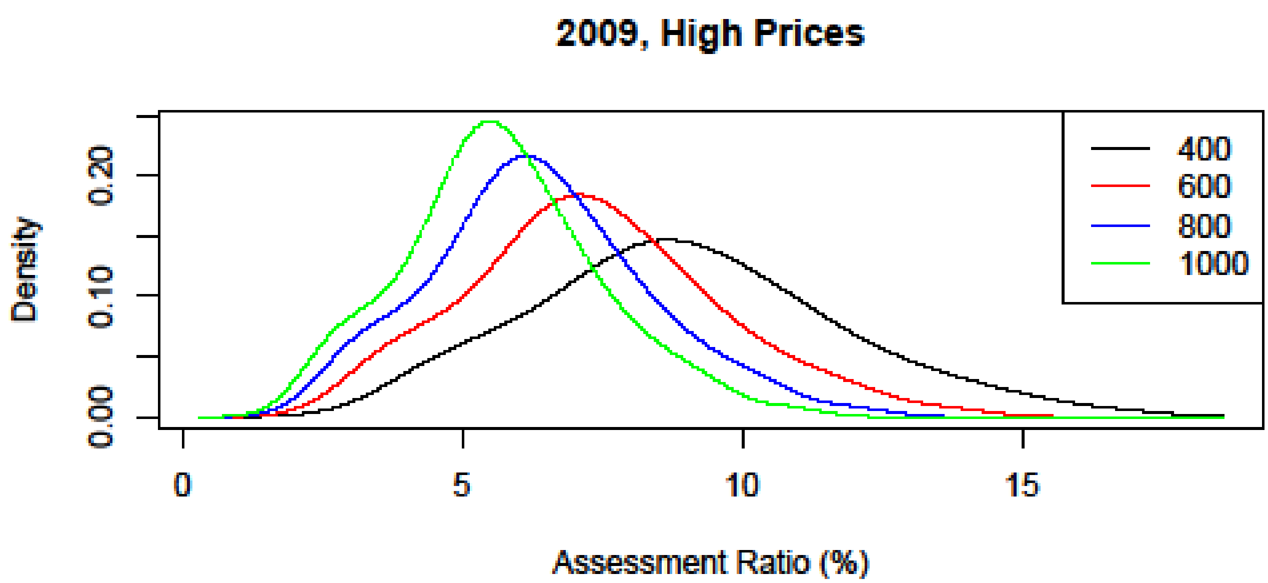


Figure 5

CPAR Quantile Estimated Densities, 2006

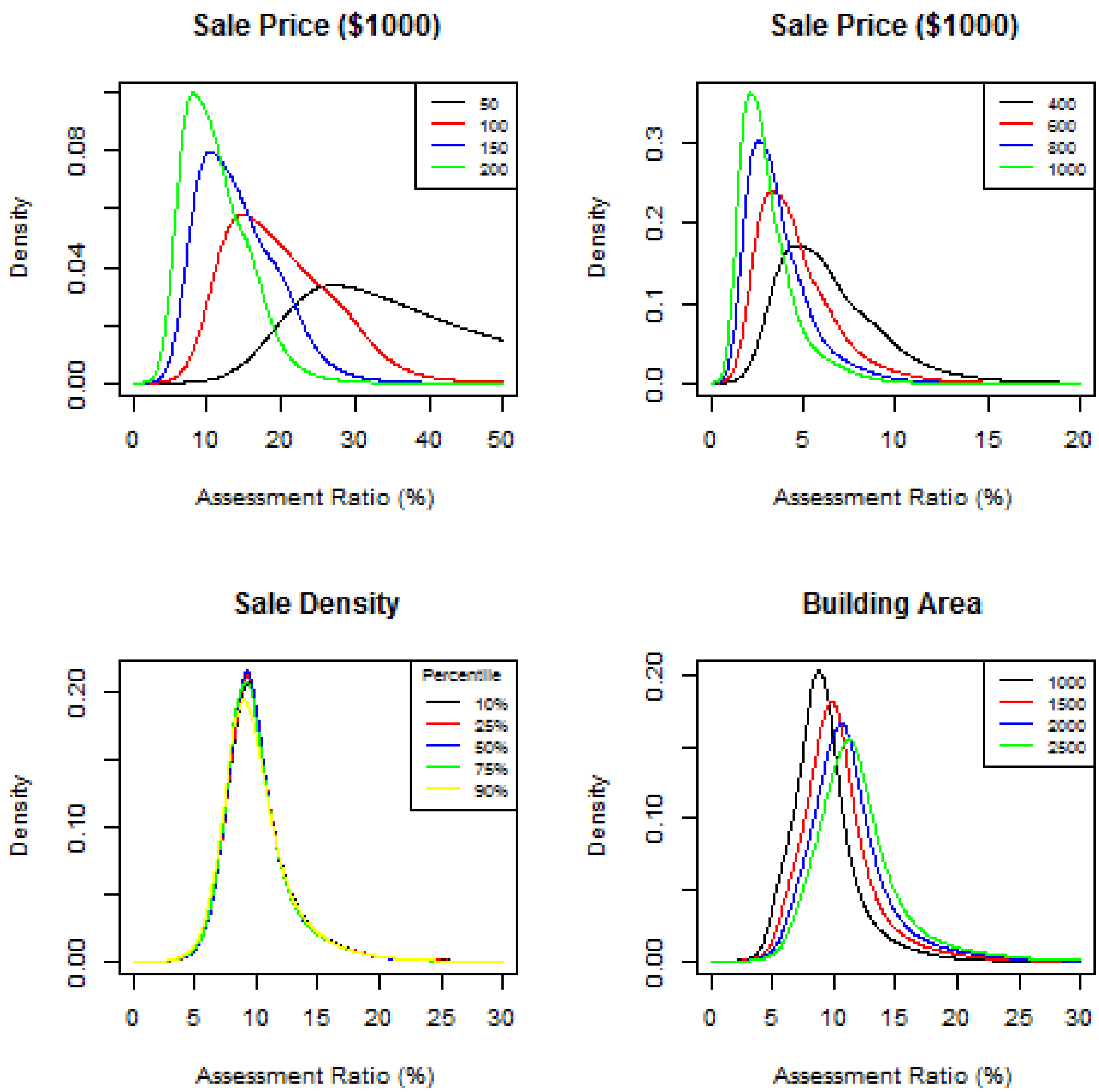


Figure 5 (cont'd)

