The Z-axis:
Elevation Gradient Effects in Urban America

Victor Yifan Ye*
Charles M. Becker†‡§

November 24, 2016

Abstract

This paper presents an in-depth analysis of hilliness effects in American urban communities. Using data from seventeen cities, we establish robust relationships between elevation patterns and density and income gradients. We find that high-income households display strong preference not only for high-altitude but also for high unevenness, leading to spatial income stratification at both the city and tract-level. We analyze potential causes of this propensity: micro-climate, crime, congestion, view effects, and use of public transit. We conclude that multi-dimensional spatial methods are crucial to investigations of cities with substantial unevenness. Moreover, redistributive social and economic policies must struggle with a fundamental, topographical dimension to inequality.

JEL: J10, R11, R12

Keywords: Elevation, Hilliness, Household Income, Population Density, Urban Gradients, Spatial Modelling

*Department of Economics and Department of Statistical Science, 213 Social Sciences, Durham, NC, 27708, United States yifan.ye@duke.edu

†Department of Economics, Campus Box 90097, 213 Social Sciences, Durham, NC, 27708, United States cbecker@duke.edu

‡Corresponding author

§We are thankful to Patrick Bayer, Graham P. Clarke, Laurence J. Kotlikoff, Suminori Tokunaga, Edward Tower and session attendees at the 63rd Annual North American Meetings of the Regional Science Association International for advice and comments. All errors and misinterpretations remain our own.

1 Introduction

In the past two decades, spatial methods have experienced a surge in attention among economists. Much empirical research has focused on bi-dimensional, geographical-coordinate-based modelling approaches and operate under location-parity, spatial equilibrium (Glaeser and Gottlieb 2009; Grimaud and Laffont 1989) conditions. Cities are largely assumed to be flat, featureless plains. Governments, workers, households and firms, in the absence of economic incentives, are assumed to be indifferent among locations.

Yet we seem to understand intuitively that locations are *not* created equally. While walking down or up a slope is undesirable, wealthy individuals seem to strongly prefer living on hillsides; cities with uneven suburbs are centrally denser and often more expensive to live in. Our understanding of these characteristics is so deeply ingrained that one instinctively associates the affixes “hills” or “heights” with high-income communities. Conversely, “bottom” or “flats” generally connote low-income areas: one can hardly imagine John Steinbeck writing *Tortilla Heights* rather than *Tortilla Flat*.

This paper presents a formal study of the economic consequences of elevation as a locational attribute in a US-centric, multi-city framework. By constructing high-resolution Digital Elevation Models (DEMs) for seventeen major metropolitan areas, we demonstrate that altitude and unevenness of terrain strongly influence income and population patterns. Specifically, we find that controlling for spatial factors, doubling of a tract’s elevation variance, defined as the standard deviation of within-tract altitude samples, is associated with a decrease in population density of 20.8%, an increase in median household income of 4.4%, and an increase in the tract Gini coefficient of 1%.

In light of recent literature on geographical features as natural amenities or disamenities (Saiz 2010; Lee and Lin 2015; Burchfield et al. 2006), our paper develops the discussion in two important aspects. First, our approach suggests that elevation effects need not be explicitly linked to actual hills or mountains: terrain does not have to be so uneven as to constrain development or provide explicit utility in terms of views or attractions to influence density gradients or income distributions. We show that cities that are not conventionally considered ”hilly” (Kansas City; Dallas, TX) may nonetheless have significant effects related to elevation variance and altitude.

Second, beyond the explanation of “view” effects as natural amenities for high-income individuals, we provide evidence for four intermediaries through which elevation patterns influence major urban gradients: micro-climate, crime, congestion, and costs of accessing public transit. *Ceteris paribus*, our models indicate that households in high-altitude, high-unevenness tracts enjoy milder weather patterns, significantly lower rates for a range of crimes, less local congestion, and they commute less by public transit and walking. These effects are robust to comprehensive controls for local income profiles, suggesting that even relatively mild levels of unevenness may drive a wedge between utility of a location derived by high and low-income households.

In light of recent literature on elevation variability effects related to housing and use of public transit (Ye and Becker 2016), our work connects transaction-level price effects to measurable differences in usage patterns of public transit among different locations of a given city. With data from a multi-city panel, we conclude that such effects are at least somewhat consistent across cities with varying degrees of public transit accessibility. It also suggests the potential for elevation gradient effects to price into housing markets through other factors.

The salient prediction of our findings is that independent of centricity, local amenities and historical conditions, spatial socioeconomic stratification contains an elevation gradient component: for a large range of cities, certain locations are intrinsically attractive to high-income individuals and unattractive to low-income ones. It is not difficult to conceptualize how outcomes predicted by sorting models of racial aversion (Courant 1978; Bailey 1966; Bayer and McMillian 2005) could be reinforced and perhaps even exacerbated by a purely geographical aspect to segregation. Such persistent and fundamental barriers to integration no doubt pose significant challenges to designers of redistributive microeconomic policies.

Section 2 reviews prior literature related to elevation effects and spatial modelling. Section 3 outlines empirical methodologies and data sources utilized in the geospatial modelling process. Section 4 discusses model robustness and presents evidence for elevation effects in density and income gradients; section 5 presents evidence for intermediary effects. Section 6 concludes.

2 Prior Literature

The classic spatial literature addresses a number of relationships between aspects of natural geography and economic outcomes. Somewhat similar to our multi-city framework, Bosker and Buringh (2015) utilize broad, geographical perspectives to explain initial locational choice of major European cities. Rappaport (2007) models nice weather as a commodity to explain US migration flows. The role of water bodies have been investigated from perspectives such as the constraining of land supply due to lakes and oceans (Rose 1989) and the economic attractiveness of coastal living (Rappaport and Sachs 2003).

Literature specifically concerning elevation effects is relatively limited. Existing research primarily focuses on flood risk as an undesirable outcome of low elevation (Scawthorn, Iemura, and Yamada 1982; Shilling, Sirmans, and Benjamin 1989). Recent work on similar topics features significantly more advanced methods such as the combining of spatial methods with hedonic housing data (Bin et al. 2011) or with dynamic models (Husby et al. 2014). While not featuring elevation as a primary investigative target, Kok, Monkkonen, and Quigley (2014) find evidence of elevation effects in San Francisco land value.

Saiz (2010) presents a study of geographical influences in housing and land markets of 73 major MSAs, finding that undevelopable land on the city periphery is a strong predictor of low housing supply elasticity. While his work directly investigates elevation effects on urban gradients, we identify two shortcomings. First, by using a 15% slope cutoff, an artificial dichotomy is created between developable versus undevelopable land, while in reality the decision to develop a location depends on a wide range of supply and demand-side factors. Second, the model assumes that elevation primarily influences urban gradients by limiting land supply, whereas phenomena such as high-income household preferences for elevation do not seem to be driven by the supply of land. Here, we show that by modelling elevation not as a binary term but as continuous factors of unevenness and altitude, these issues can be addressed successfully.

Lee and Lin (2015) build on prior work (Bleakley and Lin 2012; Lin 2015) on the role of geographic features in persistence of income distribution and present a formal theory and empirical evidence of the role of natural amenities in shaping income distributions across space. Specifically, they model natural geographical features as "anchors" for high-income households and show that proximity to hills, as with proximity to a range of other natural amenities such as coastal proximity and lakes, is associated with a positive income effect.

Our contribution to this discussion is twofold. First, while our findings confirm that a "pure" natural amenity is indeed likely to exist for actual hills as a preference for scenery, we also show that the influence of terrain is not limited to locations that are conventionally considered as having such amenities. In other words, cities that do not have salient natural features in terms of hilliness (Kansas City; Pittsburgh) are also shaped by local elevation profiles.

Second, we present evidence that elevation variance and altitude is not only an amenity to the rich but also a cost to the poor. Through the intermediary effects as outlined in Section 5, middle-income households that are likely to not pay significant premiums for scenery would nonetheless strongly prefer locations with higher altitude and unevenness. Low-income households that regularly use public transit or walk would perceive such locations as a distinct disamenity. These findings suggest that the role of elevation is richer and more nuanced than presented in the existing literature.

Using data from the Hong Kong government-subsidized housing market, Ye and Becker (2016) find significant elevation gradient effects in apartment sale prices. The paper concludes that walking up or down an incline to public transit hubs is highly undesirable: holding other factors constant, a 1-decimal-degree increase in the slope between an apartment and the closest metro station decreases its selling price by up to 1.9%. While Hong Kong may be an outlier among global cities because of its highly uneven terrain, these results are nonetheless significant. In this present paper we not only demonstrate existence of public transit-elevation effects in US cities in general, but also expand the influence of elevation to a much broader set of urban gradients.

3 Methodology

3.1 Data Collection and Summary

At the core of our analysis is the merging of two datasets: a panel of information on census-tract-level statistics, and a dataset of elevation gradients and other geospatial factors. For altitude and unevenness, we use the Microsoft Representational State Transfer (REST) Services to construct city-level DEMs.[1] Altitude figures are sampled from each city on a 0.3-by-0.3 decimal-degree square region centered on respective downtown areas at a grid resolution of 500×500, and joined to individual census tracts using boundary data from the US Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) database.[2] The mean altitude and elevation variance of each census tract is calculated by taking the average and standard deviation of all interior sample points. We henceforth abbreviate these metrics as, respectively, *mal* (mean altitude) and *elv* (standard deviation of elevation).

