DRAFT – Please do not circulate or quote without permission

Estimating Residential Land Prices and Airport Infrastructure Impacts Using Local Polynomial Regressions

Jeffrey P. Cohen, Cletus C. Coughlin, and John M. Clapp*

May 2014

Abstract

Airports facilitate many economic activities and likely affect the value of many resources, including land. Using residential home sales in Denver during 2003-2010, we use an innovative approach – Local Polynomial Regressions – to separate the value of land from the value of structures. Next, for the years in which a property was not sold, we interpolate land values for each property in our sample in each year. To assess the accuracy of our interpolations, we perform a within-sample forecasting exercise and determine that the Normalized Root Mean Squared Error is approximately 0.4%. Finally, we estimate the impacts of changes in airport infrastructure improvements on land values. We find that airfields, terminals, parking, and intermodal transportation lead to higher land values in the short-run, while "other" airport infrastructure lead to lower land values. We find similar results with a longer-run perspective in terms of the signs and significance of each of these airport capital stock variables.

Keywords: Local Polynomial Regression, Airport Infrastructure, Land Values
JEL Classification: C14, R42, R51, R53

*University of Hartford, Federal Reserve Bank of St. Louis, and University of Connecticut, respectively. The authors appreciate the assistance of Brett Fawley for his interpolations of the land values and David Lopez, Diana Cooke, and Lowell Ricketts. Cohen acknowledges the support of the Lincoln Institute of Land Policy. Clapp acknowledges support from the Center for Real Estate, University of Connecticut. The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Introduction

A common, yet challenging, issue in urban economics and local public finance is to produce an accurate estimate of the effect of infrastructure on land prices. One approach to obtain estimates of landowners' valuation of airport improvements would be to consider the implied value of land in sales of particular houses (and/or plots of vacant land) near airports, based on the product of the sales price and the ratio of the assessed value of land to total assessed value. After imputing the value of land for all properties sold, a regression analysis based on the relationship between changes in the sizes of airport capital stocks between two periods and changes in land values across space over the same time frame, after controlling for other factors that may affect the land values, can be undertaken. If positive, this estimated coefficient on the capital stock variable would represent the surplus obtained by homeowners or landowners as a result of airport improvements, after controlling for other factors that may have influenced their surplus.

Any approach using assessed values can be criticized due to the complex interaction between land values and the values of improvements, an interaction that may not be included in assessed values.[1] We introduce an alternative approach that considers the interaction between structure and land. The local polynomial regression method uses a nonparametric model for land values over time and a linear ordinary least squares model for the characteristics of the structure. A backfitting method insures independence between land and structure.

To generate land values, our empirical analysis uses housing sales data in Denver, over the period of 2003-2010; data include sales price, location, and various housing characteristics.[2,3] Our findings reveal that the interpolation approach is quite accurate. In an application of our land estimates, we examine the impact of airport infrastructure on land prices. We examine airport spending data available from the Federal Aviation Administration in several categories. Our results highlight the relationship between changes in the capital stock and land value. Most types of airport infrastructure, including airfields, terminals, parking, and intermodal transportation, are positively related to land values.

The major contributions of this paper include our application of the Local Polynomial Regression model (LRM) to separate the value of land from the value of structures.

---

[1] In most U.S. jurisdictions, property is taxed on the basis of total value, so the assessor has little incentive to be careful about separating land value and structure value.

[2] We have chosen this time frame due to the availability of airport infrastructure investment data over this period.

[3] We also analyze properties in Atlanta over the same period, but due in part to the sharp drop in the Lincoln Institute's land price index in the years 2008 through 2010, the land value estimates and the resulting analysis of their determinants are unstable. This suggests that caution should be exercised when attempting to use our approach if the time period under analysis includes sharp decreases in land prices due to a real estate "bust".

Second, we develop an interpolation method to construct balanced panel data: that is, we estimate land value at every location at each point in time. Third, the panel data allows us to use a fixed effects model to determine the impacts of changes in airport infrastructure capital on land values, which is important for the purpose of extracting land value from property owners. Fourth, more generally, our work contributes to the literature of the impacts of infrastructure improvements on housing prices.

This line of research is important because separate estimates of land and building value are used to adjust property tax assessments for structure depreciation and for changes in land value.[4] Moreover, the ratio of structure to land is used by investors to choose the time and intensity of redevelopment (Hendriks, 2005; Clapp and Salavei, 2011; Ozdilek, 2012). The application we pursue here is suitable for use by planners to evaluate the effects of infrastructure investment and to decide where to allocate urban renewal funds.

The body of the paper consists of several sections. The next section is a literature review. This review is followed by a summary of our estimation procedure for land prices, which is a semiparametric approach developed by Clapp (2004), the interpolation procedure, and the resulting land values. The following section describes the details of our data. In the next section, the determinants of the land values are examined. The key determinants are various categories of airport infrastructure. A summary of key results and several questions for further research completes the paper.

Literature Survey

Housing prices (i.e., the total price that includes land and structures) in the U.S. experienced a dramatic increase in the years leading up to 2008, followed by a substantial "bust" in the subsequent years.[5] Cohen, Coughlin and Lopez (2012) describe how some regions of the U.S. faced more of a downturn than others. Figure 1 depicts quarterly housing prices in Denver in the years 2000 through 2012, which includes the boom, bust, and nascent recovery in the U.S. housing market.

During roughly this same time period, there has been substantial variation in airport infrastructure investment, as well as depreciation, at the Denver airport. The past investment levels, net of depreciation, can be used to obtain a stock of airport infrastructure for a variety of airport infrastructure categories. These stocks of infrastructure are presented in Table 1 for Denver over the period 2003 through 2010.

A large literature examines how and to what extent transportation infrastructure becomes capitalized into housing prices, including McMillen and McDonald (2004), Weinberger (2000), and Forest, Glen, and Ward (1996). These papers use a hedonic pricing approach

[4] The relative volatility of the land value component contributes to macroeconomic risks as suggested by Davis and Palumbo (2008) and by Bourassa *et al.* (2011).

[5] See Cohen, Coughlin, and Lopez (2012) for details.

to assess the impacts of the transportation infrastructure on housing prices (or commercial rents, in the case of Weinberger). Gatzlaff and Smith (1993) use both a repeat sales approach and a hedonic housing prices approach to examine the effects of a new rail line in Miami, and find that it had only a minimal impact. However, none of these papers directly address the issue of transportation infrastructure capitalization by distinguishing the distinct impacts on land values.

During a period with wide fluctuation in housing prices as well as improvements in public services such as airport infrastructure, it could be useful to isolate for policy makers and fiscal authorities the impacts of improvements in airports on land values. This capitalization could provide the basis for land value capture based on airport improvements. Chapman et al. (2009) examine the feasibility of land value taxation for financing transportation infrastructure in Utah, and find that in addition to being a non-distortionary form of taxation, the land value tax could generate significant revenue and would be relatively straightforward to administer.

Cohen (2012) summarizes the ideas behind value capture at airports. He describes that economic theory underlying how value capture implies improvements to airports become capitalized in land values. These land rents can be taxed with a land value tax, and this tax is preferable because it is non-distortionary. However, in order to achieve land value capture in practice, it is necessary to obtain estimates of land separately from the improvements to land. This is one of a number of practical issues involved in estimation of transportation infrastructure capitalization into land values.

The separation of urban land value from structure value is an important topic made challenging by the scarcity of vacant land sales in an urban setting. Hendriks (2005) evaluates three methods used by appraisal professionals for this purpose: fractional apportionment (FAT), rent apportionment (RAT) and price apportionment (PAT) theories. He raises substantial questions about each, recommending that appraisers caution their clients about the unreliability of apportionment methods. Our LRM method is most closely related to PAT since it uses sales prices together with location and property characteristics to allocate value (i.e., predicted price from a hedonic model) between land and structure.

Longhofer and Redfearn (2009) examine how one might in practice disentangle the value of land from the value of structures on the land. Longhofer and Redfearn argue that land and structures are inseparable, as does Hendriks (2005). Their argument is that houses within a neighborhood are fairly homogeneous. For example, it may not be possible to buy a small house on a small lot in a neighborhood with much larger houses. They give an example where the supply of pools cannot adjust to the demand within a given neighborhood, so pools are priced “too high” in some neighborhoods. They use vacant land on the periphery of a city, along with the estimation technique of locally weighted regressions, to estimate the values of land throughout the city. One drawback of their approach, however, is that it requires data on vacant land sales to derive the land values for all properties.

As an alternative, Clapp and Salavei (2010) implement an "option value" approach that addresses the problem from a different perspective than Longhofer and Redfearn: existing structure relative to optimal structure at any time will influence the value of the land. The costs of adjustment are high, so it takes a long time to reach the trigger point to redevelop. The costs of rebuilding to a new optimal level are the cost of construction and the sacrificed rents from the existing structure (i.e., this is an exchange option). Thus, a house with specific characteristics, such as age, layout, and size, will have implicit characteristic prices that vary with the land value.

Identical to the approach we utilize, Clapp (2004) uses a local polynomial regression model (LRM) to disentangle land and structure prices by holding constant structure value and extracting the associated land value. Similar to Longhofer and Redfern (2009), a locally weighted regression is part of the research design.

In the context of the effect of transportation infrastructure, another issue is the timing of the capitalization effect. Clearly, market prices respond to many types of information, so price adjustments may occur at the time of the expansion announcement. What is not clear is the time path of the adjustment process. Prices may not adjust fully until the investments are in place or even later if the potential effects of the investment, such as new services and ease of use, are initially unclear.

Jud and Winkler (2006) examine the impact of announcement of construction of a new hub airport, which is expected to lead to greater noise, on housing prices in Greensboro, NC. They find that this post-announcement effect is nearly a 10 percent reduction in housing prices within 2.5 miles from the airport. Agostini and Palmucci (2008) found an announcement of new transit station construction led to an increase in nearby housing prices ranging between 3% and 8%. Similarly, McMillen and McDonald (2004) found the housing market began adjusting to a new rail line before the construction was completed. Clearly, it will be important in our context to consider both the announcements of expansions as well as the actual construction expenditures and dates, by examining first and second differences around the actual year of the expenditure. We address the timing issue by examining long-term as well as short-run effects of changes in the value of various categories of airport infrastructure stocks on land values.

Method for Separating Land and Structure Values

First, we consider the problem of obtaining the land values separately from structure prices. The Clapp (2004) local regression model (LRM) and Clapp and Salavei (2010) "option value" approach is followed here. The choice of LRM is motivated by the observation that structures are reproducible at current construction costs whereas location value (the value of the right to build a single family residence at a given location) varies substantially across space and time. By separating the structure and land components we can estimate the variation in location value over time, and correlate it with airport expansion events.[6]

We begin by motivating our problem with a standard hedonic model, a parametric method for finding implicit prices for each element of the vector of housing characteristics (structure and location), and a price index independent of these characteristics. Regress the log of sales price (*lnSP*) on a vector of house structure characteristics (Z), locational characteristics (S), and time (*t*) which is represented here in the form of annual time dummies, Q_t:

$$lnSP_{it} = \gamma_0 + Z_i\alpha + S_i\beta + \gamma_0 Q_0 + \gamma_1 Q_1 \cdots + \gamma_T Q_T + \varepsilon_{it} \qquad (1)$$

where ε is typically an iid noise term that is assumed to be normally distributed for the purposes of hypothesis testing.[7]

The cumulative log price index for a standard house in the area where the data were collected is measured by the parameters on the annual time dummies, γ. The assumption is that the parameters on structure and location are constant over time. Since they are not, we are measuring the average implicit prices, α and β over the time interval *T*. Thus, any change over time is forced into the estimates of the γ parameters; they can be considered an approximation to a pure time component that shifts the constant of the regression, γ_0.

Before estimating the local polynomial regression model, it is necessary to define the grid size and the bandwidth. The grid is composed of equally spaced time, latitude, and longitude points that span the data. In our model, there are 20 time, 15 latitude, and 15 longitude points, for a total of 4500 "knots" on the grid. The size of the bandwidth determines whether or not an observation will be used to estimate the function value at the knot. See below for a discussion of the cross-validation bandwidth selection approach. For this paper, the bandwidth is chosen to be {.3σ(time), .3σ(latitude), .3 σ(longitude)}. The knots on the grid are estimated from the function values of the observations "close to" the knots. The local polynomial regression then fits a surface to the observations conditional on the function values estimated at each knot on the grid. The LRM is designed to allow substantial nonlinearity in the spatial and time dimensions: it fits a value surface at each point in time as an alternative to estimating the set of parameters in equation (1). The LRM views price index and value surface estimates as

[6] Davis and Palumbo (2008) develop a model decomposing property value into structure and land components, and they use this model to find significant changes in land value over time and across metropolitan areas. They depend on subtracting the cost of construction from sales prices, whereas we use the implicit value of the structure. The Davis and Palumbo approach can be viewed as a robustness check.

[7] The log of sales price is the dependent variable because logarithms control for heteroscedasticity and some nonlinearity. Using sales price, SP, instead would cost degrees of freedom; See Hastie and Tibshirani (1990), pp. 52-55 for a discussion of degrees of freedom for smoothing models.

descriptive exercises that are not designed to test hypothesis about parameters. Writing the model as follows emphasizes the nonlinear and nonparametric aspect of the LRM:

$$\mathit{ln}\mathrm{SP}_{it} = f\left(Z_i, S_i, t_i\right) + \varepsilon_{it} \tag{2}$$

We allow the function *f( )* to be nonlinear because local house prices rarely move in a straight line over time and a nonlinear spatial pattern is well known. These nonlinearities, as well as the descriptive purpose of the model, make nonparametric smoothing regressions an ideal tool.

The first step to estimating the value of residential location and time at each of the 326,744 observations is to run the local polynomial regression to estimate the value of house prices given time and space, as well as to estimate the value of structural characteristics given time and space. The next step is to subtract these estimated values from the original values to determine the sales price (Y*) and value of structure (Z*) independent of time and space. The parameters of the structural characteristics are estimated by regressing Y* on X*. Finally, the partial residuals are calculated by subtracting the estimated value of structural characteristics. We use these partial residuals as input to the local polynomial regression to create a smooth function of land values at each point in the grid, which are then used to estimate the value of land at each observation.

LRM estimation methods can be introduced by imagining that a number, q, of identical houses trade at a given point in space and time, denoted by the fixed vector ($\mathbf{z}_0$, $\mathbf{s}_0$, t_0). Then, an obvious way of estimating equation (2) at the fixed point would be to average those prices:

$$\hat{f}\left(\mathbf{z}_0, \mathbf{s}_0, \mathrm{t}_0\right) = \frac{\sum_{i=1}^{q} \mathit{ln}\mathrm{SP}_{it_0}}{q} - \frac{\sum_{i=1}^{q} \varepsilon_{it_0}}{q} \tag{3}$$

The error term results from negotiation between heterogeneous buyers and sellers. Since the average error term will tend to zero as the sample size gets large, we will have a consistent estimator of a point on the value surface at the given point in time.

Actual sales prices are spread out in space and time as well as over the range of housing characteristics, **z**. If the data were densely distributed over these characteristics, then we could average prices that are "close to" any particular point in characteristic space ($\mathbf{z}_0$), physical space ($\mathbf{s}_0$) and time (t_0). This averaging process is very much in the spirit of nonparametric smoothing.

Nonparametric smoothing implements this local averaging idea by down-weighting observations that are more distant from the fixed point:

$$\hat{f}(\mathbf{z}_0, \mathbf{s}_0, t_0) = \sum_{i=1}^{q} \frac{K_h(\cdot) ln\mathrm{SP}_{it_0}}{\sum_{i=1}^{q} K_h(\cdot)} - \frac{\sum_{i=1}^{q} \varepsilon_{it_0} K_h(\cdot)}{\sum_{i=1}^{q} K_h(\cdot)} \tag{4}$$

where the weighting function, $K_h(\cdot)$, is defined such that greater distances (e.g., larger values for $S_i - \mathbf{s}_0$) imply lower values for K; h is bandwidth, a set of parameters that govern the selection of points "close to" the target vector.[8]

Bandwidth selection is a trade-off between high variance (bandwidth is too small) and high bias (bandwidth is too large). This paper uses a cross validation method for bandwidth selection: See Wand and Jones (1995, Chapter 4). Locally adaptive bandwidths are allowed by increasing bandwidth until 20 observations are within one bandwidth of the fixed point.