CPAR Quantile Estimated Densities, 2006

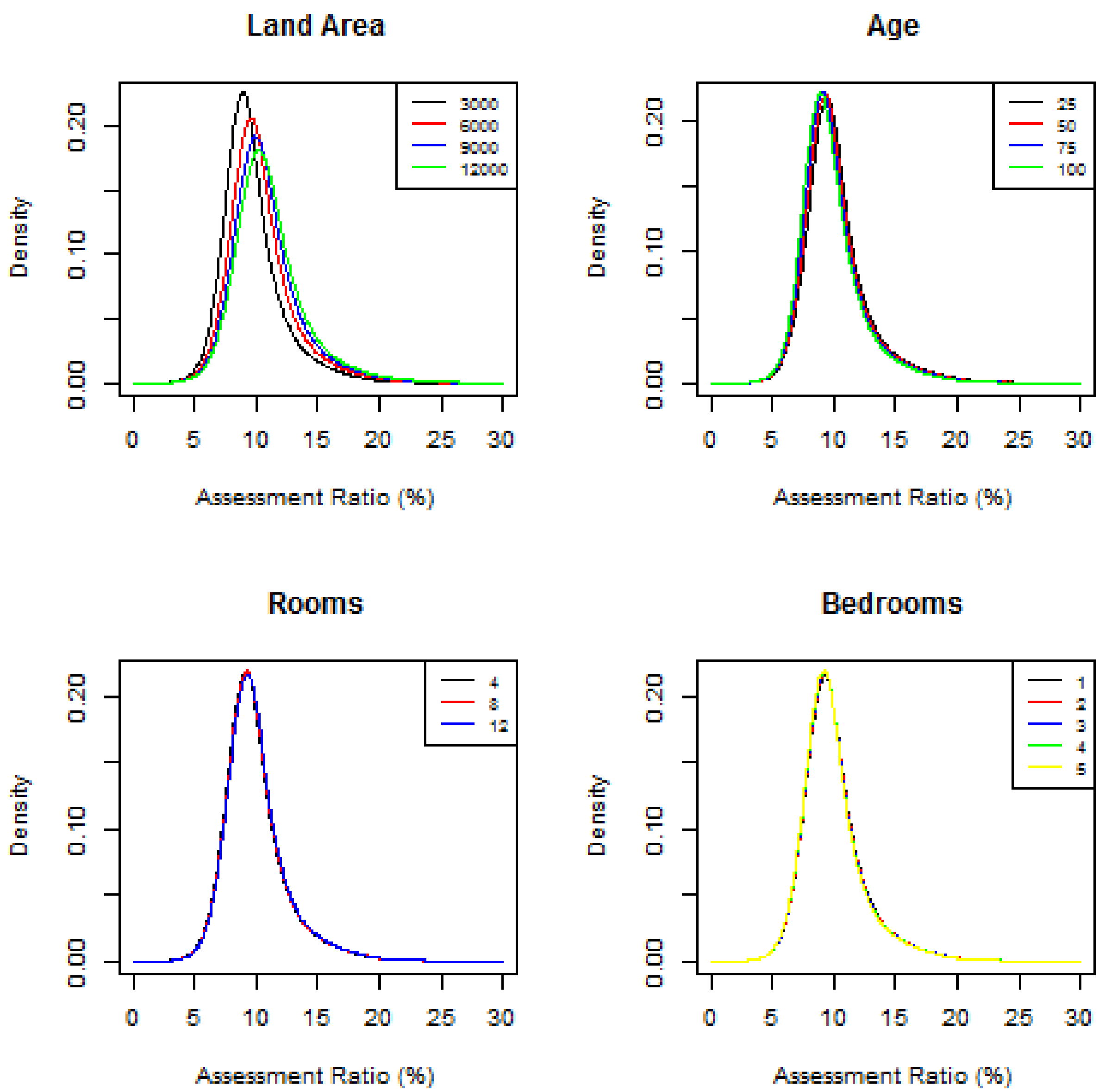


Figure 5 (cont'd)