Since tract boundaries do not usually coincide with edges of the sampling area, tracts on the border of the sampling square that, excluding water area, are not at least 80% covered by elevation sample points are excluded from the dataset.[3] To prevent *elv* estimates from being calculated from samples that are too small, we also exclude tracts that contain less than 20 elevation sample points. To address downward bias in *elv* caused by local water areas, we examine each city and remove elevation samples at the local water level.[4]

To illustrate the existence of *general* elevation effects across cities, we select cities for elevation data collection to maximize diversity in terms of both unevenness and geographical orientation. Seventeen major cities and surrounding suburbs are selected: Atlanta, Boston, Charlotte, Cincinnati, Dallas, Denver, Kansas City, Los Angeles, Miami, Nashville, New York City, Pittsburgh, Portland (Oregon), San Diego, San Francisco, Seattle, and Washington, DC. All of the ten Standard Federal Regions are represented by at least one city, and cities are divided roughly equally among Census Bureau statistical regions. Figure 1 contrasts the distribution of census tract *elv* of dataset cities.

Performing elevation sampling over the aforementioned cities yields usable elevation data for a total of 6,645 census tracts, with an average sampling density of 463.1 points per tract for a total of approximately 3.08 million samples in total, or about 181,000 per MSA. To account for outliers in elevation level (Denver), we transform altitude measurements from meters to a "relative alti-

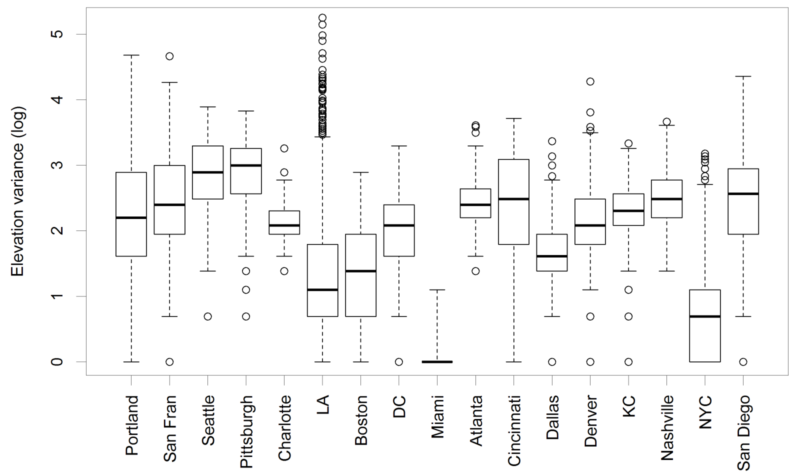
[1] DEM contour plots of select cities are presented in Figure A9 in the appendix.

[2] Decimal degrees translate to marginally different distances in regular length terms at different latitudes. For the range of latitudes in the dataset, the side length of the sampling square may vary from 22.5km (Seattle, 47.5°N) to 30km (Miami, 25.8°N).

[3] We estimate coverage by comparing the per-unit area point count of partially covered tracts with that of fully covered tracts in the same city.

[4] Sea-level elevation samples are removed from coastline cities, as are samples from major local rivers, reservoirs and lake surfaces of landlocked cities.

Figure 1: Standard boxplot of tract elevation variance (log) by city



tude" metric *ral* of standard deviations above or below the city average. To control for fixed effects of tract proximity to the ocean, we construct a separate spatial model using National Oceanic and Atmospheric Administration (NOAA) coastline profiles and calculate tract distances to the coastline.[5]

We utilize the REST routes Application Program Interface (API) to generate driving distance and time estimates from each census tract to respective downtown areas.[6] Three sets of values are collected: a hypothetical "no-traffic" estimate, an estimate assuming traffic conditions at 7:00am local time, and a third one assuming conditions at 7:00pm. By introducing a road-network based monocentricity metric and separately controlling for driving times at different intervals of the day, this approach accounts for the presence of natural barriers to commuting (New York City, San Francisco) as well as possible temporal differences in commuting time costs; i.e., heavy downtown traffic that only occurs during the morning or evening commute.[7]

The routes API is also used to construct a measurement of local traffic congestion. From the center of each census tract, we calculate "no-traffic" and 7am driving time estimates to eight different destinations at an equal linear distance of 8 kilometers, at bearings from 0° to 315° in 45-degree increments.[8] The

[5] "Coastline" is defined by NOAA guidelines. http://shoreline.noaa.gov/glossary.html

[6] The API is set to optimize for driving time and, if time estimates are equal for two or more routes, to choose the shortest route.

[7] Conceptually, we would not expect all residents of all tracts to commute regularly to downtown. However, real proximity to downtown in terms of expected travel costs is a strong predictor for a rich variety of spatial factors such as housing value and local amenities. By explicitly controlling for travel time, we effectively absorb much of the heterogeneity in income and density related to those effects.

[8] If a destination point is on water or otherwise inaccessible as estimated by the API, the

proportional relative time penalties of the 7am drive are averaged and standardized on a 0-100 scale to create a tract "congestion score", with 0 representing maximum and 100 representing minimum congestion among dataset tracts.

Elevation and spatial data are combined with demographic, income and education information on census tracts from the 2009-2013 five-year American Community Survey (ACS). Removing observations with missing data yields a remaining total of 6,383 observations usable for the regression model.[9] Population density figures are generated by dividing ACS population figures over standardized TIGER tract area size estimates. To adjust for potential spatial autocorrelation (Dubin 1991; Hubert, Golledge, and Costanza 1981) on variables of interest, we calculate city-specific spatial autocorrelation values for all response variables using between-tract point distances and a squared Euclidean distance (L2) penalty scheme.

3.2 Modelling methods

Our empirical specifications begin with a Linear Mixed Effects (LME) Model (Goldstein 1987). The reasoning for employing this relatively complex procedure is twofold. First, the implicit assumptions of LME are most consistent with our intuitive understanding of the role of elevation in urban gradients. Households, being fundamentally somewhat similar, should also be similar in their aggregated response to elevation effects. On the other hand, one would not expect that households across different cities perceive elevation variance *identically* when testing for the existence of such effects because US cities vary so much in terms of elevation profile. Our observations also take the form of an unbalanced city-level panel, where within-city observations are much more highly correlated than observations in different cities.

A Fixed Effects Linear Model (FELM) captures the commonality of elevation gradient effects but assumes that they are identical across cities by extracting a single slope for all cities. Running separate OLS regressions by city captures the variation in the effect of elevation but implicitly assumes that the effects being measured are fundamentally different. In contrast, LME negotiate a midpoint between these approaches by allowing for slopes by city under the assumption of measuring "emissions" of a common effect.

Second, LME allows for the testing of whether elevation effects differ significantly by city or is significant for a particular city with regard to the null with no need for additional mechanisms or modelling (such as model comparison tests for multiple separate regressions). We consider this a superior approach in terms of consistency to specifying separate models for tests for a general versus city-specific effect. We also note that running separate regressions by city would inflate the difference between elevation effect coefficient estimation by assuming fundamentally different effects.

furthest accessible point within the 8km-range with the same bearing is used.

[9]See Appendix Section C for a discussion on not utilizing imputation.

The primary challenge we face in drawing inferences on city-specific effects is that null hypothesis tests derived from a single LME model are generally regarded as unreliable.[10] Our solution is to combine LME with a bootstrapping model that additionally incorporates uncertainty of our data by using "pseudo-data" drawn from Gaussian distributions from the original data with ACS-reported standard errors. Appendix Section A provides a discussion of this issue as well as the methodology and advantages of combining the LME model with bootstrapping and pseudo-data.

We note that this is also an advantage of using ACS over 2010 Census data: census data collection is inevitably error-prone and errors are generally non-random. Locations with more owner-occupied residencies, high density, and within major urban areas (which may correlated with income) generally will have more reliable estimates. Because these gradients *are* our response variables, we consider it extremely important that uncertain estimates be strongly discounted in deriving coefficients.

We also note that pseudo-data bootstrapping precludes the need for examination of outliers: outliers with high certainty have good reason to be factored into the model, and the effect of outliers with low certainty will be averaged out across bootstrap iterations. Observations where only one or two variables have extreme values will also be discounted less than observations where many covariates take extreme values. It also lowers the likelihood of statistically significant findings being an artifact of the data itself: estimates of individual iterations do not necessarily have to be distributed symmetrically with regard to the mean.

A summary of tests of model performance is provided in Appendix Section B, including favorable comparisons to both single-model tests (LME, FELM, OLS) and pseudo-data bootstrapping of reduced models. See Appendix Section C for an extended discussion on computing LME model coefficient p-values and estimating R^2 under pseudo-data bootstrapping.

Elevation as a spatial attribute lacks reciprocity: it shapes how cities are built, yet only in rare cases do human activity significantly change the profile of land. Therefore it is not difficult to establish basic causality in the sense that urban gradients are shaped by elevation, yet is difficult to isolated direct causal effects from indirect effects that initially originate from elevation, yet should be attributed to other, "intermediary" urban gradients. Also, we would expect elevation to be correlated with income gradients even if there were no current or ongoing amenity effect so long as historical amenity effects exist.

We address this concern by restricting the size of models to estimate the effect of elevation on major urban gradients: for population density, only spatial factors (coastline distance, driving distance and time estimates, tract size, and presence of water) are controlled for; for household income only spatial factors and density, and for census tract Gini coefficients we include only spatial factors, density and median income[11]. This allows for the establishment of basic,

[10] See Appendix Section A for an extended discussion.

[11] We note that adjusting for tract size absorbs the variation in *elv* caused by larger tracts

"location-neutral" relationships between elevation and major gradients. Extra control variables, such as those for employment, occupancy status, income level and ethnic composition are included in models for intermediary effects in Section 5, as these analyses are intended to demonstrate strong evidence for direct causality of ongoing utility provided by elevation. We provide a summary of parameters for each model in Tables A2 and A3 in the appendix.