Equation (4) is a special case of local polynomial regression (LPR), given a specific point in space and time, $\mathbf{x}_0 = (\mathbf{z}_0, \mathbf{s}_0, t_0)$, the data, $\mathbf{X}_i = (\mathbf{Z}_i, \mathbf{S}_i, t_i)$ and $Y_i = ln\mathrm{SP}_i$. Local polynomial regression now takes the form of equation (5):[9]

$$Y_i(\mathbf{x}_0) = \beta_o + (\mathbf{X}_i - \mathbf{x}_0)\beta_1 + (\mathbf{X}_i - \mathbf{x}_0)^2 \beta_2 + \ldots + (\mathbf{X}_i - \mathbf{x}_0)^p \beta_p + \varepsilon_i \tag{5}$$

Here, the β_j (j=1,…,p) are column vectors with number of elements equal to the columns of $\mathbf{X}_i$: β_0 is a scalar.[10] Note that, when $\mathbf{X}_i$ equal $\mathbf{x}_0$ then equation (5) reduces to β_0, the parameter of interest. Thus, LPR fits a surface to the Y-values conditional on the values of $\mathbf{x}$ given by $\mathbf{x}_0$: E.g., $\mathbf{x}$ is a grid of equally spaced points that span the data; the level of Y is estimated conditional on each knot of the grid.

Kernel weights are applied when estimating equation (5):

$$\mathrm{MiN}(\hat{\beta}) \sum_{i=1}^{n} \{Y_i - \beta_0 - \ldots - (\mathbf{X}_i - \mathbf{x}_0)^p \beta_p\}^2 K_h(\mathbf{X}_i - \mathbf{x}_0) \tag{6}$$

where the weights are applied to each of the variables including the constant term (the vector of ones). The only difference between the weights in equation (6) and those in

[8] Equation (4) is the well-known Nadaraya-Watson (NW) smoother. See Clapp (2004) for details on the choice of the kernel weighting (i.e., density) function. Experts in this field have found that the choice of bandwidth is much more important than the choice of a kernel density function.
[9] The exponents in equations (5), (6) and (8) are taken element-by-element.
[10] The parameters other than β_0 allow for curvature around $\mathbf{x}_0$; a weighted average of neighboring points, equation (4), would ignore curvature. Also, comparing equations (6) and (3) show how LPR takes local averages.

equation (4) is that time has been entered as a vector rather than a scalar.[11] Thus, the parameters estimated using equation (6) can be defined as follows:

$$\hat{\beta}(\mathbf{x}_0) = (\mathbf{X}_\mathbf{x}^{T'} W_x \mathbf{X}_x)^{-1} \mathbf{X}_x^{T'} W_x Y \tag{7}$$

$$\mathbf{X}_x = \begin{bmatrix} 1 & \mathbf{X}_1 - \mathbf{x}_0 \cdots (\mathbf{X}_1 - \mathbf{x}_0)^p \\ \vdots & \vdots \quad \ddots \quad \vdots \\ 1 & \mathbf{X}_n - \mathbf{x}_0 \cdots (\mathbf{X}_n - \mathbf{x}_0)^p \end{bmatrix} \tag{8}$$

$$W_x = diag\{K_h(\mathbf{X}_1 - \mathbf{x}_0), \ldots, K_h(\mathbf{X}_n - \mathbf{x}_0)\} \ . \tag{9}$$

This regression is repeated for each point on the $\mathbf{x}_0$ grid.

LPR is a weighted OLS regression at the point $\mathbf{x}_0$, so we can test hypotheses on the $\hat{\beta}$'s by assuming that they are multivariate normal with the following covariances:

$\left(Var-Cov(\hat{\beta})\right) = \left(\mathbf{X}_x^T W_x \mathbf{X}_x\right)^{-1} \mathbf{X}_x^T W_x \mathrm{V} W_x \mathbf{X}_x \left(\mathbf{X}_x^T W_x \mathbf{X}_x\right)^{-1}$, where V is a diagonal matrix of variances for ε_i.

The treatment of time is much more flexible in equation (5) than it would be in the OLS model, equation (1). LPR treats time as an addition to the spatial dimension: that is, we grid time as finely as the data permit at each point in space. For example, to estimate the value function at 10 points in time, and at each point of a 30x30 spatial grid, we need 9,000 regressions. Each estimator gives high weight to observations that are nearby in space and time and lower weight to those that are farther away.

The semi-parametric LRM model enters because of the "curse of dimensionality." As a practical matter, there would typically be five or six variables for structural characteristics (e.g., interior area, bathrooms) in the **X** matrix. If all were represented by even a coarse grid, the data would typically be sparse near any point. The semi-parametric solution assumes linearity for the equation (1) parameters, α, on all the housing characteristics in the matrix **Z.**[12] In the LRM method, an LPR model is used to estimate these coefficients allowing for conditional on the location of the house. This approach addresses the concerns of Longhofer and Redfearn (2009) and provides statistical independence between the estimated coefficients on Z and the nonlinear part of the model. Then the residuals from this regression can be fit with an LPR model.

---

[11] The metric for time is different from that for space (and also different for structural characteristics). Cross-validation (CV) is used to select optimal bandwidths: If CV indicates that observations more distant in space should receive more weight, then a larger bandwidth will be chosen in the spatial dimension. This addresses a concern raised by Pavlov (2000).

[12] Of course, a nonlinear relationship (e.g., with building age) might be more appropriate. The point here is to focus on the highly nonlinear space-time relationships.

Following this logic, the LRM method begins by estimating equation (1), then revising the $\hat{\alpha}$'s to assure independence from the land value estimates: this is termed the "backfitting" method.[13] Then, partial residuals are taken by subtracting the estimated value of structural characteristics:

$$partres_{it} = lnSP_{it} - Z_i \hat{\alpha} \tag{10}$$

where *partres* is the partial residual from equation (1).

The nonparametric part of the LRM model is:

$$partres_{it} = q\left(S_i, t_i\right) + \varepsilon_{it} \tag{11}$$

where S_i is now defined as the latitude and longitude for house i. Typically, LPR estimation of equation (11) can deal with the two spatial dimensions and the time dimension without substantially increasing the standard error of the *q(.)* estimate. From another perspective, the method requires sufficient density of transactions near the given target point, $\left(S_i, t_i\right)$.[14] Estimation methods for standard errors reveal any problem with lack of data.

To summarize, the purpose of the LRM is to estimate location value over time, $q\left(S_i, t_i\right)$, equation (11). Since we subtracted an average value of structural characteristics, $Z_i \hat{\alpha}$ estimated so as to require independence from $q\left(S_i, t_i\right)$, the LRM estimate may be taken as a reasonable approximation to location value.[15]

---

[13] See Clapp (2004) for details on methods for estimation of $\hat{\alpha}$.

[14] A problem with temporal aggregation in the standard hedonic method - the bunching of transactions within the quarters, equation (1) - is handled nicely by the kernel weighting scheme applied to equation (11); the t_i variable is based on day, month, and year of the transactions and the t_0 target is typically the middle of the year.

[15] However, it may be objected that location value should be estimated as property value less construction costs, as suggested by Davis and Palumbo (2008). To get to this quantity, one would add back $Z_i \hat{\alpha}$ and then subtract construction costs. An approximation to construction costs can be obtained by assuming that they are invariant within the metropolitan area and that they change slowly over time as the costs of material and labor change. With these assumptions, the level of construction costs at time zero is the same for all houses in the city. One can use the Marshall Valuation Service (MVS) to approximate this level. Then percentage changes over time can be approximated by using a construction cost indexes such as those published by Engineering News-Record (ENR, http://enr.construction.com/economics/). With these adjustments, location value is estimated by:

Applications

Once we obtain the land values, holding constant for structure values, we perform two applicaitons. First, we perform an interpolation procedure to obtain estimates of land values for each house in our sample in every year over the period 2003-2010, and examine the accuracy of these land value estimates. In other words, the interpolation procedure estimates the value of land at each location, in every point of time. There are 2,613,744 (equal to 326,744 x 8) points to interpolate. Second, once we have interpolated these land values, we use them to assess the impacts of changes in various categories of airport infrastructure stock values on land values.

Land Value Interpolation

The interpolation procedure uses the grid of equally spaced knots from the local polynomial regression. Since this grid spans the data, each point to be interpolated is surrounded by 8 knots. To understand this grid, one might imagine a cube (or more generally, a prism with 6 sides and 8 vertices), with the point to be interpolated in the middle. The method of tri-linear interpolation (Bourke, 1999) approximates the land value in the center of the cube (or prism) using the values on the lattice points.

Figures 5-12 show the land values in quintiles for each year 2003 through 2010. The quintiles are not calculated by separate years, but rather for the entire period. For a given figure, the darker it appears, the higher the estimated land values. Thus, an eyeball comparison of Figure 5 with Figure 8 suggests that land prices tended to rise somewhat between 2003 and 2005. This is similar to Figure 2. Subsequently, land prices tended to decline for a few years and then flattened out. To confirm these results and indicate the general movement of land prices, here the medians in natural logs by year: 2003 – 9.90, 2004 – 9.92, 2005 – 9.95, 2006 – 9.94, 2007 – 9.90, 2008 – 9.84, 2009 – 9.83, and 2010 – 9.84.

The next step is to test the accuracy of the tri-linear interpolation using the land values from the local polynomial regression procedure as our standard of measure. We omit the land values from 20% of the 326,744 observations and use the 80% of known observations to "forecast" the missing 20%. We run the local polynomial regression to fit a smooth surface of land values based on the partial residuals of the 80%. The output is a grid, slightly different than in the initial case because it is produced using only 80% of the data. For the forecasting procedure to work, it is necessary that the 20% of omitted observations do not contain the most extreme values for time, longitude, and latitude. Also, since the partial residuals come from the LPR program, the same grid size and

---

$$\hat{q}\left(\mathrm{S}_i, t_i\right) + Z_i\hat{\alpha} - C_{it}$$

where C_{it} is an estimate of construction costs for house *i* at time *t*. This procedure can be considered as a robustness check.

bandwidth should be used for this step. The method of tri-linear interpolation estimates the 20% of missing land values, based on the grid defined by the 80%. We compare the 20% of land value estimated from the interpolation procedure to their estimated produced by the local polynomial regression.

The land values that we obtain from the LRM and interpolation procedure are in natural logarithms. We perform several exercises to show the accuracy of the interpolated land values. We calculate the Root Mean Squared Error (RMSE) and the Normalized Root Mean Squared Error (NRMSE), for the difference between the land values (in natural logs) and the interpolated land values for the 20% sample that is omitted. We also separately calculate the NRMSE for the land values converted into dollars. The NRMSE can be viewed as a unit-less measure that is analogous to a percentage, and is calculated by dividing the RMSE by the difference between the two extreme points among the combined set of the land values and interpolated land values. In our exercise, we find that the NRMSE is virtually identical, and equal to approximately 0.39%, for the two different units of land value. This small NRMSE supports the notion that our land value interpolation approach is very accurate.

Estimating the Impact Airport Capital Stocks on Land Values

The unit of observation is the individual house transaction (not repeat sales). Specifically, we consider major improvements such as terminal expansions, airfield improvements, parking structures, roads/transit/rail, and all other expenses, to construct airport capital stocks for each of these categories. These capital stocks control for depreciation. We identify the impact of a major improvement in year t off of change from before to after the event. Specifically,

o We expect distance from the airport to attenuate the effect.

o We expect long run changes (more than one year before and after changes in the airport capital stock) to be different from short-run effects (one year or less before and after). See Clapp and Ross (2004). The effect should build up to a larger total over a longer term, if the cause is a permanent increase in the number/frequency of airline service.

We also initially include cross-sectional dummies to control for unobservables. Also, those professional jobs requiring a lot of travel might locate closer to the airport, especially after expansion. However, we don’t have the identifying demographic groups that Clapp and Ross (2004) had. For our paper, a strategy for allowing sorting is to allow sufficient lags after airport expansion for those valuing this to bid up the price of housing benefiting from the expansion. We evaluate increasingly long intervals around the expansion event, as described above, to deal with the lag issue.

While we would also be able to control for any increase in airport noise using methods similar to earlier work by Cohen and Coughlin (2008), there are few houses in the noisy zones. Also, any heterogeneity due to noise can be captured through our individual-level Fixed Effects (FE). Since our model is based on the FE, there is also little point in trying to collect demographics at the CBG level. The reason for individual transactions is that houses within the CBG will differ in their access to the expanded airport. By lining all the transactions up around the expansion events (time zero is event date, regardless of calendar date), and including calendar year FE along with the individual level FE, we control for omitted variables other than the expansion. However, we do not have enough "events" in enough MSAs to do the statistical tests used in event studies in the finance literature. The most important explanatory variables are distance from the airport interacted with the amount and type of expansion. It may also be the case that some expansions don't increase congestion but only make the terminal facilities more attractive.

For our estimation of effects of infrastructure stocks on land values, we estimate the following model after obtaining extrapolated land values for each house in each year (see Appendix for description of the extrapolation approach):

$$L_{i,t} = c_0 + c_1 *A_{1,\,i,t} + c_2 *A_{2,\,i,t} + c_3 *A_{3,\,i,t} + c_4 *A_{4,\,i,t} + c_5 *A_{5,\,i,t} + \alpha_i + \tau_t + \varepsilon_{i,t} \qquad (12)$$

In this model, $L_{i,t}$ is log of land value, normalized by log of acres[16] for property i in year t; $A_{1,\,i,t}$ through $A_{5,\,i,t}$ represent airport infrastructure stocks for property i in year t for airfields, terminals, parking, roads/rails/transit, and "other", respectively; $A_{1,\,i,t}$ through $A_{5,\,i,t}$ are weighted by the distance from house i to the airport; α and τ are individual and time fixed effects, respectively; and $\varepsilon_{i,t}$ is an iid error term with mean zero, constant variance and zero covariance across observations; and for Denver, i=1, 2,…, 178,731; t=2003, 2004,…, 2010.

To compare the short-run versus long run impacts on land values of changes in airport infrastructure, we employ a year-over-year change approach, which leads to the following model:

$$\Delta_d L_{i,t} = c_1 *\Delta_d A_{1,\,i,t} + c_2 *\Delta_d A_{2,\,i,t} + c_3 *\Delta_d A_{3,\,i,t} + c_4 *\Delta_d A_{4,\,i,t} + c_5 *\Delta_d A_{5,\,i,t} + \theta + \Delta_d \varepsilon_{i,t} \qquad (13)$$

where Δ_d is the d-year-over-year change, d=1,2. When d=1, this represents the short-run impacts of changes in airport infrastructure on land values; d=2 represents the medium or long-term impacts. Note that examining year-over-year changes causes the cross-sectional fixed effects to drop out, and there are a new set of time-specific fixed effects, θ, which includes a constant (intercept) term.

[16] There is evidence that land values increase with the square root of lot size, so the fact that we are using logs is important since it prevents excess acreage from having the same effect on value as the building pad.

Denver Analysis

Figure 1 depicts housing prices in Denver in the years 2000 through 2012. During the period of our data sample for airport infrastructure investment (2003 through 2010), housing prices rose by about 8% in the boom years (2003 through 2007), while they fell by about 7.4% during the bust years (2007 through 2010). There were somewhat larger fluctuations in land prices, with a steadily decreasing land price from Davis and Palumbo (2008), as can be seen in Figure 2. Specifically, between 2003 and 2007, land prices in Denver fell by 11.5%, while during the years 2007 through 2010 land prices fell by 30.7%. This trend can be seen in Figure 2, which covers the broader period of 2000 through 2010. It is noteworthy that land prices rose dramatically between 2000 and 2003, before the steady subsequent decrease in land prices.