CPAR Quantile Estimated Densities, 2006

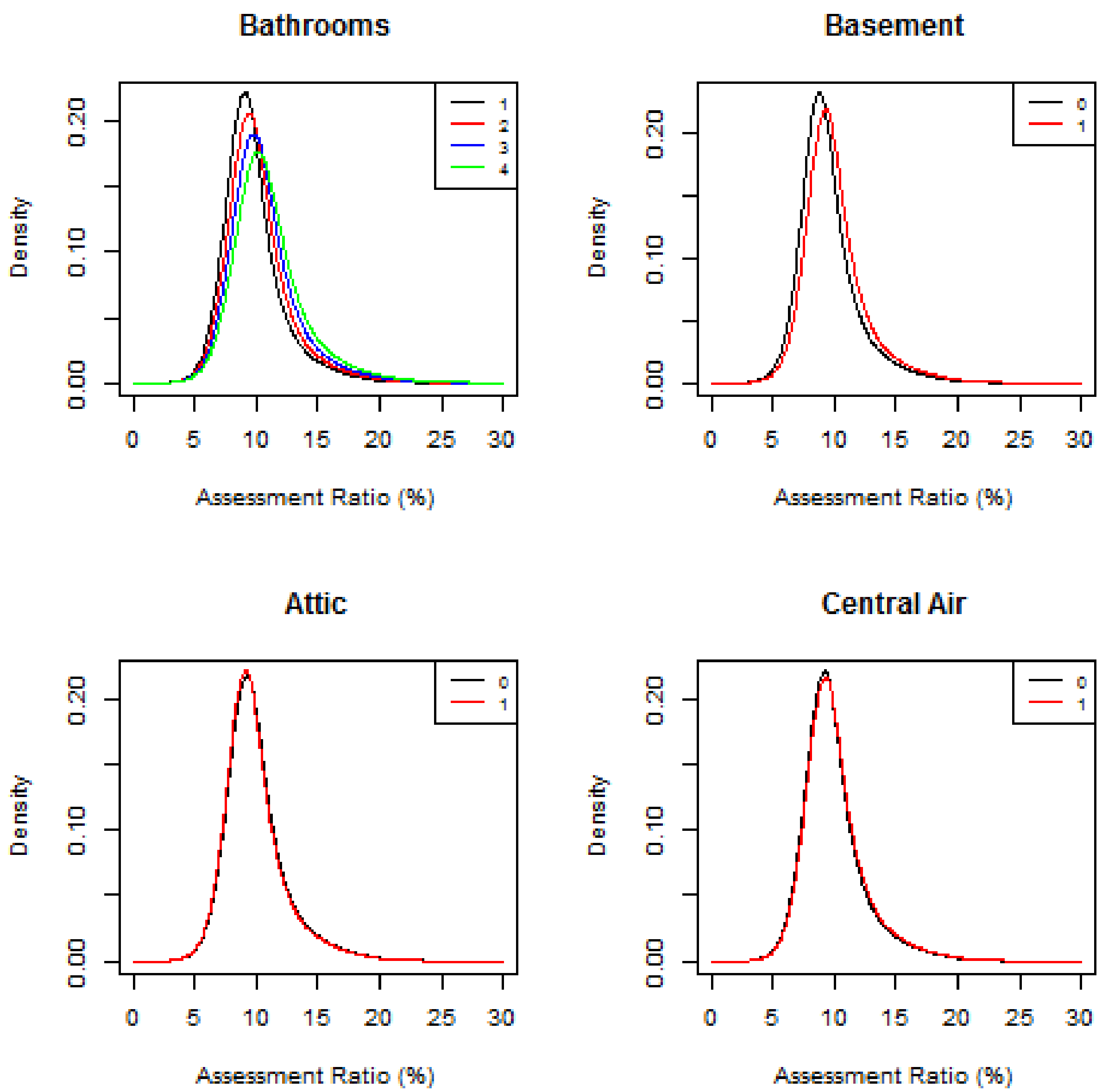


Figure 5 (cont'd)