While this structure may seem somewhat semi-recursive in nature, it avoids the difficulty of interpretation that including extra control variables in the main gradient models or fewer control variables for the intermediate effect regressions may cause. However, disentangling elevation gradient effects from other spatial factors, most importantly monocentricity and proximity to water, is still necessary because downtown areas and areas close to a lake or ocean are likely to be the flattest parts of a city. In a similar vein we control for spatial autocorrelation, given that many, if not all urban gradients we investigate (density, income, crime, congestion) are expected to display high levels of spatial dependency.

4 Elevation Effects in Major Urban Gradients

Controlling for spatial factors outlined in Section 3.2 and autocorrelation, we find strong evidence for a negative relationship between tract *elv* and population density ($p<0.01$), and a positive relationship between elevation variance and median household income ($p=0.014$).[12] The relationship between the tract Gini coefficient and *elv* is marginally significant ($p=0.07$). Holding spatial factors constant, we estimate that doubling of *elv*, assuming city-average $ral = 0$, is associated with a decrease in density of 20.8%, an increase in tract median income of 4.4%, and an approximately 1% increase in the tract Gini.[13] The relationship between *ral* and population density is highly significant but positive ($p<0.01$), assuming zero elevation variance.

We note that there is strong evidence that density and income gradients of different cities are associated differently with *ral* and *elv*. As Figure 2 indicates, while there exists a shared component in the relationship between *elv* and population density, the mean of coefficient distributions and the possibility of rejecting the null hypothesis differ significantly by city.[14]

The key finding from Figure 2 is that hilliness is associated with reduced

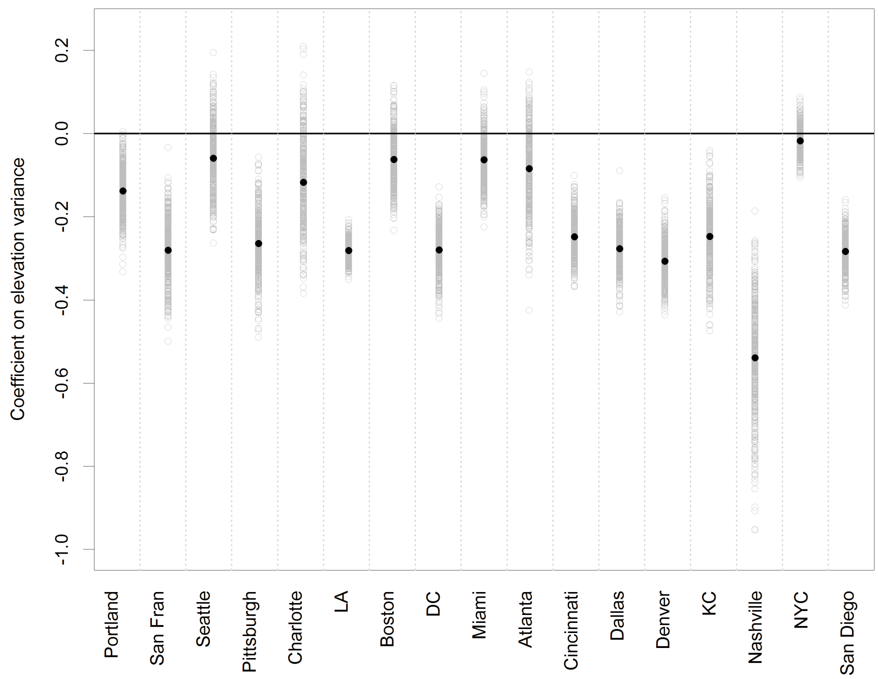
having more elevation sample points and hence generally having higher variance.

[12] For ease of interpretation, we translate coefficient distributions into null hypothesis confirmation rates under a one-tail assumption. Note that because of the limited number of coefficient draws, it is not feasible to inference beyond 99% confidence level. For coefficients with less than 5 draws that reject the alt hypothesis, the notation "$p<0.01$" is used.

[13] To clarify, the effect size is 1% of the coefficient and not one percentage point. In other words, doubling *elv* is associated with the change from a given GINI of 0.4 to 0.404, not 0.41.

[14] Random effects for elevation variance on population density for Seattle, Charlotte, Boston, Miami, Atlanta and NYC cannot be established at 95% certainty. $p<0.001$ for unpaired t-tests between the distribution of coefficients on Nashville and that of all other cities.

Figure 2: Distributions of random effect coefficient estimates of log(elv) by city, population density model



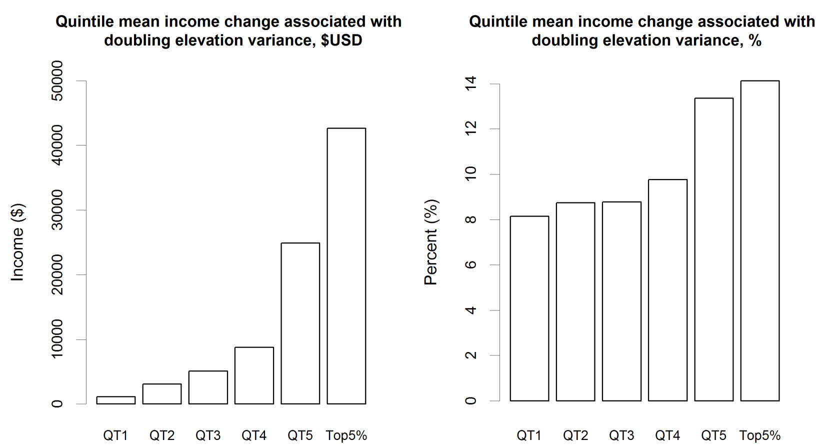
population density in most metropolitan areas yet the effect varies considerably across them. A metro area such as Nashville has, in effect, two parts, largely divided by the Cumberland River. The northeast is characterized by low-lying, low altitude variance, high density areas. Heading south toward Williamson County, hilliness increases markedly, as does socio-economic status of the inhabitants, which in turn is associated with decreased density. In contrast, New York City and Boston are less hilly, especially near the center and hence socio-economic differentiation is less associated with topographic variation. While a formal investigation of city-specific elevation effects is beyond the scope of this paper, we present select by-city estimates in Figures A5, A6, and A7.

A positive altitude-density relationship may seem counterintuitive. However, in reality increase in altitude is almost always strongly associated with an increase in unevenness: the coefficient on the *elv*-*ral* interaction effect is negative and also highly significant (b=-0.06, $p<0.01$). Consequently, incorporating interaction effects, the negative elevation variance-density connection strengthens as altitude grows, and the positive altitude-density relationship flattens and eventually slopes downward as *elv* increases. At 3x dataset-average elevation variance (24.3), a increase in *ral* of 1SD translates to a decrease in population density of 4.3%. Fitted curves for both effects, assuming dataset-average density at average altitude and elevation variance, are presented in Figure A8.

To better understand the elevation-income relationship, we perform pseudo-data bootstrapped random effects regressions on tract mean annual household income data by quintile, controlling for spatial factors, density and autocorrelation of income levels.[15] The coefficient on *elv* is strongly significant and positive for all quintiles: mean quintile income increases by $1,132, $3,074, $5,087, $8,777 and $24,895, respectively, per doubling of elevation variance. Mean income of the top 5% of households by annual income increases by $42,628.

Two observations can be made with regard to quintile income effects: first, income is higher for all quintiles in tracts with high elevation variance. *Ceteris paribus*, even low-income communities in high-variance tracts are richer than counterparts in low-variance tracts. Second, the proportional changes in quintile income levels are similar: assuming dataset-average annual income levels, percentage increases in income associated with doubling *elv* are between 8-13% for all quintiles. Interaction effects are also significantly positive and similar across quintiles: doubling *elv* translates to percentage increases in quintile income of 12-17% at 1-SD above city-average altitude. Changes in absolute and proportional income are contrasted in Figure 3.

Figure 3: Absolute/proportional change in quintile income level, doubling elevation variance at $ral = 0$



These results are consistent with our findings on tract Gini and median household income. Significant, positive income effects that are somewhat consistent across income groups translate to a modest increase in the tract Gini.

[15]Random slopes on elevation parameters are removed to reduce computational load. Z-tests are performed on single LME models with mean quintile income data as the response to confirm output significance. Using both methods, all elevation variance coefficients are significant at the 99% level.

Given the evidence in Figure 3, the higher within-tract income inequality of high *elv* tracts is not caused by the presence of impoverished low-income communities, but rather by the existence of wealthier high-income communities. As within-tract income distributions tend to be strongly long-tailed, median income changes are naturally less drastic than changes to the mean.

The income effect of elevation is reflected in a broad range of demographic pattern behaviors. After controlling for spatial factors, density and autocorrelation, we find that individuals in high *elv*, high *ral* tracts are older, more likely married, have smaller households, and are more highly educated. Holding other factors constant, a doubling of *elv* is associated with a 4.4% increase in the proportion of tract residents above the age of 64 (rising to 6.0% at ral = 1), a 1.3% increase in the tract median age (2.4% at ral=1), a decrease in the average household size of 0.026 members (0.036 members at ral = 1), a decrease in the average married household size of 0.045 members (0.084 members at ral = 1), a 1.2% increase in the proportion of married individuals (2.7% at ral = 1), a 5.2% increase in the proportion of bachelor's degree holders (8.4% at ral = 1), and a 8.7% increase in the proportion of graduate and professional degree holders at city-average altitude(13.9% at ral = 1).[16]

The purpose of these models is not simply to reiterate well-established behavioral and demographic patterns of high-income households, but to show that these patterns contain, to some extent, an *elevation-oriented component*. Much economics literature has focused on the modelling and much policy has addressed on the mitigating of education gaps, inequality, and segregation in American cities. Yet while historical and cultural aspects of these issues are at least in principle resolvable, there are few ways to redistribute utility gained from hilliness and altitude to those who live on flat, low-lying land.