Descriptive statistics for the housing data are presented in Table 2 for Denver. There were over 326,000 sales observations for single family residential homes that sold between 2003 and 2010 in Denver. The average house in Denver had 3 bedrooms, with approximately 2 full baths and 0.18 half-baths. The average sale price was approximately $280,000, and was located about 19.3 miles from the airport. The closest house was 4.8 miles from the airport while the furthest house was 56 miles away. Figures 3 and 4 show the locations of the Denver home sales relative to the airport for the years 2003 and 2010, respectively.

The annual Denver International Airport capital stock data for 2003 through 2010 are listed in Table 1. As can be seen by examining the data for each category (airfields, terminals, parking, intermodal transportation, and other) in each year, there is variation in these capital stocks over time and across different categories. We used investment data to construct capital stocks for each of these categories of airport infrastructure, using the perpetual inventory method. Specifically, we deflated the investment series using a national deflator for government investment obtained from the 2013 Economic Report of the President, and the initial (or seed) value for the capital stock for each category is obtained as the average of the investment data for the years 2001 through 2004, multiplied by the estimated service life for each category of investment. The depreciation rate was assumed to be the inverse of the service life, and the capital stocks followed a straight line depreciation path. Additional details on the capital stock construction can be found in the data appendix. Once we constructed the capital stocks, our approach was to assign a capital stock value for each category to each single family residence sold, by weighting the capital stock by the property's inverse distance from the airport. Thus, properties more distant from the airport are viewed as having less airport capital.

We examine how land prices, obtained for each SFR sold in 2003-2010, are impacted by investment in airport infrastructure over time, for Denver and Atlanta. After implementing the LPR approach to obtain land values for each of these cities, and then interpolating to obtain a land value for each house in each year of our sample, we regress (for each city separately) the log of land values (normalized by log of acres) on a constant, on each of 5 categories of airport infrastructure capital stocks (airways, terminals, parking, roadways/railways, and other), as well as a set of cross-sectional and time fixed effects.[17] In these regressions, we normalize the capital stocks by each house's inverse distance to the airport (and a robustness check, we also normalize by inverse of distance-squared, which has little impact on the results). The distance is calculated as the Euclidean distance from the house to the airport using latitude and longitude data for each point.

Tables 4 and 5 present the second-stage regressions of the log of land values on the various infrastructure categories for Denver (weighted by the inverse of the distance from the airport). Our approach is to consider first- and second-differences. In the first-difference specification, the parameter estimates on the infrastructure variables are considered to be short-run effects.

Given the complex nature of the urban area southwest of the airport, we choose to focus this part of our analysis on the properties in the northwest quadrant of the airport. Figures 13-20 show the locations of the properties examined and their associated land values for 2003-2010. Relative to the entire sample, land values used for this part of our analysis tend to be less than in the other part of the Denver area. The southwest region is close to downtown Denver, and there are likely a broad variety of economic factors that can be expected to influence land values. Although there were 54,439 home sales between the years 2003-2010 in the northwest region, it is less developed than the southwest region and there are fewer other factors (such as other types of infrastructure, business activity, etc) that might be expected to impact land values. After interpolating the land values in all years for the houses sold in the northwest region, we have over 381,000 land value observations.

Table 4 presents the regression results with the one-year change in land values as the dependent variable, and the one-year changes in each infrastructure category as the independent variables. We also included year fixed effects in the model, however the cross-sectional fixed effects drop out when looking at the year-over-year changes. In the very short run, the coefficients for airfields, parking, intermodal transportation, and roads, rail, and transit are positive, while the coefficient for other infrastructure is negative. All variables reveal a statistically significant relationship with land prices. One interpretation of these results is that expenditures to improve or expand airports generates excitement among potential air travelers very quickly, while some of the "other" activities, such as de-icing equipment improvements, can lead to other externalities such as pollution runoff.

Support for such an interpretation is also contained in Table 5. The results for the two-year changes in land values regressed against the two-year changes in airport infrastructure categories, presented in Table 5, can be considered the longer-term effects. Identical to Table 4 each of the infrastructure variables is statistically significant. The key similarities are that the expenditures on airfields, terminals, parking, intermodal, are all positive, while the "other" category is still negative.[18]

[17] Figures 5-12 show quintiles for the interpolated land values in Denver in the years 2003-2010.
[18] As an alternative, the announcements of expansions, coupled with their values, might be an appropriate approach to organize the expansion data. We are unable to implement such an

Conclusion

We implement a Local Polynomial Regression framework to estimate land values for single family houses sold in Denver between 2003 and 2010. We interpolate the land values for every property in our sample in every year between 2003 and 2010, and validate the accuracy of the interpolated estimates with a within-sample forecasting approach. We then use these interpolated land value estimates to assess how various types of airport infrastructure investment affect land values. For the very short-run, we find that the coefficients for airfields, terminals, parking, intermodal transportation are positive, while the coefficient for "other" infrastructure is negative. Meanwhile, for the longer-term, we find similar relationships in terms of the signs and statistical significance of these relationships. One interpretation of these results is that some of the expenditures to improve or expand a terminal can produce results very quickly, depending on whether parts of the terminal are opened before renovations of the entire terminal are completed. On the other hand, building an entirely new terminal, expanding a parking facility or building a new runway can take years to begin and complete. In some cases, it takes time before the value of the expenditures is clear to market participants.

In future work, we aim to assess the robustness of these findings for airports in other cities. One such example is Atlanta, given its size and volume of operations.

The results demonstrate the importance of obtaining reliable estimates of land values, particularly during a period of a pronounced boom and bust. In a boom, the price of land is too high relative to its fundamental value, so eventually there is a bust, leading to a dramatic drop-off in land prices relative to one of the boom years. Although there were fluctuations in land prices in Denver, they remained plausible in the bust years, and we believe this is reflected in the results of our differencing analysis. Further research could explore additional data for such complex environments of sharply declining land values. Perhaps gathering additional data would be a more promising approach when the extra data spans over years when land prices were increasing or stable. Unfortunately the time period of our study (2003-2010) was limited by the availability of airport infrastructure data from the FAA.

---

alternative because the airport investment data we obtained from the Federal Aviation Administration does not distinguish between the announcement of expansions and the time of expenditures associated with the expansions.

References

Agostini, C.A and G.A. Palmucci. 2008. "The Anticipated Capitalisation Effect of a New Metro Line on Housing Prices," *Fiscal Studies* 29(2): 233-256.

Bourassa, S.C., Hoesli, M., Scognamiglio, D., Zhang, S., 2011. "Land leverage and house prices." *Regional Science and Urban Economics*. 41 (2): 134-144.

Bourke, P. 1999. "Interpolation Methods", http://paulbourke.net/miscellaneous/interpolation/ (accessed 5/9/2014).

Chapman, J.I., G. Cornia, R. Facer, L. Walters. 2009. "Alternative Financing Models for Transportation: A Case Study of Land Taxation in Utah," *Public Works Management & Policy* 13(3): 202-214.

Clapp, J. M. 2004. "A Semiparametric Method for Estimating Local House Price Indices." *Real Estate Economics* 32(1): 127-160.

Clapp, J.M. and S.L. Ross. 2004. "Schools and Housing Markets: An Examination of School Segregation and Performance in Connecticut," *Economic Journal* 114: F425-F440.

Clapp, J.M., Salavei, K., 2010. "Hedonic Pricing with Redevelopment Options: A New Approach to Estimating Depreciation Effects." *Journal of Urban Economics*. 67(3): 362-377.

Cohen, J.P., C. C. Coughlin, D. A. Lopez. 2012. The Boom and Bust of U.S. Housing Prices from Various Geographic Perspectives, *Federal Reserve Bank of St. Louis Review* 94(5): 341-367.

Cohen, J.P. 2012. "Commentary: Airport Improvement Fees, Benefit Spillovers, and Land Value Capture Mechanisms", in Ingram, G. and Y.H. Hong, eds., *Value Capture and Land Policies*, Lincoln Institute of Land Policy.

Cohen, J.P. and C.C. Coughlin. 2008. "Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices" *Journal of Regional Science* 48(5): 859-878.

Cohen, J.P. and C.J. Morrison Paul. 2003. "Airport Infrastructure Spillovers in a Network System," *Journal of Urban Economics*, 54(3): 459-473.

Davis, M. and M.G. Palumbo. 2008. "The price of residential land in large US cities," *Journal of Urban Economics*, 63, 352-384, data located at Land and Property Values in the U.S., Lincoln Institute of Land Policy, http://www.lincolninst.edu/resources/

Forrest, D., J. Glen and R. Ward. 1996. “The Impact of a Light Rail System on the Structure of House Prices” *Journal of Transport Economics and Policy*, 15-29.

Gatzlaff, D.H. and M.T. Smith. 1993. “The Impact of the Miami Metrorail on the Value of Residences near Station Locations” *Land Economics* 69(1): 54-66.

Hastie, T. J., Tibshirani, R. J. 1990. “Generalized Additive Models,” Chapman & Hall/CRC, 1990.

Hendriks, Dave (2005). “Apportionment in Property Valuation: Should We Separate the Inserparable?” *Journal of Property Investment & Finance*, 23(5), 455-470.

Ingram, G. and Y.H. Hong, eds., 2012. *Value Capture and Land Policies*, Cambridge: Lincoln Institute of Land Policy.

Jud, G.D. and D.T. Winkler. 2006. “The Announcement Effect of an Airport Expansion on Housing Prices” *Journal of Real Estate Finance and Economics* 33: 91-103.

Longhofer, S. and C. Redfearn. 2009. “Estimating Land Values Using Residential Sales Data.” Lincoln Institute of Land Policy Working Paper.

McMillen, D.P. and J,F. McDonald. 2004. “Reaction of House Prices to a New Rapid Transit Line: Chicago’s Midway Line, 1983-1999” *Real Estate Economics* 32(3): 463-486.

Ozdilek, Unsal (2012) An Overview of the Enquiries on the Issue of Apportionment of Value between Land and Improvements” *Journal of Property Research* 29(1), 69-84.

Pavlov, A.D. 2000, “Space-Varying Regression Coefficients: A Semi-parametric Approach Applied to Real Estate Markets,” *Real Estate Economics*, V.28 (2), 249-283.

Ruppert, D., Wand, M.P., Holst, U. and Hossjer, O. 1997. “Local Polynomial Variance Function-Estimation,” *Technometrics*, V. 39, No. 3, 262-273.

Wand, M. P., and Jones, M. C. 1995. “Kernel Smoothing,” Chapman and Hall.

Weinberger, R. 2000. “Commercial Rents and Transportation Improvements: The Case of Santa Clara County’s Light Rail.” Lincoln Institute of Land Policy Working Paper.

Data Appendix

-Capital stocks:

We use the perpetual inventory method, along with annual data on new airport investments in several different categories (airfields; terminals; parking; rail, road and transit; and “other”) to obtain separate estimates of capital stocks for each of these categories. We assume the depreciation rate = 1/service life, where service life of airport terminals and airfields = 25 years; service life of parking = 40 years; service life of roads/rail/transit = 44 years; service life of "other" = 25 years.

The 25 year number for airfields and terminals came from airports council international, used in Cohen and Morrison Paul (2003).

The highways and streets service life: 60 years (0.0152); state and local railroad equipment: 28 years (0.0590); For the roads, rail, and transit variable, we take the average of these two service lives and use 44 years for the service life. Source: http://www.bea.gov/scb/account_articles/national/0797fr/table3.htm:

Parking: http://www.chamberlinltd.com/extending-the-service-life-of-parking-structures-a-systematic-repair-approach/

Initial capital stocks are average of 2001, 2002, 2003, and 2004 expenditures, times the service life for that category

-Land price indexes:

We interpolated land values for all years for each house, using a method devised by Clapp (2004) and subsequently modified by Brett Fawley, Diana Cooke, and us. Details are available from the authors upon request.

Subsequently, we add back the values of the time dummy variables from the hedonic regressions, then deflated the land values by the CPI for Denver.

Figure 1 – Single Family Home Sales Prices, Denver

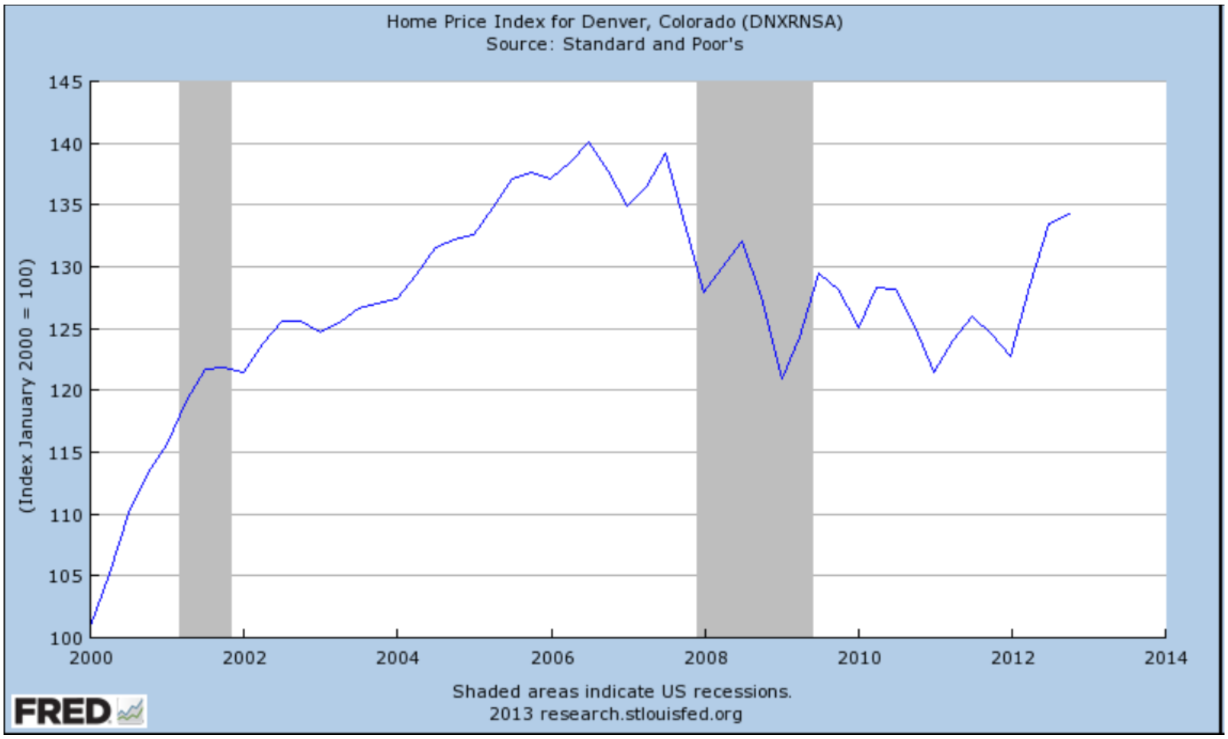


Figure 2

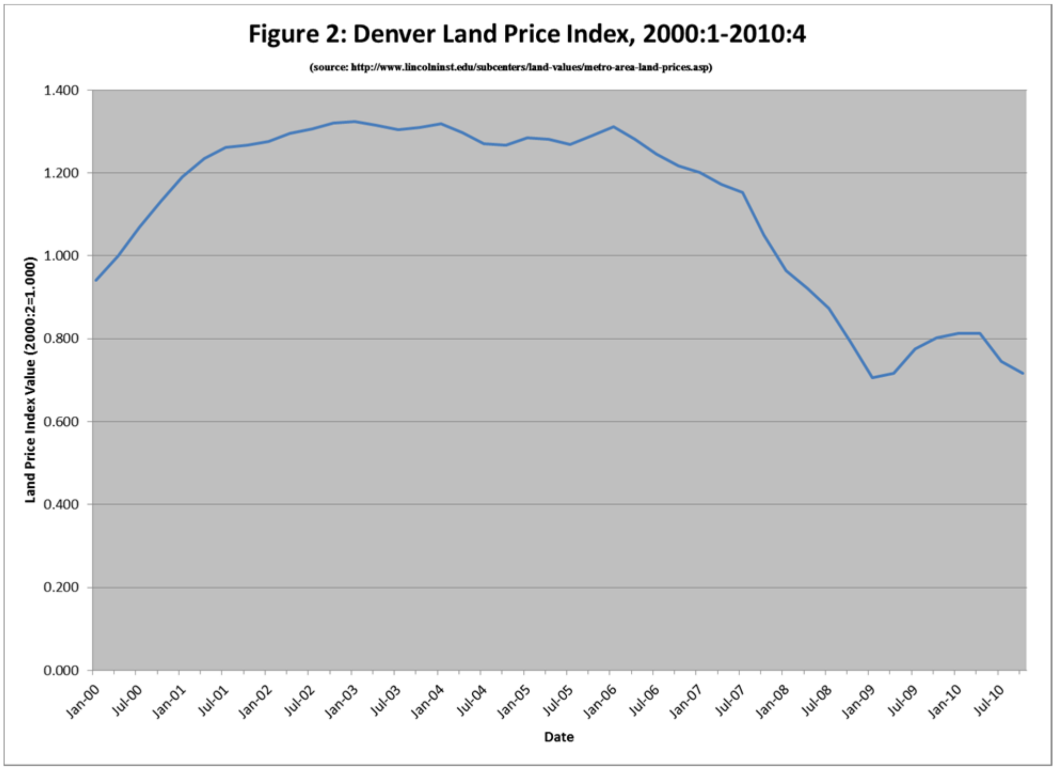


Figure 2: Denver Land Price Index, 2000:1-2010:4
(source: http://www.lincolninst.edu/subcenters/land-values/metro-area-land-prices.asp)
Land Price Index Value (2000:2=1.000)
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Jan-00
Jul-00
Jan-01
Jul-01
Jan-02
Jul-02
Jan-03
Jul-03
Jan-04
Jul-04
Jan-05
Jul-05
Jan-06
Jul-06
Jan-07
Jul-07
Jan-08
Jul-08
Jan-09
Jul-09
Jan-10
Jul-10
Date