CPAR Quantile Estimated Densities, 2006

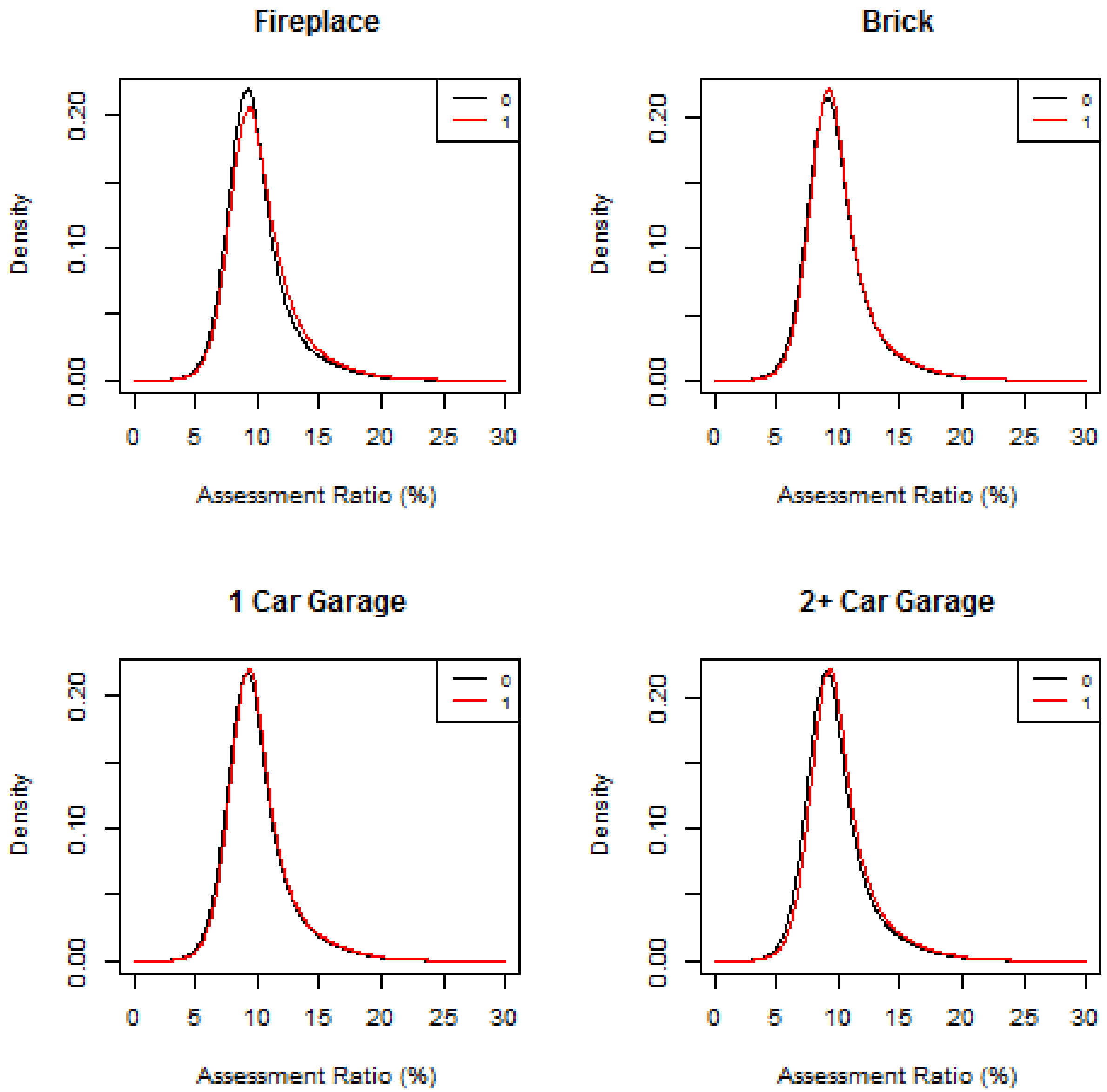


Figure 5 (cont'd)