It is important to note that our findings are causal only in the sense of identifying spatial distributions of income stratification and not of the degree of stratification. Although it is likely that relatively more uneven cities are indeed more stratified income-wise, it is also possible that elevation gradients merely determine of where will stratification occur. We do not address this question here because, given the focus on census-tract-level effects in this paper, our current dataset is not particularly suitable for drawing inferences at the multi-city-level.

One way to investigate the relationship between elevation gradients and the level of income stratification across cities is the merging of lower-resolution elevation data with a large panel of Metropolitan Statistical Areas. Our work in progress combines this approach with a city-scale counterfactual simulation model (Bayer, McMillian, and Rueben 2004) of the housing market in an effort

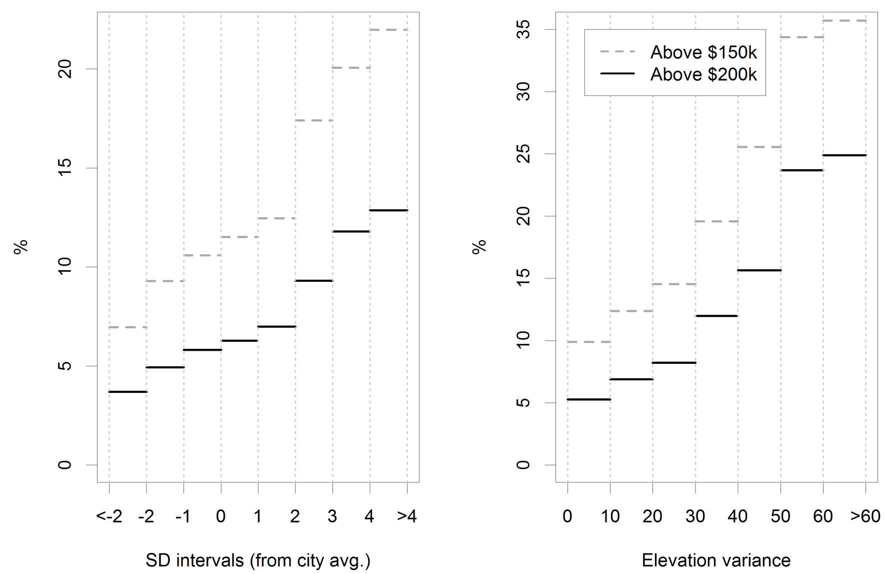
[16] p<0.01 for coefficient on *elv* for regressions on married household size, proportion of bachelor's degree holders and proportion of graduated degree holders. p=0.01 for regression on median age, p=0.014 for regression on proprotion of elderly individuals, p=0.024 for regression on household size, and p=0.154 for regression on proportion of married individuals. All interaction effects significant at p<0.01 except regressions on elderly individual proportion and household size (p=0.014 and p=0.032, respectively).

to conclusively determine whether cities would become less stratified if elevation features were removed.

5 Intermediary Effects

Our findings in Section 4 indicate that holding spatial and autocorrelation effects constant, high-income household strongly prefer locations that are both hilly and at higher altitudes. Conceptually, the most straightforward explanation of this preference is view effects: a combination of high altitude and high elevation variance leads to access to beautiful scenery, a strictly luxury good among housing amenities. Low-income households will have price-elastic demand for scenery, and hence be priced out of high-*ral*, high-*elv* neighborhoods.

Figure 4: Percentage of high-income households by annual income bracket/elevation block[17]



Although view effects are difficult to test directly, there is suggestive evidence. First, as shown in Figure 4, the proportion of high-income households is progressively higher in high-*elv*, high-*ral* tracts. In tracts with elevation variance greater or equal to 60, 35% of households have an annual income of above $150,000, more than three times the dataset average of 11.1%.[18] Second,

[17]Each percentage estimate denotes average level of tracts with elv/ral between the two bounds.

[18]31 out of the 6,383 observations in the data, approximately 0.5%, report $elv \geq 60$.

elv and $elv \cdot ral$ effects remain positive and significant ($p<0.01$) for the regression on the proportion of households with annual income above \$200,000 after controlling for spatial factors, autocorrelation, income level, income inequality, demographics, proxies for local economic conditions, public transit use, local congestion *and* local weather conditions.[19] We consider this result suggestive of a "pure" elevation amenity that exists independent of local economic conditions and intermediary effects, most conceivably through the form of scenery.

However, there are two reasons to suspect that view effects do not tell the whole story. First, the proportion of high-income households, generally low at below city-average ral and low elv levels of 0-20, increases on the margin at these levels of altitude and variance. Yet it is conceptually unlikely for there to be any scenery effects at such variances and altitudes. Second, all income quintiles are of significantly higher income in tracts with greater elv. Yet it seems improbable that any but the one to two highest-income quintiles are strongly influenced by view effects when choosing locations of residency.

Considering Figure 4 in light of our finding that elevation effects are significant in cities without obvious scenery locations or strong unevenness, it seems likely that elevation is not only a historical determinant of income distributions and not only has an impact through scenery, but influences income gradients through other amenity effects which are attractive to middle-income households and perhaps unattractive to low-income households. Based on these observations, we propose four additional explanations of why high elevation-variance and high-altitude may be positively correlated with income: better micro-climate, difficulty of accessing public transit and lack of walkability, lack of local congestion, and lower crime rates.

5.1 Micro-climate effects

While it is well-established that individuals prefer nice weather as a good with high price elasticity of demand (Rappaport 2007) and, in geography literature, that there is a strong relationship between topography and climate (Geiger, Aron, and Todhunter 2009; Linacre 1982), it is not immediately clear whether micro-climate effects are observable at the census-tract level for dataset cities and, if so, whether they are correlated with our estimators of altitude and unevenness. To show that climate effects exist, we merge our dataset with NOAA historic weather data on monthly temperature levels by zip code. Random-intercept regressions are performed on aggregate metrics of weather against elv, ral, $elv \cdot ral$ and spatial control variables. Population density is included as a proxy for temperature effects caused by human activity.

Other factors constant, we find that census tracts with higher elv and ral enjoy better weather. Doubling of elv is associated with a decrease in the

[19] The elv term is positive but not significant for the regression with proportion of households with annual income above \$150,000 ($p=0.13$), and $elv \cdot ral$ is positive and marginally significant ($p=0.05$). A list of variables used and details are presented in Table A2.

standard deviation of monthly average temperatures by 2.1% ($p<0.01$) at city-average altitude, and a decrease of 2.5% at 1-SD above city average altitude. Doubling variance is also associated with a decrease in the average monthly maximum-minimum temperature spread of 2.6% and a decrease in the annual number of Cooling Degree Days (CDDs) of 2.8% .[20] $elv \cdot ral$ is also associated with marginally more rainfall: at 1-SD above city-average ral, doubling variance is associated with 0.8% more annual rainfall ($p<0.01$).

These estimates suggest that there are statistically significant micro-climate effects at the zip-code level. Areas with higher elv and ral are cooler, receive slightly more rainfall, and enjoy lower volatility in temperature both between months of the year and within each month. These effects are not trivial: an increase in elv of one or two orders of magnitude translates into significantly smoothed temperature cycles independent of general climatic effects associated with centricity, local water bodies or human activity.

5.2 Public transit and congestion

We investigate public transit effects by analyzing commuting methods of individuals in high-variance, high-altitude tracts: using ACS estimates, we perform pseudo-data bootstrapped random-intercept regressions on the percentage of tract workers commuting by public transit and by walking. In addition to spatial factors and autocorrelation of use of alternative methods of transit, we introduce control variables for tract income level, income inequality, and demographics, as well as proxies for local economic conditions.[21]

Holding all other factors constant, we find that workers commute significantly less via public transit and walking in tracts with high elv. At city-average altitude, doubling elv translates to a 3.7% reduction in the percentage of workers who commute by public transit ($p \approx 0.03$). Doubling elv at 1-SD above city-average altitude is associated with a reduction of 7.1%. The interaction effect $elv \cdot ral$ is not significant for effects on percentage of workers commuting by walking. However, the elv effect for walking is significant, with the percentage of commute-by-walking workers decreasing by 5.6% per doubling of elv.

The difference in altitude effects perhaps can be explained by the extra supply-constraint of public transit use versus walking: individuals also use public transit less when there are fewer options available. Mass transit systems, by the nature of cost-optimization, almost always will be restricted by unevenness along routes. Here, we attribute the elv effect to undesirability of walking down or up steep slopes, which applies equally to commuting by walking and to walk-oriented public transit hubs, and the ral effect for public transit to a lack of transit options in tracts with higher relative altitude.

[20] All $p<0.01$. Annual CDDs are calculated by subtracting 65°F from the daily average temperature when the daily average is greater than 65°F, and then summing over all days of the year.

[21] Detailed specifications are provided in Appendix Table A2

Conceptually, the effect of extra costs of public transit is threefold. First, when considering the existence of a city-level equilibrium of public transit costs, low-income households may outbid high-income households in locations where public transit is relatively accessible (Glaeser, Kahn, and Rappaport 2008), resulting in central concentration of poverty. Our findings imply that households that heavily utilize public transit would prefer flatter instead of uneven areas. This provides poor, public-transit dependent populations with incentives to reside in generally flatter downtown areas, even if suburbs are well-endowed with public transit options.