Figure 3 – Prices of Single Family Home Sales in 2003 for Denver

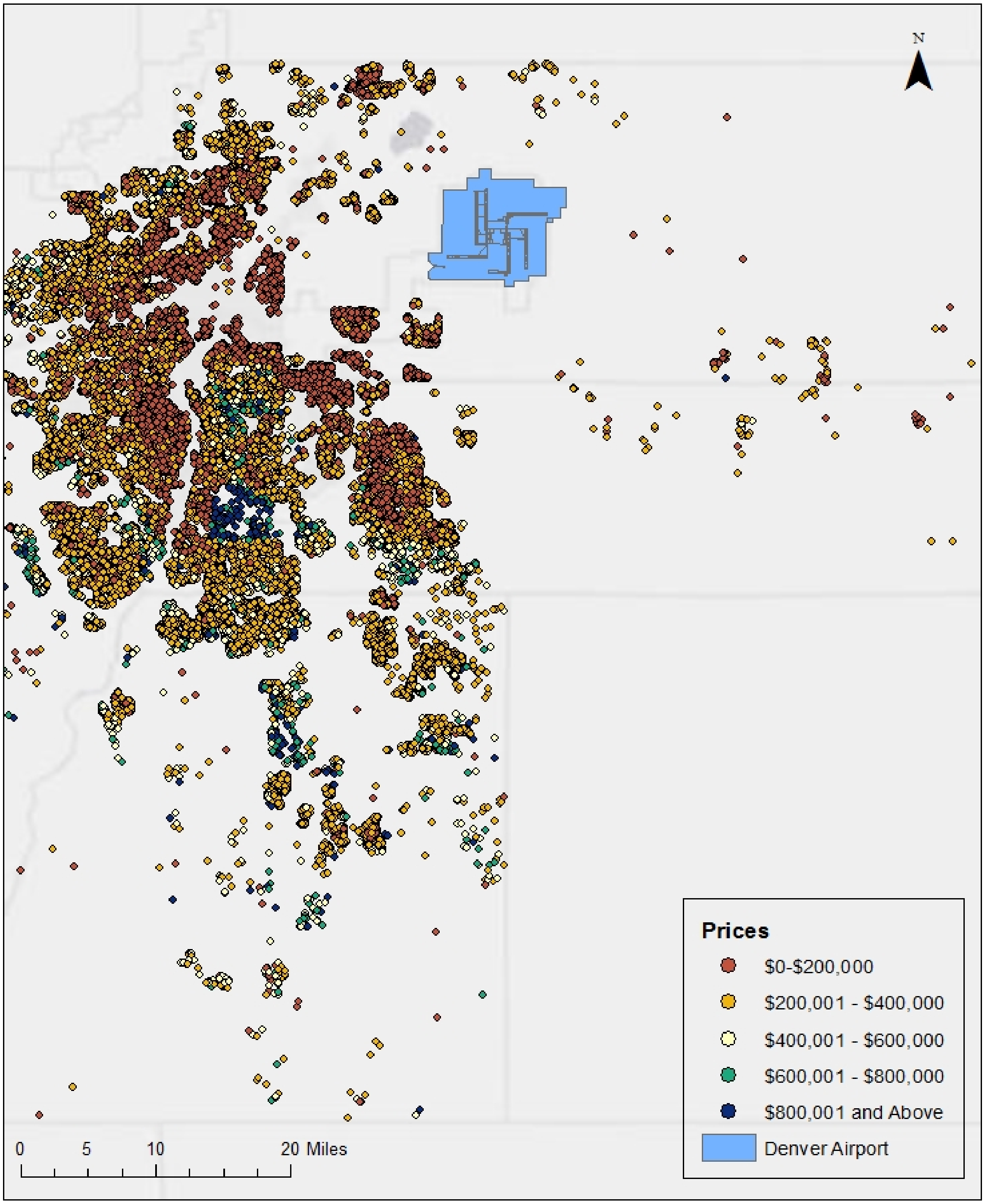


Figure 4 – Prices of Single Family Home Sales in 2010 for Denver

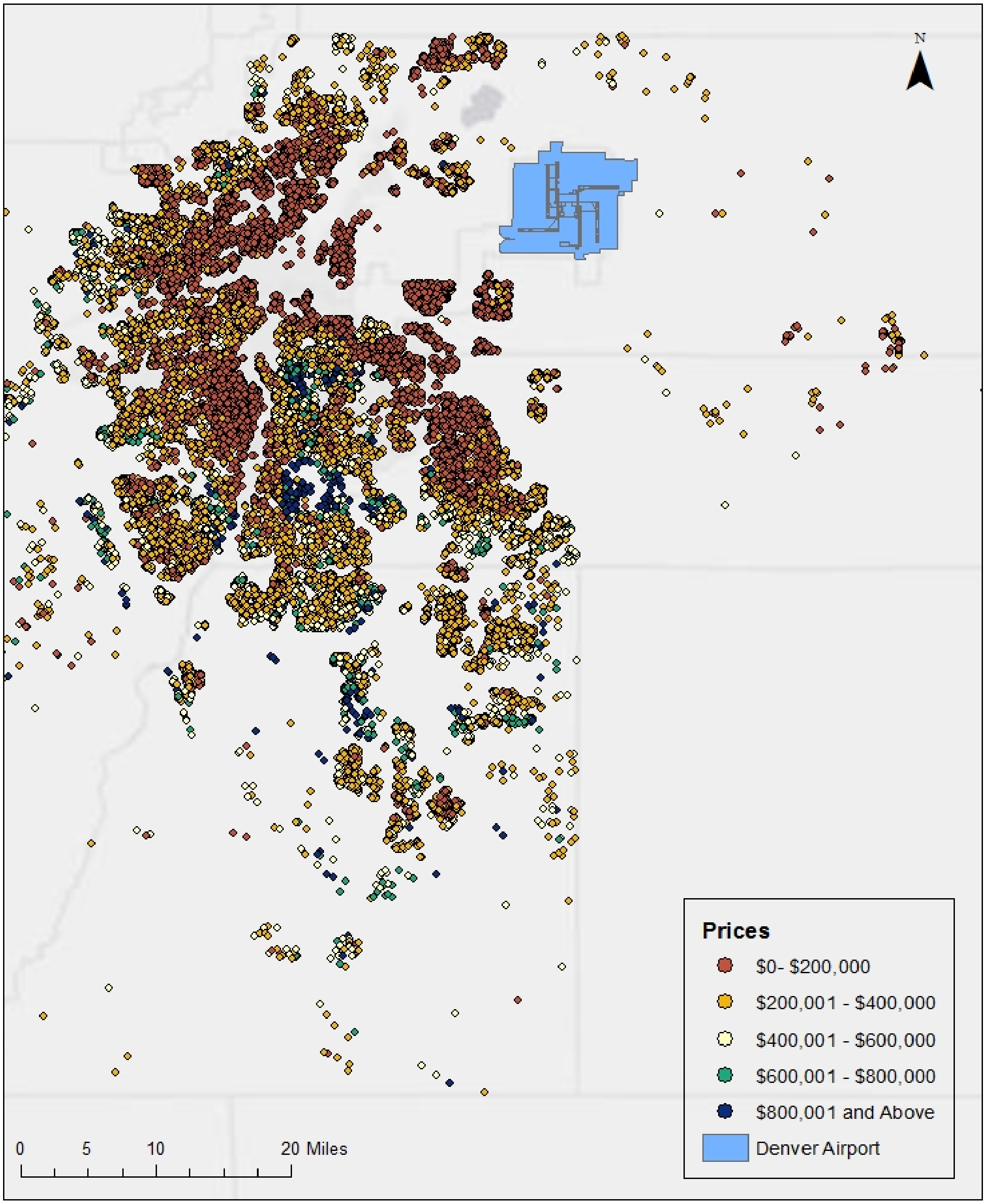


Figure 5 – 2003 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2003 are estimated directly by local polynomial regression. Values for homes not sold in 2003 use linear interpolations.)

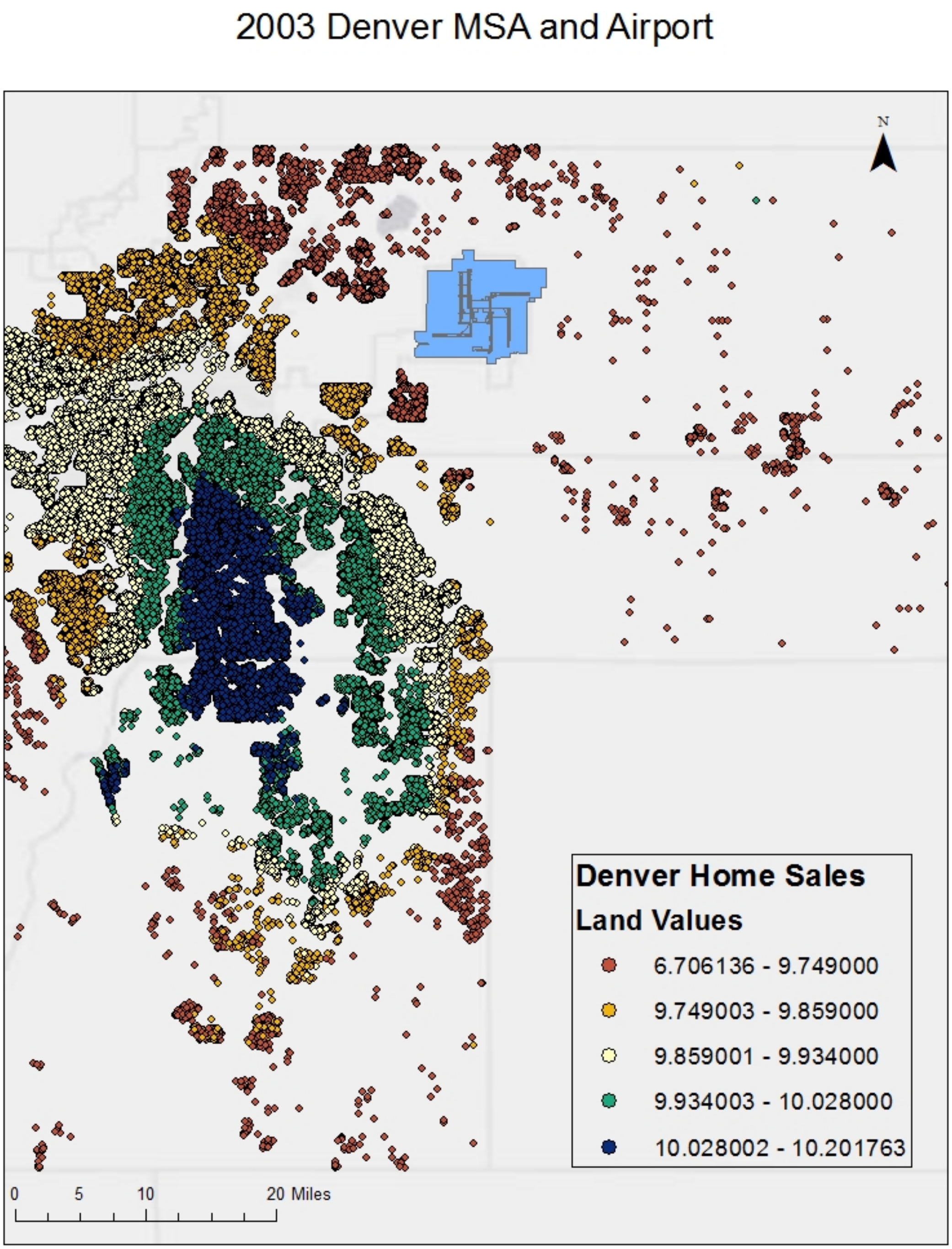


Figure 6 – 2004 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2004 are estimated directly by local polynomial regression. Values for homes not sold in 2004 use linear interpolations.)

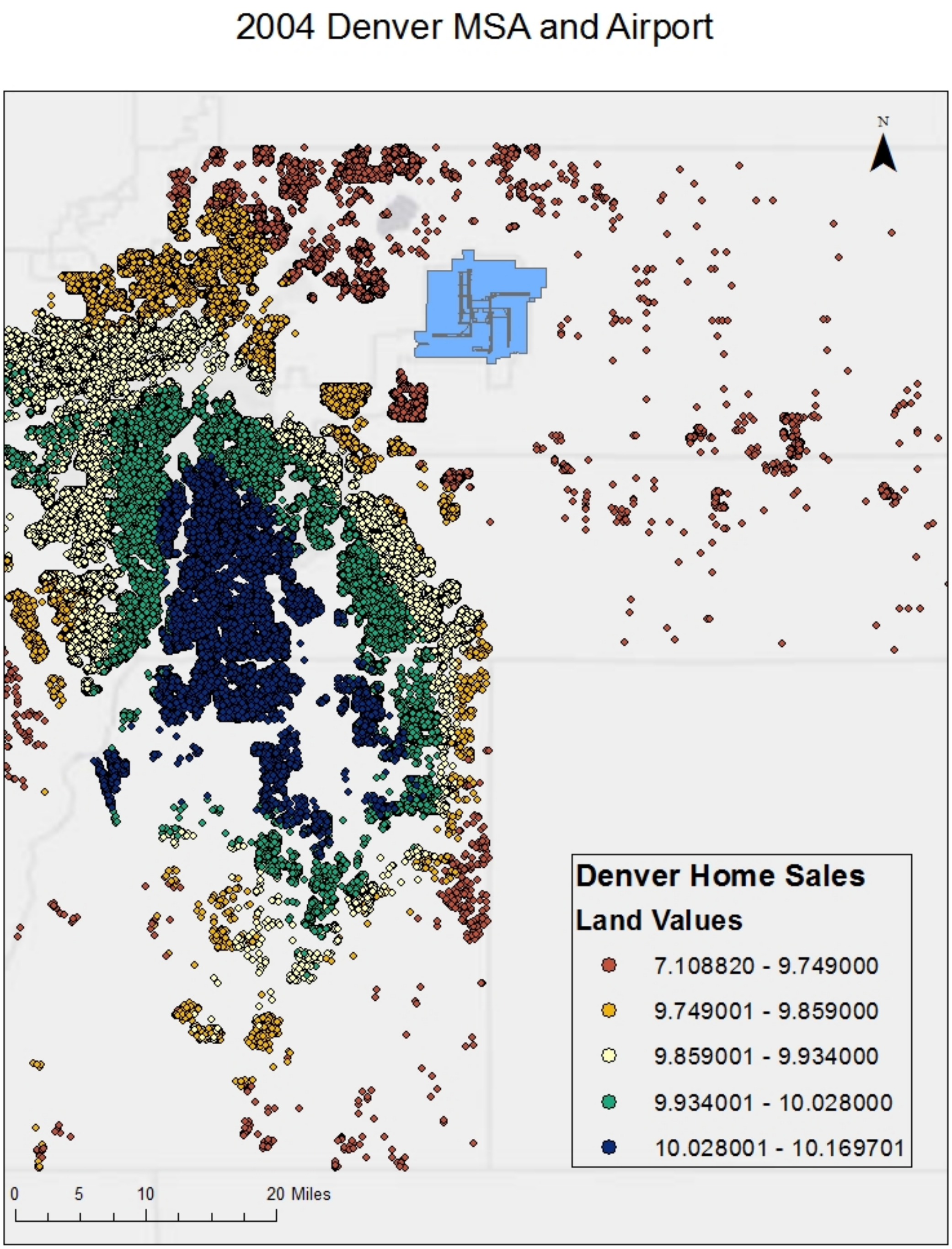


Figure 7 – 2005 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2005 are estimated directly by local polynomial regression. Values for homes not sold in 2005 use linear interpolations.)