CPAR Quantile Estimated Densities, 2006

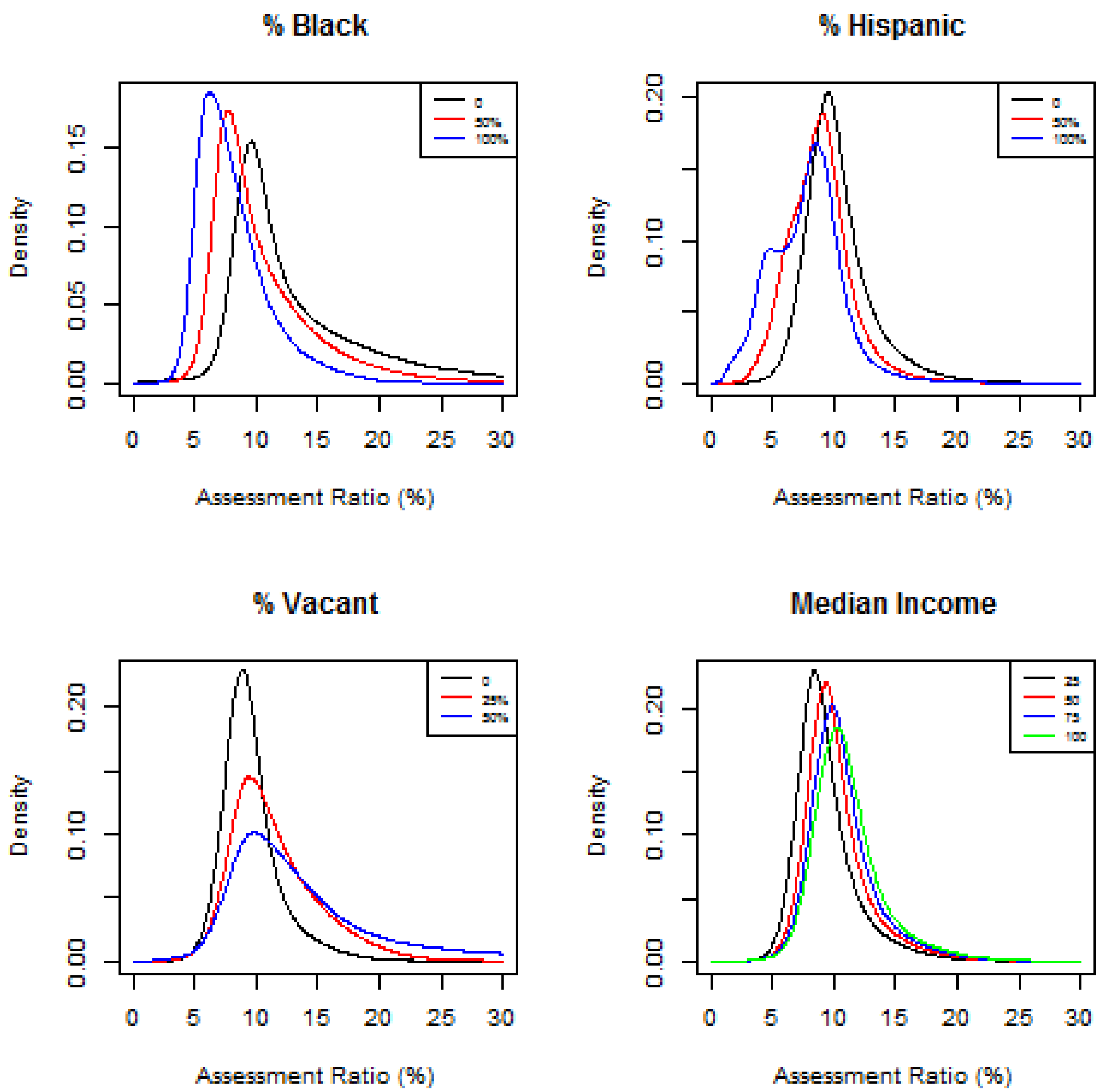


Figure 6

Statutory v. Estimated Quantile