Second, high-income households that do not use public transit will pay a "flatness premium" when not selecting areas with relatively high elv and ral, and display opposite preferences. Third, public transit hubs may also be associated with higher levels of local traffic and commercial activity (Kahn 2007). If major hubs are mostly located in flat regions of a city, local high-income communities, in an attempt to avoid close proximity to hubs, also may choose areas with high elv and ral.

Performing the same regression procedure on congestion scores, we find the coefficient on ral to be negative and highly significant ($p<0.01$). $elv \cdot ral$ is positive and somewhat suggestive ($p=0.134$). Holding other factors constant and elv at zero, a 1-SD increase in ral decreases the congestion score by 0.82 points. This effect is reduced at higher elevation variance: at dataset-average elv, the same increase in ral only decreases the congestion score by 0.59 points.

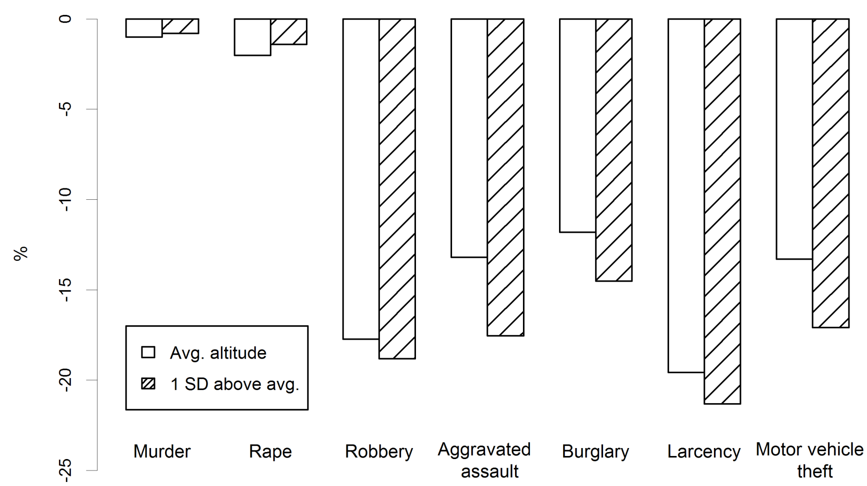
The direction of the relationship between elv and congestion is consistent with intuition: uneven and winding roads are harder to navigate and contribute to local buildups of traffic. The stronger altitude effect is most convincingly explained by proximity to major roads: freeways and other large road network components are usually not located in relatively high-altitude areas. Therefore, high-income households may also view high-altitude location as preferable because they avoid negative externalities associated with major roads, while low-income households that do not own vehicles would be relatively indifferent or even averse to such locations since major roads are also associated with better public transit options.

5.3 Crime

To investigate the relationship between elevation patterns and local crime, we merge our dataset with tract-level crime rate data from the 2000 National Neighborhood Crime Study (NNCS). The study, conducted by Peterson and Krivo (2000), provides detailed crime rate statistics by type of offense for seven offense types and 9,593 census tracts. Of the NNCS study tracts, 2,195 coincide with tracts in our dataset and are used in the following analysis.

Controlling for spatial factors, autocorrelation, tract income level profiles, demographics and proxies for local economic conditions, we find that tracts with higher ral and elv enjoy significantly lower crime rates. However, this effect does not apply uniformly for all crime types: hilliness effects are highly significant for robbery, aggravated assault, burglary, larceny and motor vehicle theft (MVT; all $p<0.01$), but less so for rates of murder and rape.[22] The size of the effects also differs tremendously: doubling *elv* translates to a decrease in rates of theft categories and aggravated assault by 12.6%-21.6%, but to a decrease in murder rate of only 1% and a decrease in the rape incident rate of 2.2%. Interaction effects are inconsistent in significance, but uniformly negative for theft categories and aggravated assault.[23] Figure 5 presents a contrast of crime rate changes associated with doubling of *elv*.

Figure 5: Crime rate reduction by category corresponding to doubling of *elv*



The difference between reductions of crime rates is convincingly explained via the nature of these crimes. Categories such as larceny and MVT are generally "foot" crimes: for example, it is unlikely that a person would drive one vehicle to steal another. For such crime types, unevenness of terrain introduces significant cost: if criminals stand to profit equally from all areas, they would choose to commit crimes in flatter areas to avoid walking up or down slopes. It is also possible that uneven terrain introduces addition risk by hindering one's view in certain directions and hence complicating one's plan of escape. Being on foot, it is also conceivably more difficult to evade capture from law enforcement, which would have ready access to vehicles regardless of local *elv* or *ral*.

In contrast, murder and rape are much more commonly committed as domestic crimes and between acquaintances. Their occurrences are hence not as significantly affected by elevation effects as the other crime categories. Given

[22] $p=0.036$ for murder rate. $p=0.02$ for rape incident rate.

[23] $p=0.16$ for robbery rate, $p<0.01$ for aggravated assault rate, $p=0.016$ for burglary rate, $p=0.118$ for larceny rate and $p<0.01$ for rate of MVT.

the significance of these effects, we conclude that areas with high *elv* are likely inherently "safer" than flatter areas of a city, especially for non-violent property crimes. If safety is priced into the cost of housing and other local amenities, these locations are expected to have higher concentrations of high-income households.

We note that while we cannot absolutely rule out the possibility that we are measuring the residual of household wealth effects when estimating the effect of elevation on crime and transit effects, we consider it highly unlikely for two reasons. First, we are controlling not only for mean and median income in the regression but the entire profile of quintile income levels. This means that any residual on wealth levels must not be picked up by all quintile income values, monocentricity variables which control for local bid-rent curve slopes, or the spatial autocorrelation term that approximate expected local levels of the response variable given a smoothed spatial distribution.

Second, our statistical method is specifically designed to account for uncertainty of both whether we observe a representative sample of tracts and whether our observed observations are reliable. If our coefficient estimates are robust to both alternative data values (as distributed in the pseudo-dataset) *and* alternative sets of data (as drawn with bootstrapping), then there is little reason to doubt that such estimates are "real" in the sense that they are separable from other effects as estimated in the model.

While these four effects clearly are not the only intermediaries between elevation and income, establishing their existence is nonetheless illuminating. Contrasted with view effects, these factors are capable of influencing the decision-making of middle and upper-middle class households: relatively few households are affluent enough to afford a truly beautiful view, yet many can afford to enjoy relatively mild weather, commute regularly by car, live far away from noisy, dirty freeways, and pay for low crime. In contrast, households that do not own vehicles will be less responsive to lack of congestion and strongly prefer public transit accessibility These effects are consistent with income gradient effects that predict consistently higher income in high-*elv*, high-*ral* tracts.

6 Conclusion

We endeavor in this paper to demonstrate that topographical structure–specifically, the standard deviation of elevation and relative altitude of a given location–matters for density, economic and social outcomes. We show that elevation should not only be considered as a natural amenity arising from the existence of hills but as a much more nuanced effect separable as preferences for altitude and elevation variance.

The unmistakable finding is that topographic structure does matter, that its importance interacts with many different variables, and that it differs from one MSA to another. Elevation variance and altitude are robust predictors of density and income gradients at both the city-level and the census-tract-level. Other spatial factors constant, tracts with greater unevenness and higher altitude are less dense, consist of households with significantly higher average income, and have higher Gini coefficients. We expect these effects to not only apply to cities with extreme unevenness but also to those with relatively mild elevation variance.

We further show that elevation effects are rooted in powerful economic incentives. We present evidence that uneven locations deter public transit use, have decreased walkability and lower crime rates, and enjoy less local congestion and smoothed micro-climate cycles. Similar to the draw of natural amenities, these effects should attract the rich to uneven, high-altitude locations. However, they would also push the poor toward flat, low-altitude areas.

We conclude that the role of elevation gradients in American urban communities should not be neglected and that the richness of intermediary effects presents a significant challenge in terms of addressing fundamental utility advantages of certain locations over others. The finding is an important and sobering one for those who hope to design policies that would reduce social inequality and racial segregation.

References

Bailey, M. J. (1966). "Effects of race and other demographic factors on the values on single-family homes". In: *Land Econ* 42, pp. 215–220.

Bayer, Patrick and Robert McMillian (2005). "Racial Sorting and Neighborhood Quality". In: *NBER Working paper* 11813.

Bayer, Patrick, Robert McMillian, and Kim Rueben (2004). "An Equilibrium Model of Sorting in an Urban Housing Market". In: *NBER Working paper* 10865.

Bin, Okmyung et al. (2011). "Measuring the Impact of Sea-level Rise on Coastal Real Estate: A Hedonic Property Model Approach". In: *Journal of Regional Science* 51.4, pp. 751–767.

Bleakley, Hoyt and Jeffrey Lin (2012). "Portage and Path Dependence". In: *The Quarterly Journal of Economics*. DOI: 10.1093/qje/qjs011. URL: http://qje.oxfordjournals.org/content/early/2012/04/19/qje.qjs011.abstract.

Bolker, Benjamin M. et al. (2008). "Generalized linear mixed models: a practical guide for ecology and evolution". In: *Trends in Ecology and Evolution* 24.3.

Bosker, Maarten and Eltjo Buringh (2015). "City seeds: Geography and the origins of the European city system". Journal of Urban Economics.

Burchfield, Marcy et al. (2006). "Cause of Sprawl: A Portrait From Space". In: *Quarterly Journal of Economics* 121, pp. 587–663.

Courant, Paul (1978). "Racial Prejudice in a Search Model of the Urban Housing Market". In: *Journal of Urban Economics* 5, pp. 329–345.