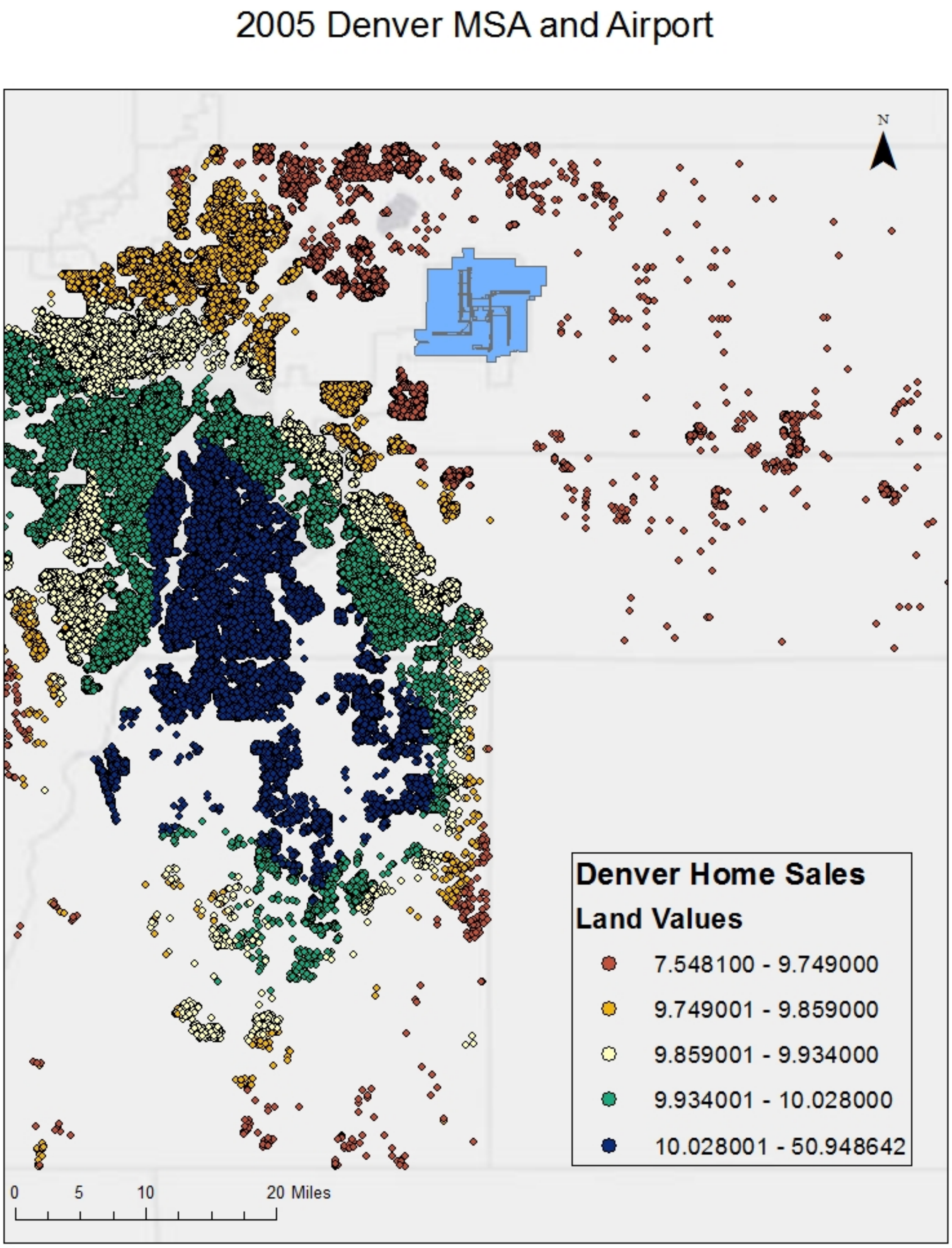


Figure 8 – 2006 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2006 are estimated directly by local polynomial regression. Values for homes not sold in 2006 use linear interpolations.)

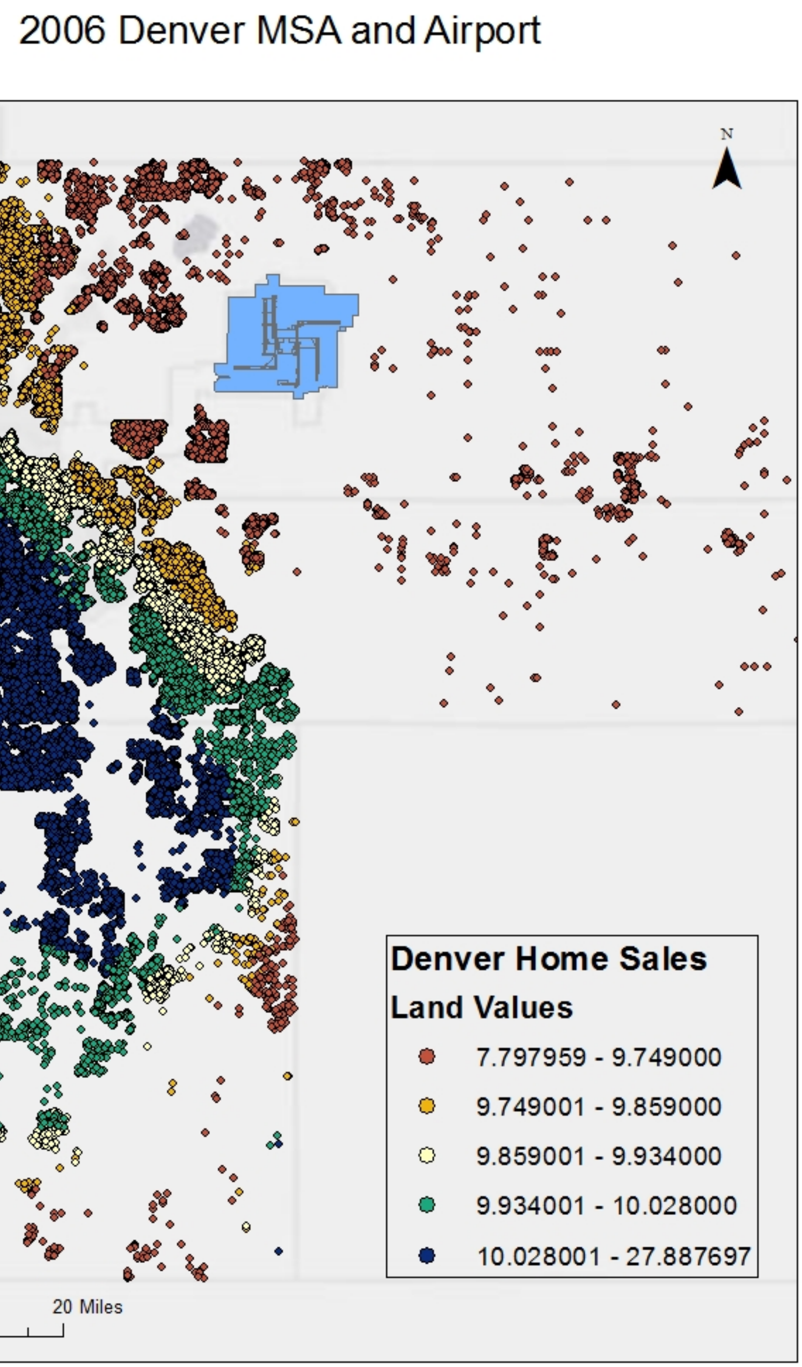


Figure 9 – 2007 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2007 are estimated directly by local polynomial regression. Values for homes not sold in 2007 use linear interpolations.)

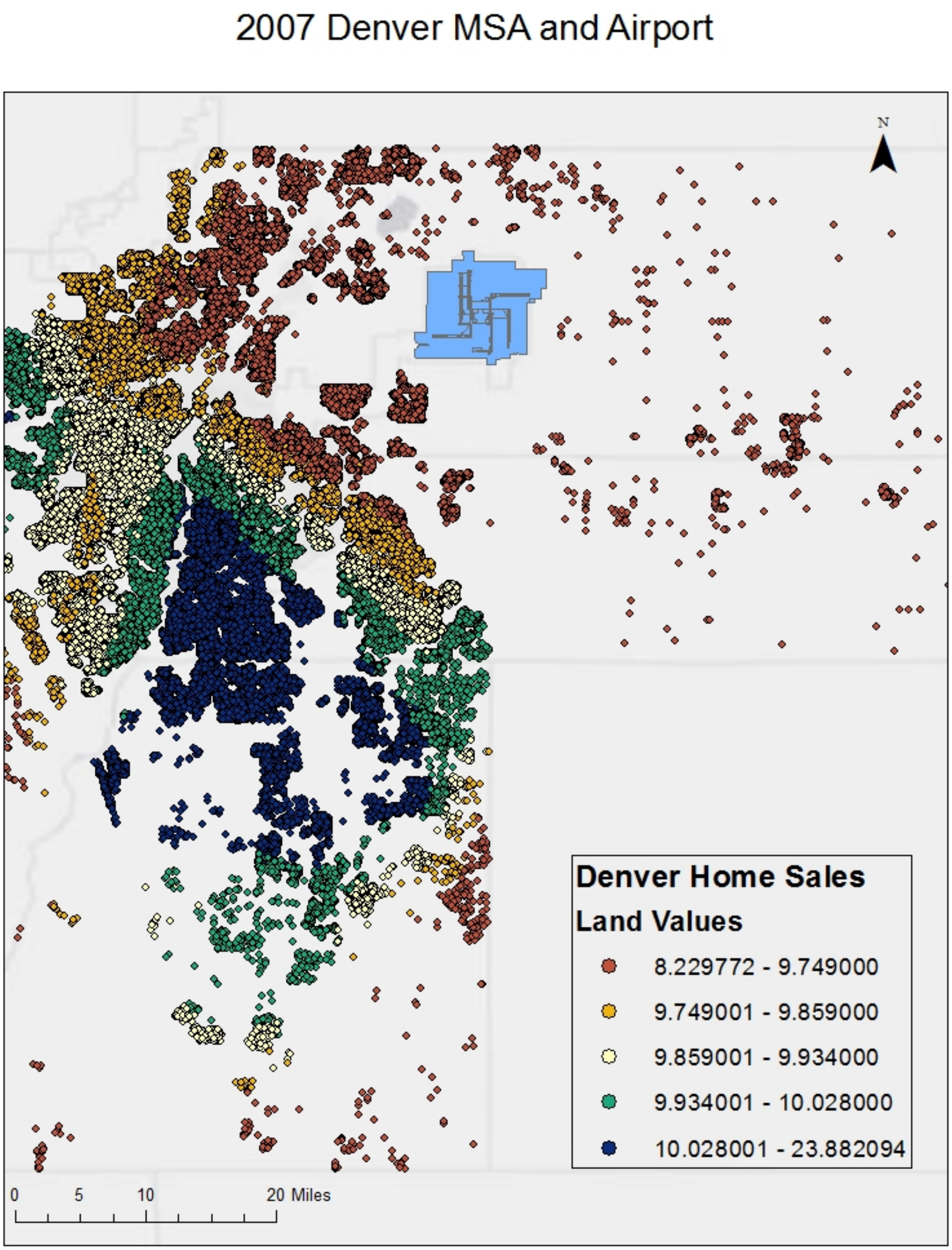


Figure 10 – 2008 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2008 are estimated directly by local polynomial regression. Values for homes not sold in 2008 use linear interpolations.)

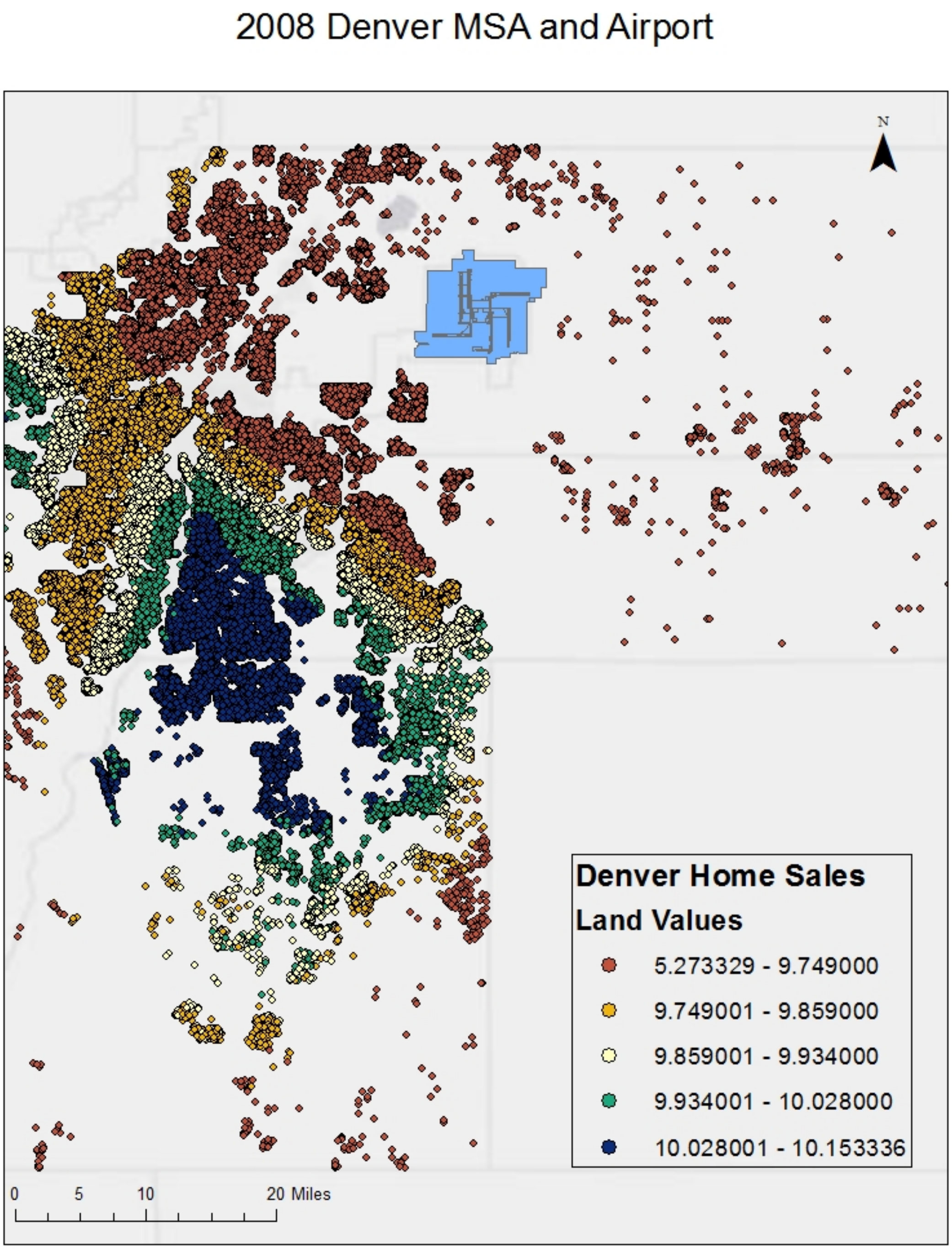


Figure 11 – 2009 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2009 are estimated directly by local polynomial regression. Values for homes not sold in 2009 use linear interpolations.)