Distributions of Tax Payments

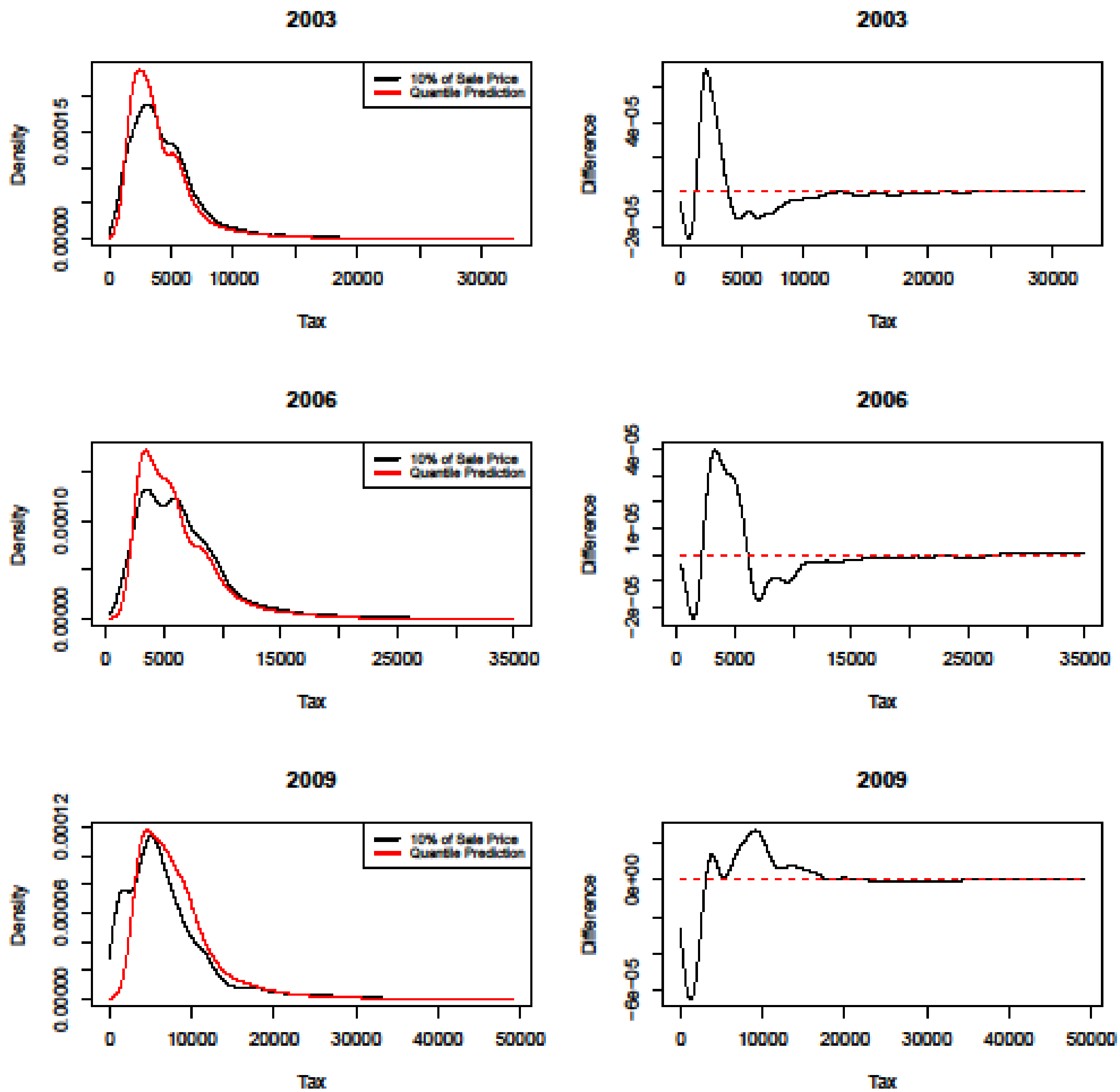


Figure 7

Statutory v. Estimated Quantile Distributions of