Dubin, Robin A. (1991). "Spatial Autocorrelation and Neighborhood Quality". In: *Regional Science and Urban Economics* 22, pp. 433–452.
Geiger, Rudolf, Robert H. Aron, and Paul Todhunter (2009). *The Climate Near the Ground*. VII. Washington, DC: Rowman & Littlefield.
Glaeser, Edward L. and Joshua D. Gottlieb (2009). "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States". In: *Journal of Economic Literature* 47.4, pp. 983–1028.
Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport (2008). "Why do the poor live in cities? The role of public transportation". In: *Journal of Urban Economics* 63 (1), pp. 1–24.
Goldstein, Harvey (1987). *Multilevel Models in Education and Social Research.* Oxford: Oxford University Press.
Grimaud, A and J.J. Laffont (1989). "Existence of a Spatial Equilibrium". In: *Journal of Urban Economics* 25.2, pp. 213–218.
Hubert, L J, R G Golledge, and C M Costanza (1981). "Generalized Procedures for Evaluating Spatial Autocorrelation". In: *Geographical Analysis* 13, pp. 224–232.
Husby, Trond G. et al. (2014). "Do Floods Have Permanent Effects? Evidence From the Netherlands". In: *Journal of Regional Science* 54.3, pp. 355–377.
Kahn, Matthew E. (2007). "Gentrification Trends in New Transit Oriented Communities: Evidence from Fourteen Cities that Expanded and Built Rail Transit Systems". In: *Real Estate Economics* 35.2, pp. 155–182.
Kok, Nils, Paavo Monkkonen, and John M. Quigley (2014). "Land use regulations and the value of land and housing: An intra-metropolitan analysis". In: *Journal of Urban Economics* 81.1, pp. 136–148.
Lee, Sanghoon and Jeffery Lin (2015). "Natural Amenities, Neighborhood Dynamics, and Persistence in the Spatial Distribution of Income". In: *Federal Reserve Bank of Philadelphia Working Papers.*
Lin, Jeffrey (2015). "The Puzzling Persistence of Space". In: *Federal Reserve Bank of Philadelphia Business Review* (Q2), pp. 1–8.
Linacre, Edward (1982). "The effect of altitude on the daily range of temperature". In: *International Journal of Climatology* 2 (4), pp. 375–382.
Nakagawa, Shinichi and Holger Schielzeth (2013). "A general and simple method for obtaining R2 from generalized linear mixed-effects models". In: *Methods in Ecology and Evolution* 4 (2), pp. 133–142.
Peterson, Ruth D. and Lauren J. Krivo (2000). *National Neighborhood Crime Study (NNCS), 2000.* Government Document.
Pinheiro, J C and D Bates (2000). *Mixed-effects models in S and S-PLUS.* New York: Springer.
Rappaport, Jordan (2007). "Moving to Nice Weather". In: *Regional Science and Urban Economics* 37, pp. 375–398.
Rappaport, Jordan and Jeffery D. Sachs (2003). "The United States as a Coastal Nation". In: *Journal of Economic Growth* 8, pp. 5–46.
Rose, Louis A. (1989). "Topographical Constraints and Urban Land Supply Indexes". In: *Journal of Urban Economics* 26, pp. 335–347.

Saiz, Albert (2010). "The Geographic Determinants of Housing Supply". In: *Quarterly Journal of Economics* 125.3, pp. 1253–1296.

Scawthorn, Charles, Hirokazu Iemura, and Yoshikazu Yamada (1982). "The influence of natural hazards on urban housing location". In: *Journal of Urban Economics* 11.2, pp. 242–251.

Shilling, James D., C. F. Sirmans, and John D. Benjamin (1989). "Flood insurance, wealth redistribution, and urban property values". In: *Journal of Urban Economics* 26.1, pp. 43–53.

Snijders, Tom A.B. and Roel J. Bosker (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* 2nd ed. Thousand Oaks: SAGE.

Ye, Yifan and Charles M. Becker (2016). "The (Literally) Steepest Slope: Spatial, Temporal, and Elevation Variance Gradients in Urban Spatial Modelling". In: *Economic Research Initiatives at Duke (ERID) Working Paper.*

Appendix

A Regression Methodology

The evaluation of fixed effect statistical significance in LME models is a topic of contention in modern statistics (Bolker et al. 2008; Pinheiro and Bates 2000), and equivalents of p-value-based measures in generalized linear models, such as the Wald Z-test, are generally considered insufficiently conservative and hence unreliable. Variations of Chi-square tests perform better, but usually are limited to balanced panels. Moreover, the reliability of ACS census tract data, as suggested by reported standard errors of estimates, varies greatly between tracts and categories. Our procedure therefore must account for disparities in data accuracy across these dimensions and confirm or reject significance given this extra aspect of uncertainty.

Our solution is also twofold. First, we utilize error estimates provided by ACS by drawing large numbers of "pseudo-datasets" from Gaussian distributions centered on the actual tract-figure estimates with variance determined by the reported error. Percentage estimates are drawn from normal distributions truncated on $[0, 100]$ and count estimates drawn from distributions truncated to non-negative values. This approach introduces diffused uncertainty across both tracts and regression parameters: by iterating over all pseudo-datasets, the influence of high-error tracts and variables is reduced. Second, we perform bootstrapping over observations in each pseudo-dataset to characterize "normal" uncertainty of regression models independent of errors associated with ACS. For each bootstrapped pseudo-dataset, we perform an LME regression with random intercepts by city and by bearing indicators (being Northwest, Southwest, Northeast and Southeast of downtown), as well as random slopes by city on elv, ral and the interaction effect $elv \cdot ral$.

More formally, we describe our model using the following notation. We begin with a dataset of n observations $(y_i, \mathbf{x}_i)$. The bootstrapping process involves drawing, with replacement, an identically-sized set of n pseudo-observations $(y_j^*, \mathbf{x}_j^*)$ from an "original" pseudo-dataset $(y_i^*, \mathbf{x}_i^*)$ with $i \in [1 : n]$.[24] For census tract i, variable of interest y_i, a set of d covariates $\mathbf{x}_i = (x_{i1} \dots x_{id})'$ and corresponding reported standard errors σ_{y_i} and $\sigma_{x_i} = (\sigma_{x_{i1}} \dots \sigma_{x_{id}})'$, pseudo-data are drawn according to $y_i^* \sim N(y_i^*; y_i, \sigma_{y_i}^2)$ and:

$$\begin{aligned} \mathbf{x}_i^* &= (x_{i1}{}^*, x_{i2}{}^* \dots x_{id}{}^*)' \\ &\sim (N(x_{i1}{}^*; x_{i1}, \sigma_{x_{i1}}^2), N(x_{i2}{}^*; x_{i2}, \sigma_{x_{i2}}^2) \dots N(x_{in}{}^*; x_{in}, \sigma_{x_{in}}^2))' \end{aligned} \tag{1}$$

where respective normal distributions are truncated for y_i and $\mathbf{x}_i$ that are percentage or count variables. Variables in the "original" data that lack available

[24] The identical set sizes preserve the original sizes of standard errors. If for some reason it is desirable to artificially inflate standard errors, one can simply draw fewer observations than the size of the original dataset.

error estimates enter directly into the bootstrapping process. For re-sampled observation $(y_j^*, \mathbf{x}_j^*)$ of city k and bearing m, we fit:

$$\hat{y_j^*} = \mathbf{B_{E_k}}\mathbf{E}_j + \mathbf{B_x}\mathbf{x}_j^* + a_k + a_m + s_{y_j} + \varepsilon \tag{2}$$

where $\mathbf{B_{E_k}}$ is a vector of slopes on elv, ral and $elv \cdot ral$ specific to city k, $\mathbf{E}_j = (elv_j, ral_j, elv_j \cdot ral_j)'$, $\mathbf{B_x}$ is a vector of coefficients $(B_1, B_2, ...B_d)$ corresponding to the d covariates, a_k is a fixed effect of being in city k, a_m a fixed effect of having bearing m, and s_{y_j} the spatial autocorrelation factor for observation j and variable of interest y_j under L2 penalization. Variables are transformed for better fit and standardized prior to the regression process.

Compared to a standard, single-model LME regression, our approach is significantly more conservative: it allows for uncertainty within the data itself and is not necessarily confined to equal-tail-ness or normality of standard errors. The effects of observations and covariates of observations that are highly uncertain given the data collection process naturally will be pulled toward zero by the high draw-to-draw variance. Conversely, higher weight is given to observations that are well-collected as they are drawn from tighter distributions. We demonstrate in section B that this method not only outperforms simpler models with the same re-sample bootstrapping procedure, but also presents better estimations of significance than Wald Z-tests via a single LME regression performed on actual ACS estimates.