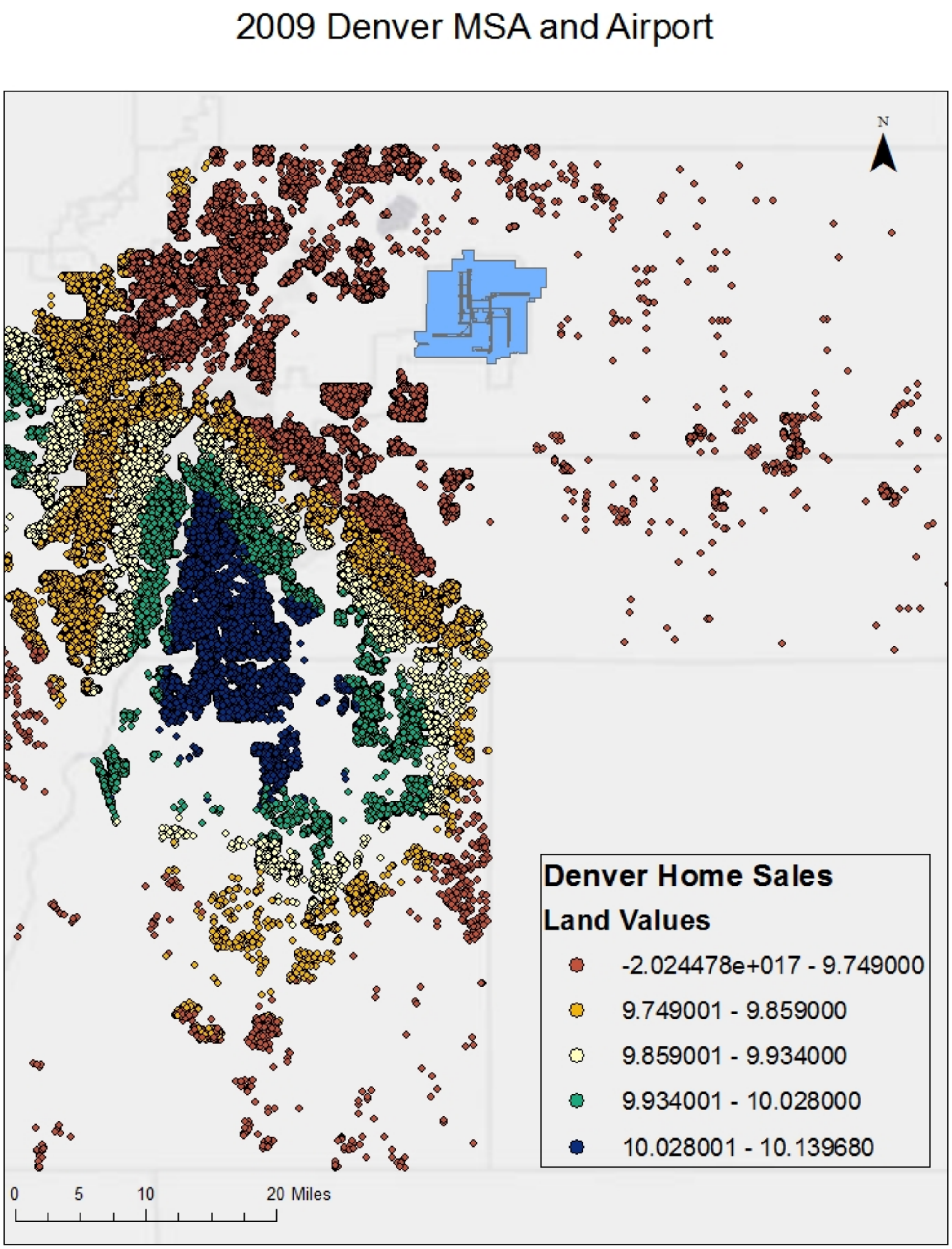


Figure 12 – 2010 Land Values for Homes Sold Between 2003 and 2010, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2010 are estimated directly by local polynomial regression. Values for homes not sold in 2010 use linear interpolations.)

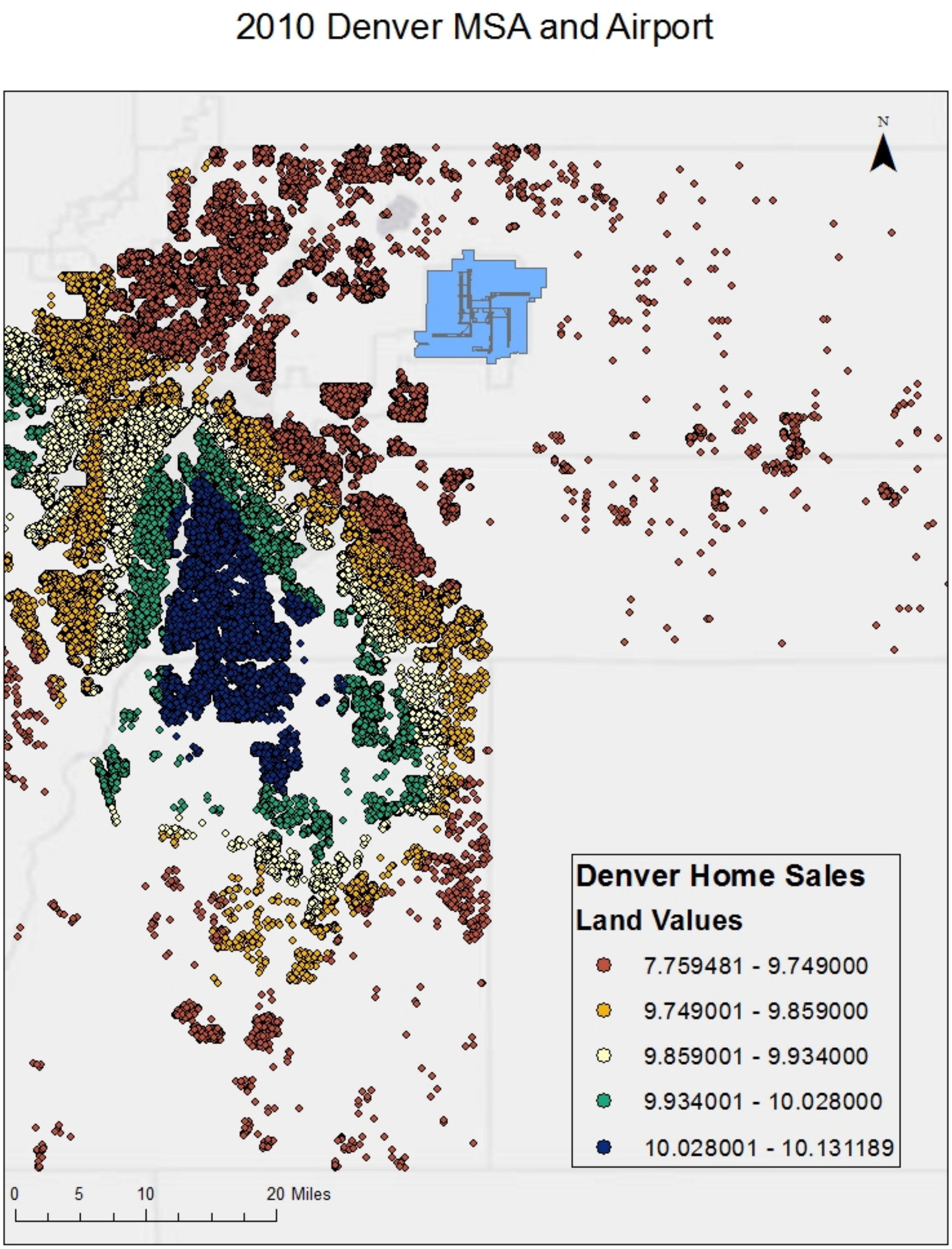


Figure 13 – 2003 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2003 are estimated directly by local polynomial regression. Values for homes not sold in 2003 use linear interpolations.)

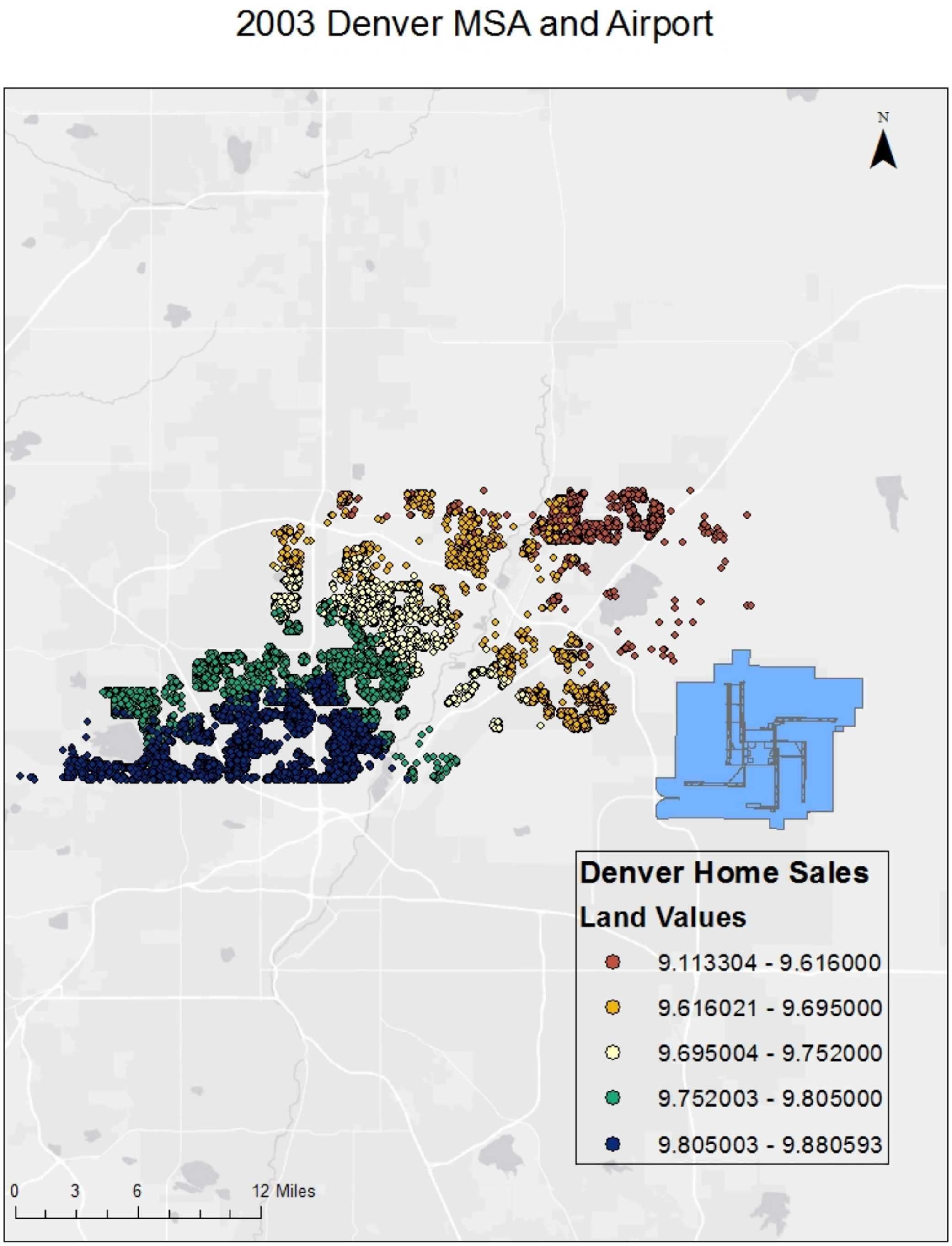


Figure 14 – 2004 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver
(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2004 are estimated directly by local polynomial regression. Values for homes not sold in 2004 use linear interpolations.)

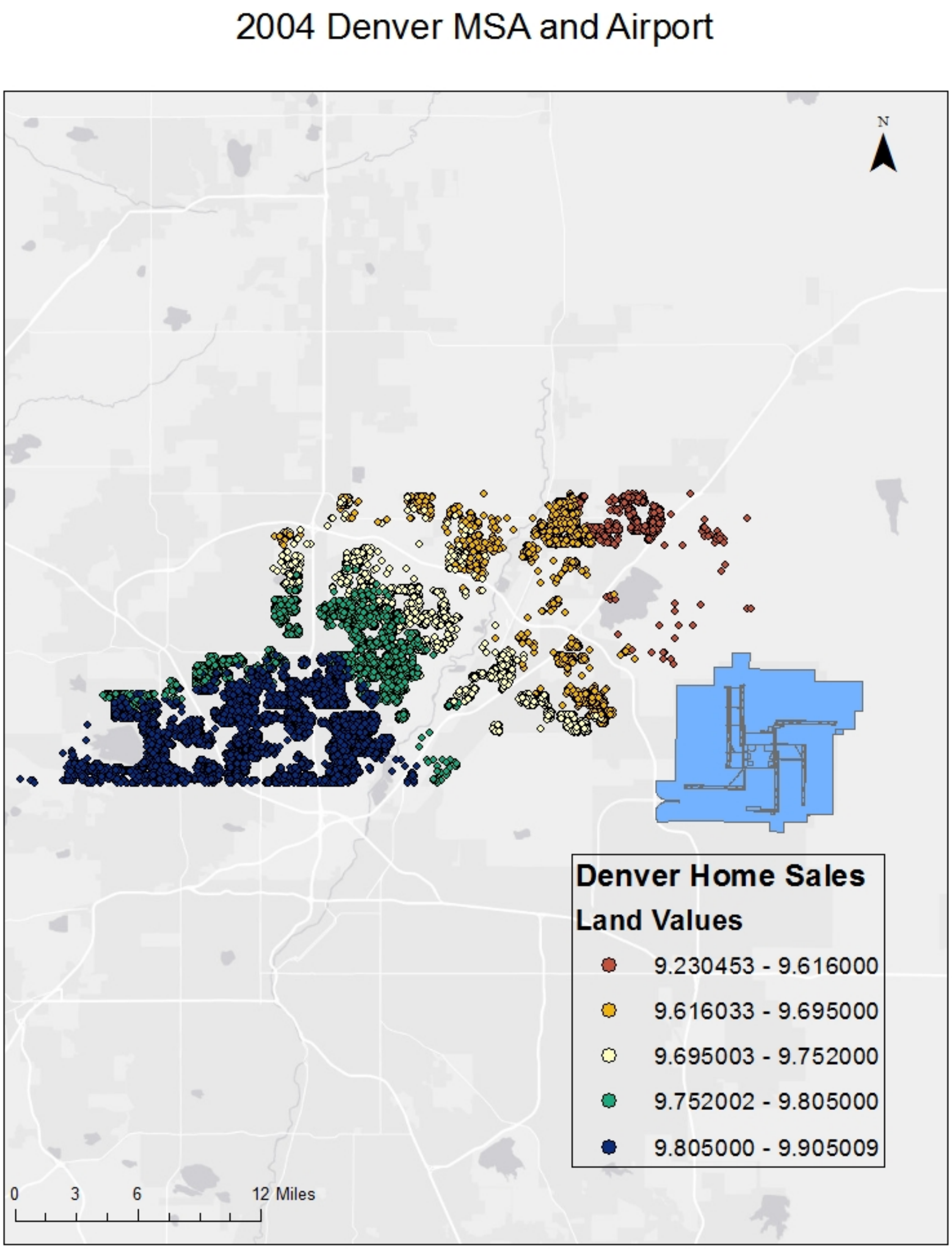


Figure 15 – 2005 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2005 are estimated directly by local polynomial regression. Values for homes not sold in 2005 use linear interpolations.)