Effective Tax Rates

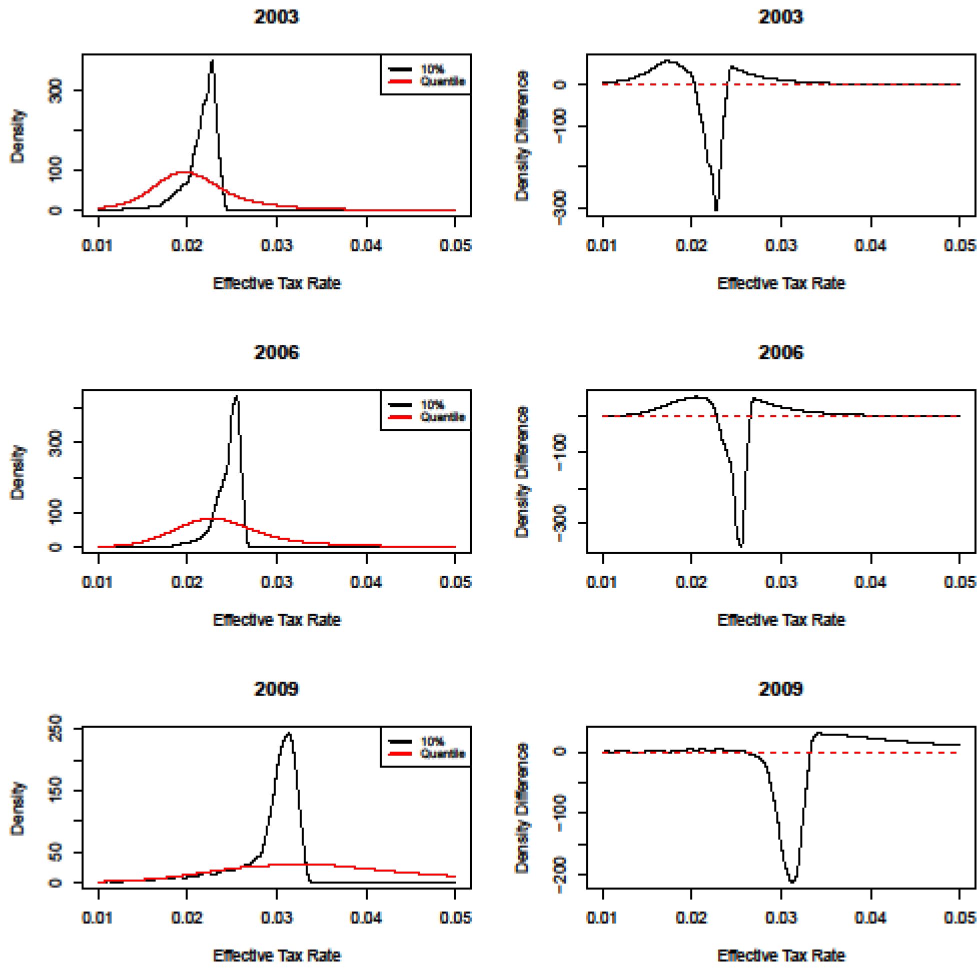


Figure 8

Statutory and Quantile Effective Tax Rates by Sale Price