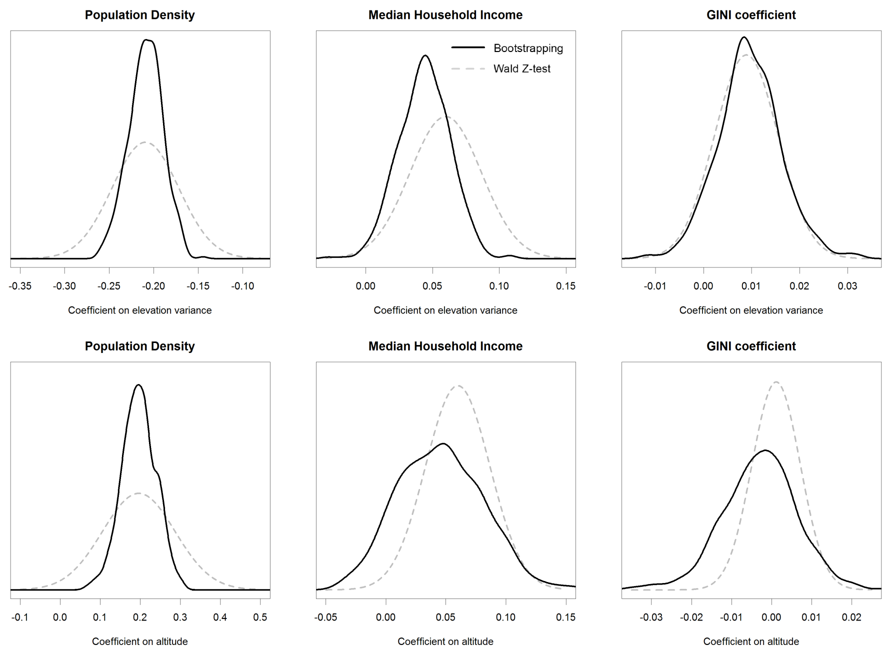
There is yet another key advantage of pseudo-data-bootstrapping: the ability to test for random effects on individual cities. While beyond the scope of this paper, obtaining a set of random effect coefficient draws allows one to perform alternative hypothesis tests on elevation gradient effects of a particular city, test for differences between city-specific effects and the fixed effect, or test for effect differences among multiple cities. We present suggestive evidence of cities responding differently to elevation gradients in section 4, but do not present an in-depth discussion of these issues on a city-by-city basis.

B Model Performance Checks

We perform 500 LME regressions with bootstrapped pseudo-datasets for each of the following three response variables log-transformed: population density, median tract household income and the tract Gini. In all cases elv and is also log-transformed. We first examine the performance of significance tests via bootstrapping compared to fixed-effect standard-error tests with the same LME models, but a single iteration on actual ACS estimates. Figure A1 contrasts densities of bootstrapped coefficient distributions for $\log(elv)$ and ral against estimated Gaussian densities of the same coefficients using the single LME model, obtained using a Wald Z-test.

The comparison is strongly in favor of the pseudo-data-bootstrapping method. Coefficients that have high levels of certainty under the single LME (variance

Figure A1: Coefficient densities from bootstrapping versus single LME model

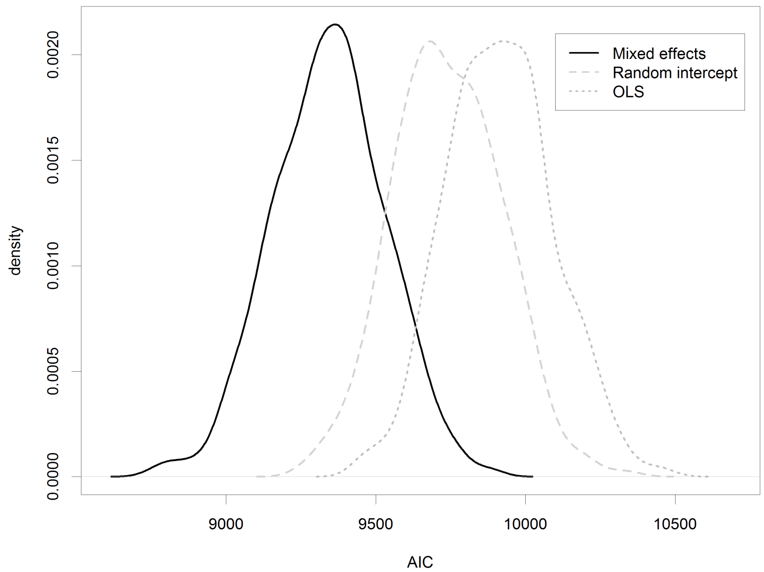


and altitude on density) are pulled into significantly denser distributions under bootstrapping, suggesting an advantage in estimating the "true" effect size given strong evidence for the existence of an effect. Other coefficients lose significance under bootstrapping and are pulled towards zero at the density mean (altitude on median household income and Gini), suggesting that estimations of effects with less certainty of existence are penalized by the two-stage randomization process and correspondingly reduced in size. Effects with high confidence of rejecting the null hypothesis and those with low confidence hence are clearly distinguished.

We further contrast model performance of the full, LME method with random-intercept-only and OLS models, utilizing the same coefficients, pseudo-data and bootstrapping approach. Holding elevation, control variables and the bootstrap procedure constant, we obtain Akaike information criterion (AIC) estimates for 500 iterations of the full LME regression, a reduced random effects regression with intercepts by city and bearing, and a standard OLS regression with no panel information. AIC densities are compared for the population density model in Figure A2; similar comparisons for tract median household income and Gini models are available in Figures A3 and A4.

In all three cases (density, income, Gini), both the LME regression and the

Figure A2: AIC densities of full LME regression versus reduced models, Population density model



random intercept regression clearly outperform OLS according to AIC.[25] This is expected, given that it is highly unlikely that no group fixed effects exist at all across cities. Evidence of superiority of LME over random intercepts is strong for models for population density and tract median household income ($p<0.001$ via unpaired t-tests), and somewhat significant for the tract Gini model ($p=0.0525$).

C Notes for Supplementary Material

We provide summarized regression output for all models used in the paper in the Supplementary Appendix, as well as additional descriptive statistics concerning the dataset. All coefficient sizes are estimated as the simple average of estimated values for all bootstrap iterations. All models are estimated using 500 draws from the full pseudo-dataset. Bootstrap P-values are calculated using a one-sided condition: the number of positive draws as a fraction of 500 for negative coefficient estimates, or the number of negative draws as a fraction of 500 for positive coefficient estimates. If there are less than 5 draws for a given variable that have opposite signs as the estimated coefficient, the notation "$p<0.01$" is used. We also present density plots of bootstrap coefficient values for *elv*, *ral* and the interaction effect. Zero is marked on plots to better visualize the relationship between each elevation effect and the null hypothesis.

[25] $P<0.001$ for all comparisons via unpaired t-tests.

With regard to data formatting, we have elected to simply remove non-complete observations before applying the regression model. Data augmentation is not used for the following two reasons. First, our approach necessarily requires that all ACS variables include reported standard errors for the pseudo-data bootstrapping procedure. Even if the values of missing data can be relatively well-estimated via a first-stage imputation process or similar methods, estimating missing standard errors is conceptually much less straightforward. The concern is that data are usually missing for a specific reason: reported standard errors should be much higher for tracts with missing data should the data actually be present. If estimated standard errors are large, these observations will contribute little given the bootstrapping process if they are included. If estimated standard errors are small, the risk that imputed data are over-utilized and drive the results becomes a concern.

Second, an observation with one missing covariate typically misses a number of other covariates. Quintile mean income is generally available as a group of six covariates, which are either all present or missing. Out of the 271 observations with at least one missing covariate or standard error value, only 11 miss fewer than 10. This suggests that even if multiple imputation is applied, most observations with missing data will require the estimation of a substantial number of covariates to be usable. This issue significantly lowers the quality of the observations with imputed data. Since at no point in our bootstrapping approach do we explicitly use the "full" dataset as collected (unless the exact same set of observations is drawn at random), benefits of performing imputation are limited to small gains to the representativeness of the data structure and a marginally larger amount of information from which inferences are drawn. We hence conclude that the potential costs of imputation outweigh the benefits.

Note that unless the coefficients on ral and $elv \cdot ral$ point in the same direction, a positive or negative sign on ral does not necessarily suggest a correspondingly positive or negative altitude effect for any given response variable. Given a dataset-average $\ln(elv)$ of $\ln(8.1) \approx 2.09$, the coefficient on ral must be at least twice as large in absolute value to overwhelm an opposite-direction coefficient on $elv \cdot ral$ at dataset-average elevation variance. It is however straightforward to draw inferences on elv at city-average altitude because $ral = 0$ at the city average by design.

The estimation of R^2 values for linear mixed effects models remains a topic of debate in modern statistics. Here we opt for a two-estimator approach outlined by Nakagawa and Schielzeth (2013), which presents for each model a fixed effects (marginal) R^2 value and a fixed-and-random effects (conditional) R^2 value. The former, given our fully Gaussian specification, can simply be regarded as being equivalent to the R^2 value of the OLS model created by removing all random effects from the LME. The latter includes variance explained by the former as well as variance explained by random intercepts. Variance for random slopes are not included as it is somewhat difficult to integrate such components with regular notions of statistical explanatory power (Snijders and Bosker 2011).

There are two advantages of presenting both marginal and conditional R^2 values. First, the complementary nature of the values reduces issues that may arise from presenting only one out of the two. Given a large number of groups relative to the dataset size, the conditional R^2 value will be dominated by between-group variance and fail to adequately reflect explanatory power of fixed effects: in an extreme case, when the group size is very large yet observation count per group is small, the conditional R^2 will be close to 1 no matter how many fixed effects terms are included. Presenting the marginal R^2 allows the explanatory power of fixed effects to be examined independent of random effects, yet we would have no knowledge of the role of the random intercept if only the marginal R^2 were presented. Second, visualizing both R^2 distributions through bootstrapping allows us to obtain non-quantitative, yet nonetheless useful, information about the importance of between-group variance relative to within-group variance for each model. In the Supplementary Appendix, we provide side-by-side density plots for the two R^2 values for each model. Through bootstrap densities, it is possible to not only draw inferences on general explanatory performance of a given model, but also on how *consistent* performance is with regard to random variations in the data.