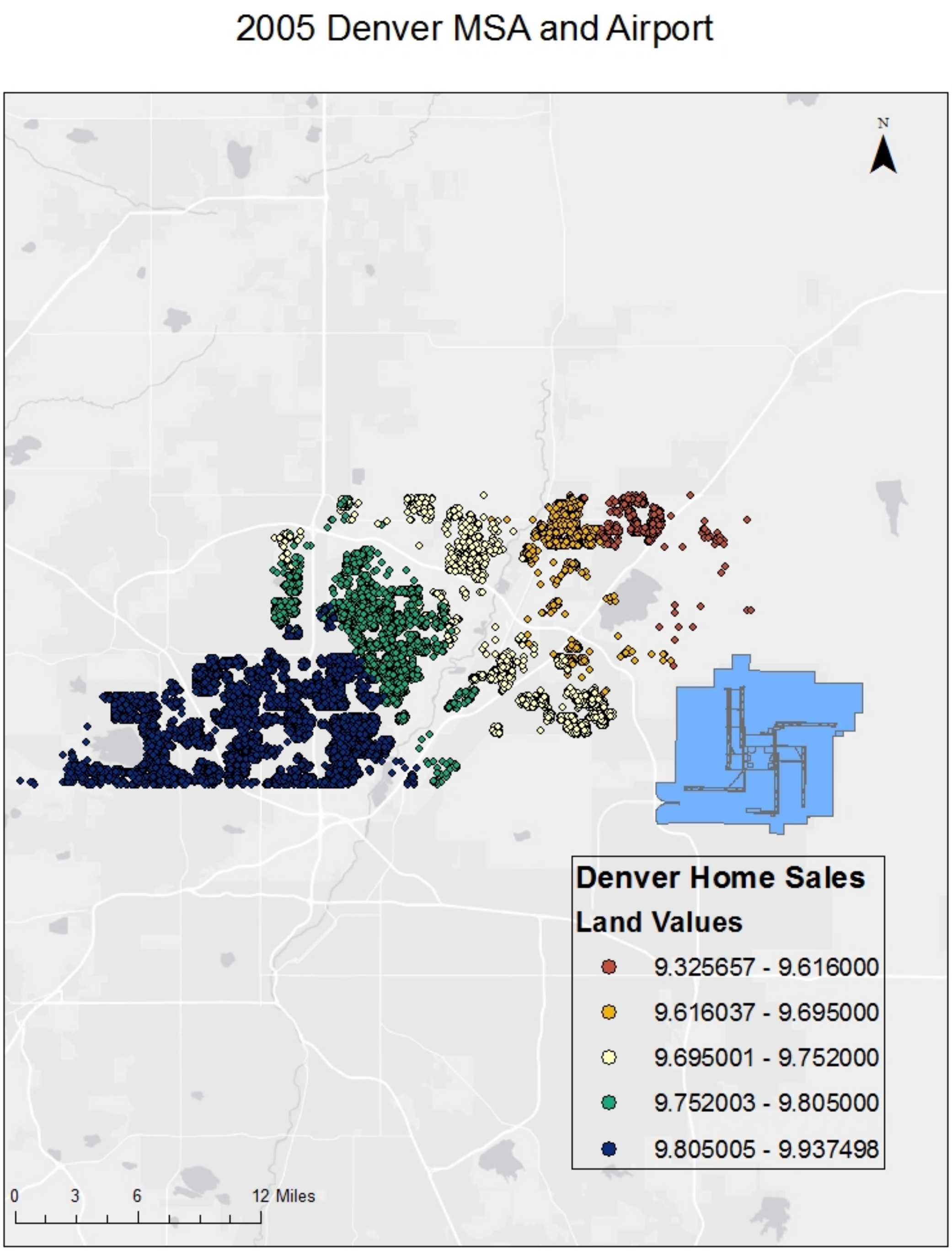


Figure 16 – 2006 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2006 are estimated directly by local polynomial regression. Values for homes not sold in 2006 use linear interpolations.)

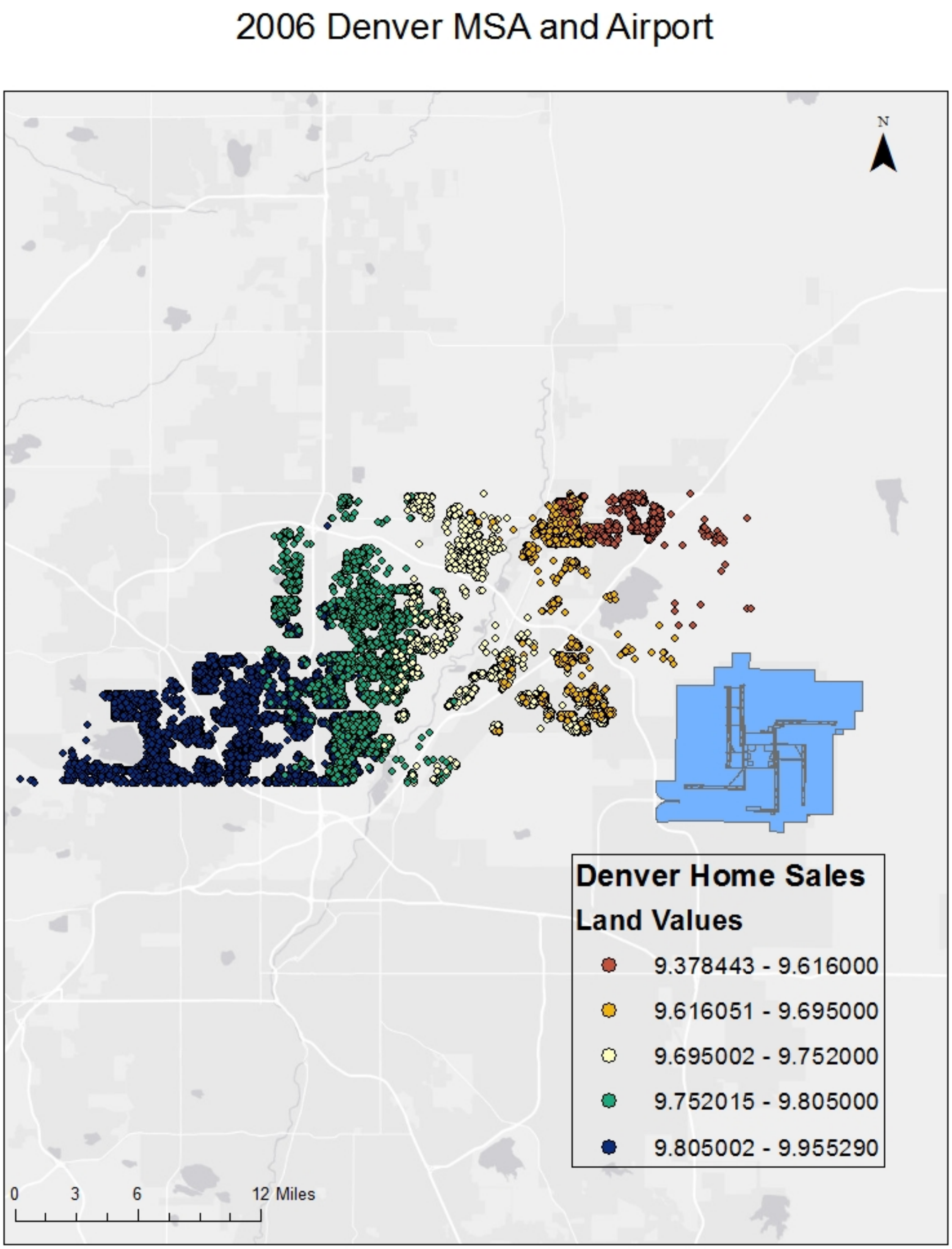


Figure 17 – 2007 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2007 are estimated directly by local polynomial regression. Values for homes not sold in 2007 use linear interpolations.)

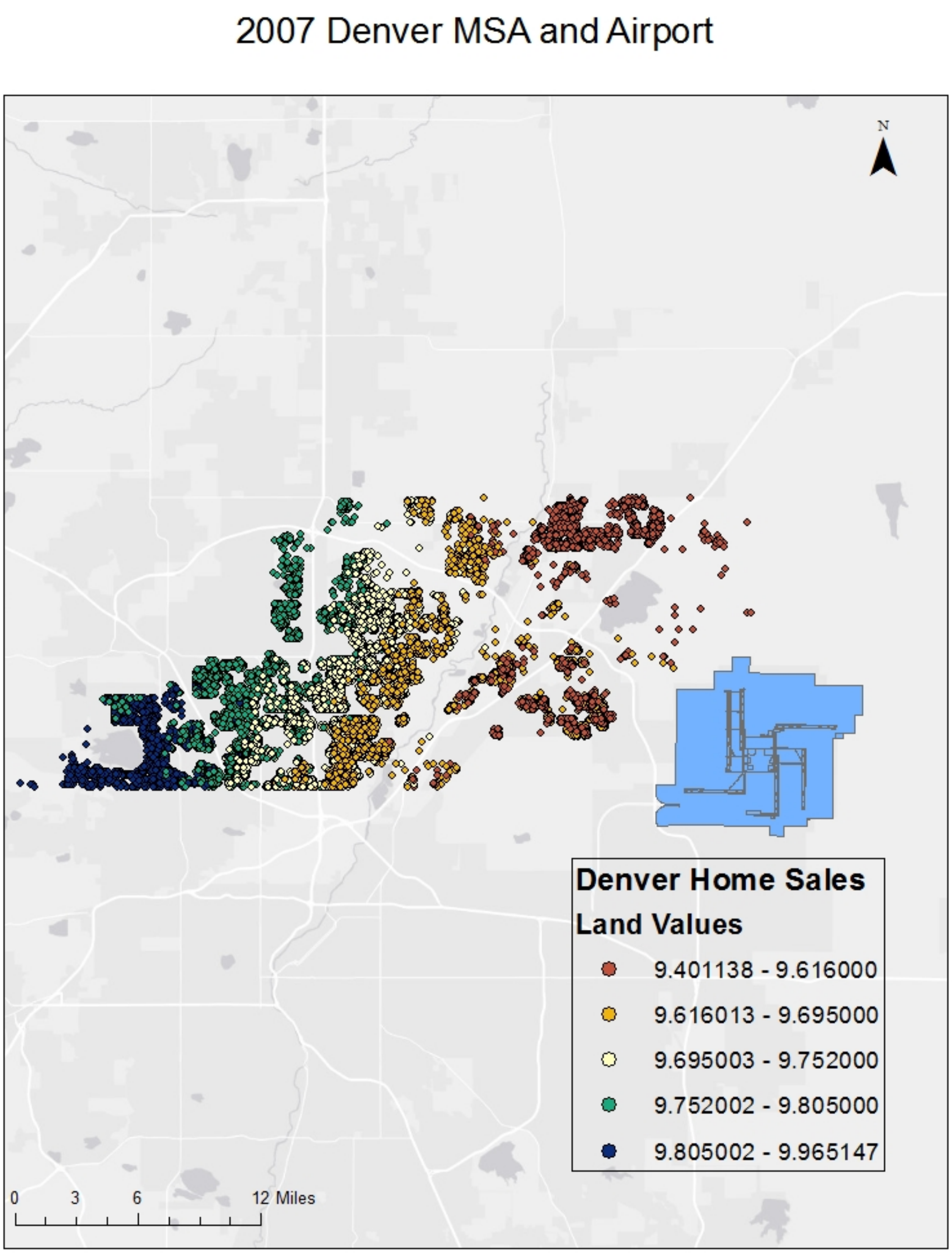


Figure 18 – 2008 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2008 are estimated directly by local polynomial regression. Values for homes not sold in 2008 use linear interpolations.)

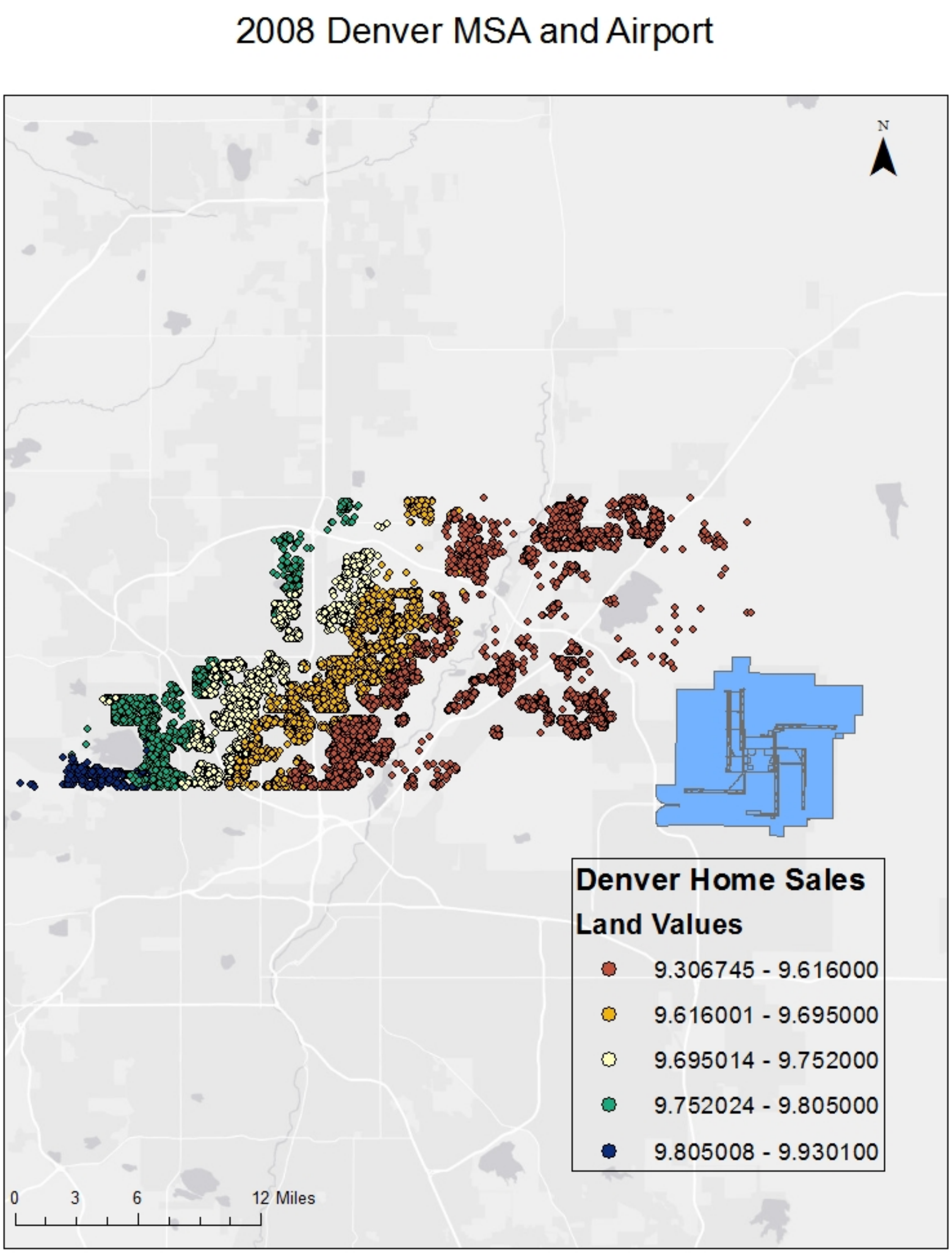


Figure 19 – 2009 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2009 are estimated directly by local polynomial regression. Values for homes not sold in 2009 use linear interpolations.)

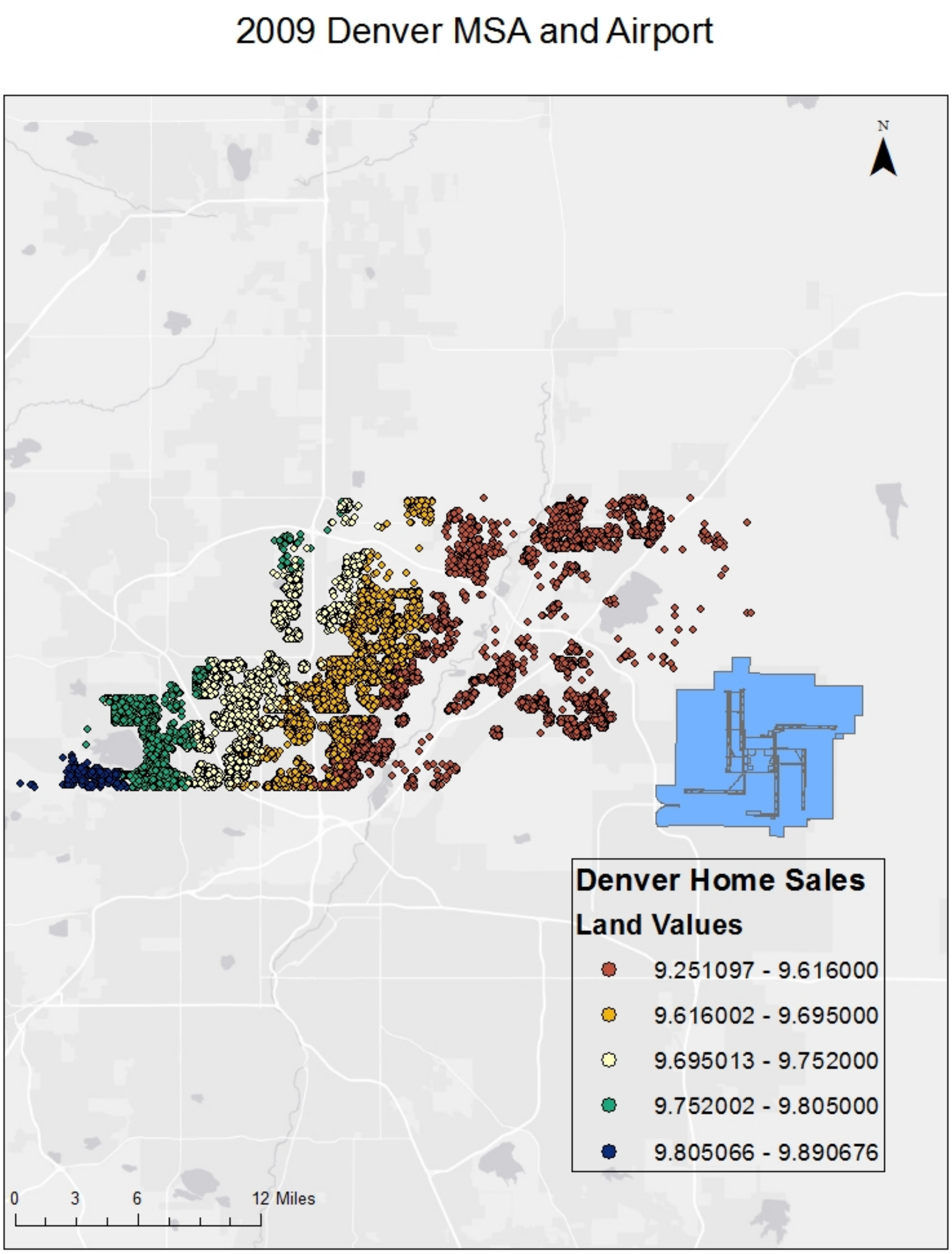
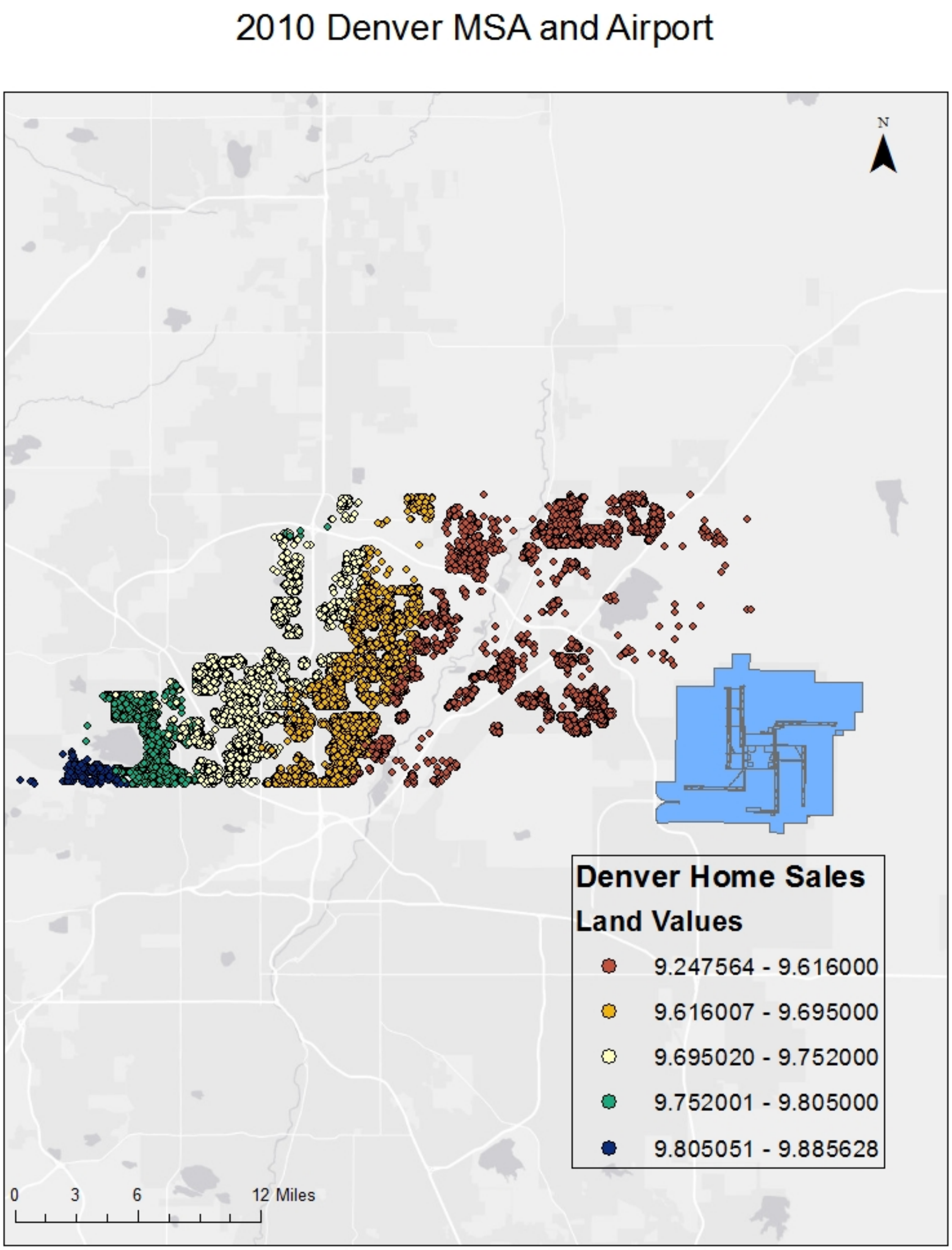


Figure 20 – 2010 Land Values for Homes Sold Between 2003 and 2010 with Latitude > 39.8439 and Longitude < -104.6733, Denver

(Note: Land Values are expressed in Natural Logs. Values for homes sold in 2010 are estimated directly by local polynomial regression. Values for homes not sold in 2010 use linear interpolations.)



‘

Table 1: Airport Infrastructure Capital Stocks, Denver International Airport, 2003-2010 (millions of dollars)
Note: capital stock estimates are in constant (2003) millions of dollars, net of depreciation.

	2003	2004	2005	2006	2007	2008	2009	2010
Airfield	759.3	735.2	710.6	695.0	678.8	664.0	675.6	670.2
Terminal	592.1	581.4	570.7	591.8	596.6	601.3	590.9	589.1
Parking	88.9	88.4	87.3	96.5	131.2	136.0	139.6	138.9
Road, Rail & Transit	104.4	107.1	111.9	112.3	110.4	113.0	111.7	114.3
Other	385.2	379.1	367.8	358.9	348.3	350.6	345.1	349.6
Total	1929.8	1891.3	1848.4	1854.4	1865.3	1865.0	1863.0	1862.1

Table 2: Descriptive Statistics, Denver Single Family Home Sales, 2003-2010

Variable	Mean	Std. Dev.	Variance	Minimum	Maximum	Valid
Sale Price (Log)	12.4315	0.5486	0.3010	6.9078	15.4249	326,744
Yr2003	0.1393	0.3463	0.1199	0	1	326,744
Yr2004	0.1530	0.3600	0.1296	0	1	326,744
Yr2005	0.1528	0.3598	0.1295	0	1	326,744
Yr2006	0.1384	0.3453	0.1192	0	1	326,744
Yr2007	0.1232	0.3286	0.1080	0	1	326,744
Yr2008	0.1112	0.3144	0.0989	0	1	326,744
Yr2009	0.0955	0.2940	0.0864	0	1	326,744
Yr2010	0.0865	0.2811	0.0790	0	1	326,744
No of Bedrooms	3.1587	0.8403	0.7061	1	13	326,744
No of Full Baths	2.2924	0.8853	0.7838	1	12	326,744
No of Half Baths	0.3259	0.4938	0.2439	0	5	326,744
No of Fireplaces	0.7669	0.7313	0.5348	0	10	326,744
Garage Dummy	0.9103	0.2858	0.0817	0	1	326,744
Basement Dummy	0.8031	0.3977	0.1581	0	1	326,744
Stories Dummy	0.4839	0.4997	0.2497	0	1	326,744
Adams County Dummy	0.1900	0.3923	0.1539	0	1	326,744
Denver County Dummy	0.2268	0.4188	0.1754	0	1	326,744
Douglas County Dummy	0.1782	0.3827	0.1465	0	1	326,744
Araphoe County Dummy	0.2114	0.4083	0.1667	0	1	326,744
Jefferson County Dummy	0.1935	0.3951	0.1561	0	1	326,744
Longitude	-104.9384	0.1441	0.0208	-105.4648	-103.765	326,744

Latitude	39.6936	0.1440	0.0207	39.1305	40.242	326,744
Longitude Squared	11012.0871	30.2435	914.6665	10767.1109	11122.82	326,744
Latitude Squared	1575.6048	11.4253	130.5373	1531.1990	1619.42	326,744
Lat*Lon	-4165.3879	16.7768	281.4625	-4206.4213	-4098.49	326,744
Age	34.0694	27.3488	747.9552	0	145	326,744
Age Squared	1908.6075	2870.4277	8238049	0	21025	326,744
Land Square Feet (Log)	9.0213	0.6916	0.4765	6.2146	18.1084	326,744

Table 3 – Hedonic Regressions, Denver SFR Home Sales, 2003-2010

Valid cases:	326744	Dependent variable:	Sales Price (Log)
Missing cases:	0	Deletion method:	None
Total SS:	98336.134	Degrees of freedom:	326717
R-squared:	0.591	Rbar-squared:	0.591
Residual SS:	40228.422	Std error of est:	0.351
F(24,178706):	18150.942	Probability of F:	0

Variable	Coefficient Estimate	Std. Error	T-Value	P-Value	Std. Estimate	Corr. With Dep Var
Constant	-4233.666469	312.963735	-13.527658	0.00	---	---
Yr2003	0.056438	0.002659	21.222047	0.00	0.035622	0.000693
Yr2004	0.085489	0.002612	32.724294	0.00	0.056102	0.030246
Yr2005	0.121738	0.002614	46.576639	0.00	0.079846	0.069579
Yr2006	0.111406	0.002662	41.845352	0.00	0.070121	0.060327
Yr2007	0.065053	0.002724	23.879809	0.00	0.038971	0.018106
Yr2008	-0.051856	0.002784	-18.623067	0.00	-0.029721	-0.094189
Yr2009	-0.041102	0.002882	-14.261565	0.00	-0.022025	-0.081952
No of Bedrooms	0.014671	0.000925	15.866329	0.00	0.022472	0.336672
No of Full Baths	0.164688	0.001078	152.781209	0.00	0.265771	0.578731
No of Half Baths	0.151836	0.001667	91.099069	0.00	0.136678	0.246221
No of Fireplaces	0.170455	0.001058	161.099640	0.00	0.227216	0.545866
Garage Dummy	0.170616	0.002378	71.760975	0.00	0.088873	0.299812
Basement Dummy	0.138199	0.001707	80.941398	0.00	0.100176	0.354783
Stories Dummy	0.014131	0.001730	8.167266	0.00	0.012872	0.356637
Adams County Dummy	0.003398	0.003325	1.021937	.307	0.002430	-0.2141020
Denver County Dummy	0.183828	0.002970	61.887987	0.00	0.140325	-0.091960
Douglas County Dummy	0.016086	0.004209	3.821778	0.00	0.011221	0.241680
Araphoe County Dummy	-0.021698	0.003374	-6.430422	0.00	-0.016150	-0.009083

Longitude	-107.719767	4.971578	-21.667117	0.00	-28.295438	-0.064119
Latitude	-72.155186	4.580548	-15.752523	0.00	-18.935222	-0.291193
Longitude Squared	-0.799777	0.021529	-37.149268	0.00	-44.090696	0.064119
Latitude Squared	-1.115896	0.033850	-32.965914	0.00	-23.240076	-0.291193
Lat*Lon	-1.529108	0.036917	-41.419968	0.00	-46.762224	0.240276
Age	-0.014724	0.000091	-160.94147	0.00	-0.734021	-0.304837
Age Squared	0.000122	0.000001	156.981420	0.00	0.639077	-0.197236
Land Sq Feet (Log)	0.183304	0.001149	159.507617	0.00	0.231085	0.329327

Table 4 – Regression of one-year change of Land Value on airport infrastructure capital stocks (normalized by distance to the airport), Denver International Airport

Note: land values are in real terms

Dependent Variable: One-Year Change Land Level
Method: Panel Least Squares

Sample (adjusted): 2003 2010 if latitude>39.8439 and longitude <-104.6733
Periods included: 7
Cross-sections included: 54439
Total panel (balanced) observations: 381073

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2087.111	1.028625	-2029.030	0.0000
One-Year Change Airfield	0.000131	2.86E-06	45.98280	0.0000
One-Year Change Terminal	0.000362	1.69E-06	214.2468	0.0000
One-Year Change Parking	0.000116	2.19E-06	53.25792	0.0000
One-Year Change Rd, Rl, Tr	0.001748	1.62E-05	107.9825	0.0000
One-Year Change Other	-0.000254	3.23E-06	-78.58469	0.0000

Effects Specification

Period fixed (dummy variables)

R-squared	0.983675	Mean dependent var	-1887.896
Adjusted R-squared	0.983675	S.D. dependent var	1916.389
S.E. of regression	244.8578	Akaike info criterion	13.83926
Sum squared resid	2.28E+10	Schwarz criterion	13.83961
Log likelihood	-2636873	Hannan-Quinn criter.	13.83936
F-statistic	2087400	Durbin-Watson stat	0.831727
Prob(F-statistic)	0.000000		

Table 5 – Regression of two-year change of Land Value on airport infrastructure capital stocks (normalized by distance to the airport), Denver International Airport

Note: land values are in real terms

Dependent Variable: Two-Year Change Land Level
Method: Panel Least Squares

Sample: 2003 2010 IF LATITUDE>39.8439 AND LONGITUDE<-104.6733
Periods included: 6
Cross-sections included: 54439
Total panel (balanced) observations: 326634

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4895.050	2.295313	-2132.629	0.0000
Two-Year Change Airfield	0.000152	4.54E-06	33.45443	0.0000
Two-Year Change Terminal	0.000368	4.94E-06	74.51339	0.0000
Two-Year Change Parking	0.000156	3.01E-06	51.75401	0.0000
Two-Year Change Rd, Rl, Tr	0.002013	3.61E-05	55.80649	0.0000
Two-Year Change Other	-0.000203	9.40E-06	-21.55340	0.0000

Effects Specification

Period fixed (dummy variables)

R-squared	0.974601	Mean dependent var	-4427.153
Adjusted R-squared	0.974600	S.D. dependent var	2880.269
S.E. of regression	459.0389	Akaike info criterion	15.09618
Sum squared resid	6.88E+10	Schwarz criterion	15.09654
Log likelihood	-2465452	Hannan-Quinn criter.	15.09628
F-statistic	1253296	Durbin-Watson stat	0.545849
Prob(F-statistic)	0.000000		