Note that in some cases, the reported marginal R^2 is expected to be extremely low by virtue of the model parameters. This is the case for microclimate effects where differences in annual and monthly temperature variance between cities is naturally far greater than differences across various areas of the same city. Similarly, a fixed-effects-only model of proportions of African American residents that does not control for spatial autocorrelation is expected to have low explanatory power. Since we do not actually draw inferences or present coefficient values based on any fixed-effects-only estimates in the paper, this is not an issue of concern.

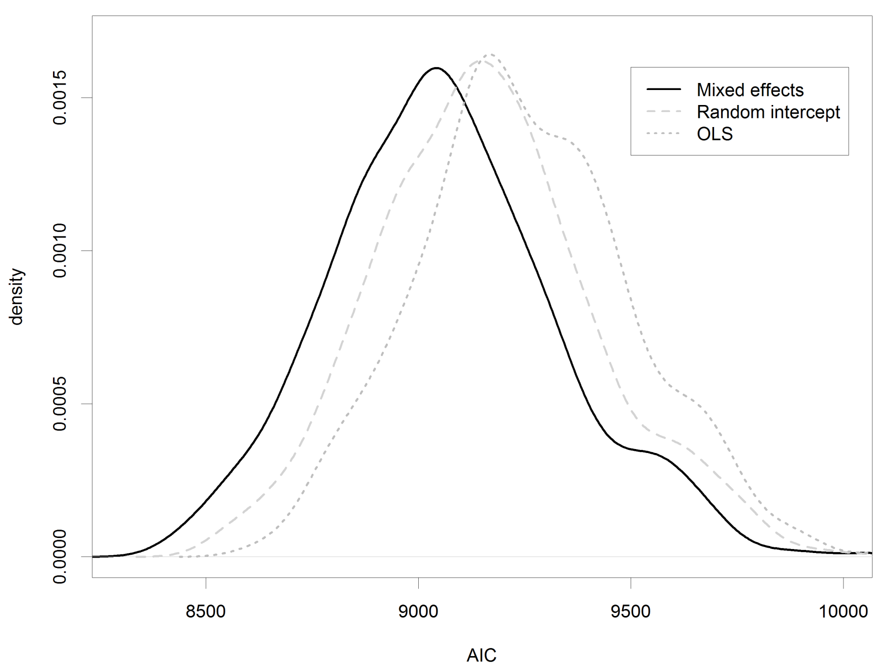
One anomaly that we note in the data is the unusually low rape incident rate of Los Angeles. (0.004 incidents per thousand persons). This is not because of a lack of crime observations for the Los Angeles area: crime data are available for 424 out of the 796 dataset LA census tracts. Neither is the issue likely caused by missing data: most census tracts with crime data have zero rape incidents during the period covered by the NNCS survey. The distinction appears to be that a far greater percentage of the LA tracts with crime data lack rape incidents. We do not know if this is simply a result of chance given the overlap between our elevation data and NNCS, or if there were potential data collection issues with the original study. However, we have performed the same regression analysis on rape incident rates in section 6.3 with data from LA omitted. Results are similar: the coefficient on *elv* changes to -0.029 from -0.022 and the corresponding p-value to 0.015 from 0.02. Because conclusions remain similar (foot crime effects are still far greater than either the effect on rape or murder), we have reported figures with LA data included in the paper for consistency.

Tables and Figures

Table A1: Correlation and Pearson product-moment significance tests between elevation effects and select census tract statistics

	Population density (log)	Median household income (log)	Tract Gini (log)	% of public transit commuters (log)	Poverty rate (log)
elv	-0.55	0.187	-0.076	-0.344	-0.188
PPMCC	<0.001	<0.001	<0.001	<0.001	<0.001
ral	-0.03	0.131	-0.062	-0.044	-0.142
PPMCC	0.01	<0.001	<0.001	<0.001	<0.001

Figure A3: AIC densities of full LME regression versus reduced models, median household income model



	Elevation (mixed effects by city)[a]	Elevation (Global Slope)[b]	Spatial factors[c]	Population density[d]	Median income[e]	Tract Gini[f]	Mean quintile incomes[g]	Public transit & Walk[h]	Congestion[i]	Autocorrelation[j]	Random slopes (Bearing /City)	Income Indicators[k]	Demographics[l]	Weather[m]
Population density[n]	*		*							*	*			
Median household income	*		*	*						*	*			
Tract Gini	*		*	*	*					*	*			
Mean quintile income		*	*	*						*	*			
View Effects		*	*	*	*	*	*	*	*	*	*	*	*	*
Public Transit&Walk		*	*	*	*	*	*			*	*	*	*	
Crime		*	*	*	*	*	*			*	*	*	*	
Congestion		*	*	*	*	*	*			*	*	*	*	

Table A2: Summary of variables used in regression models

[a] elv,ral, and $elv \cdot ral$ with random slope effects are included. elv and interaction effect are log-transformed.

[b] elv,ral, and $elv \cdot ral$ without random slope effects.elv is log-transformed.

[c] Spatial factors include log water area in tract,log driving time to downtown for no congestion, 7am and 7pm scenarios, log driving distance to downtown, $log(waterarea) \cdot elv$, total housing unit counts, log tract land area, and log distance to coastline.

[d] Population density is log-transformed.

[e] Median household income is log-transformed.

[f] Tract Gini coefficient is log-transformed.

[g] Mean income for five quintiles and mean income for top 5% of high-earning households. Estimates are log-transformed.

[h] Percentage of workers who use public transit and percentage of workers who walk to work.

[i] Congestion Score as estimated using REST data on a 0-100 scale.

[j] Log spatial autocorrelation is controlled for in all models. L2-penalization is stratified by city.

[k] Income indicators include housing occupancy rate, log unemployment, and log poverty rate.

[l] Demographic variables include percentage of Asian Americans, percentage of African-Americans, percentage of individuals over the age of 64, average household size, average household size of married couples, percentage of bachelor's degree holders, and percentage of graduate degree holders. All variables are log-transformed except household size variables.

[m] Weather variables include Annual Cooling Degree Days, standard deviation of monthly-average temperatures from 1971 to 2000, and the average monthly maximum-minimum temperature spread. All variables are log-transformed.

[n] All dependent variables except congestion score are log-transformed. View effects are estimated by the proportion of households with income level greater than $200,000 and $150,000, respectively. Crime rates and autocorrelation of crime estimates are based on NNCS data for available tracts.

	Elevation (mixed effects by city)	Elevation (Global Slope)	Spatial factors	Population density	Median income	Tract Gini	Mean quintile incomes	Public transit & Walk	Congestion	Autocorrelation	Random slopes (Bearing /City)	Income Indicators	Demographics	Weather
Local weather[a]		*	*	*							*			
African-American Proportion (Full)		*	*	*						*	*			
African-American Proportion (Reduced)		*	*								*			
AFAM Proportion (further reduced)		*									*			
Household Size[b]		*	*	*						*	*			
Marital Status[c]		*	*	*						*	*			
Median Age		*	*	*						*	*			
Elderly proportion[d]		*	*	*						*	*			
Education[e]		*	*	*						*	*			

Table A3: Summary of variables used in regression models (continued)

[a]Variables include Annual Cooling Degree Days, standard deviation of monthly-average temperatures from 1971 to 2000, and the average monthly maximum-minimum temperature spread. All variables are log-transformed. Variables below are log-transformed unless otherwise noted.

[b]Household size variables include size for all households and household size for married couples.

[c]Estimations of the percentage of tract population that is married are used.

[d]Elderly individuals defined as those about the age of 64.

[e]Variables include percentage of individuals with Bachelor's degrees (non-associate) and percentage of individuals with graduate and professional degrees.

Figure A4: AIC densities of full LME regression versus reduced models, Gini coefficient model

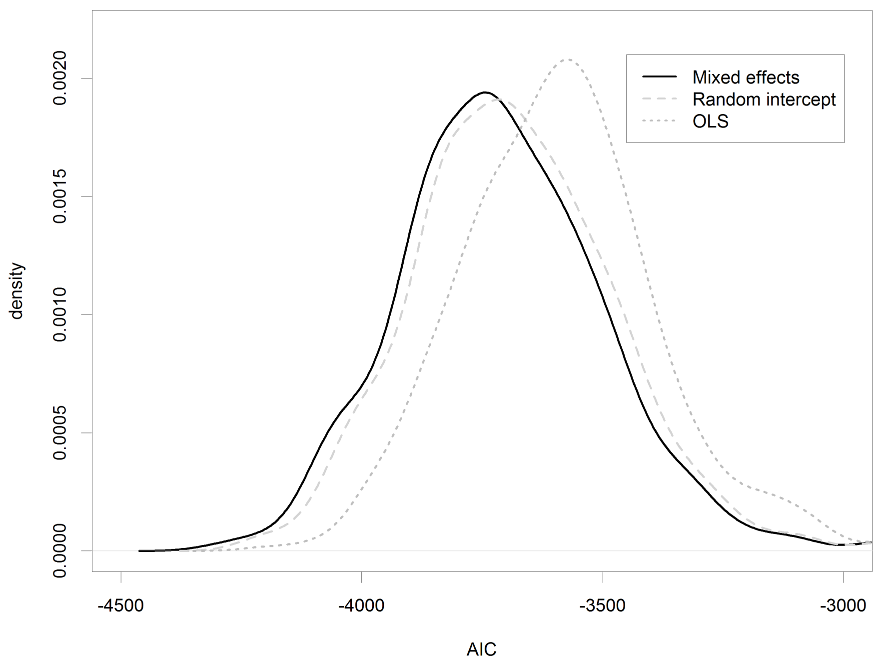


Figure A5: Distributions of random effect coefficient estimates of *ral* by city, population density model

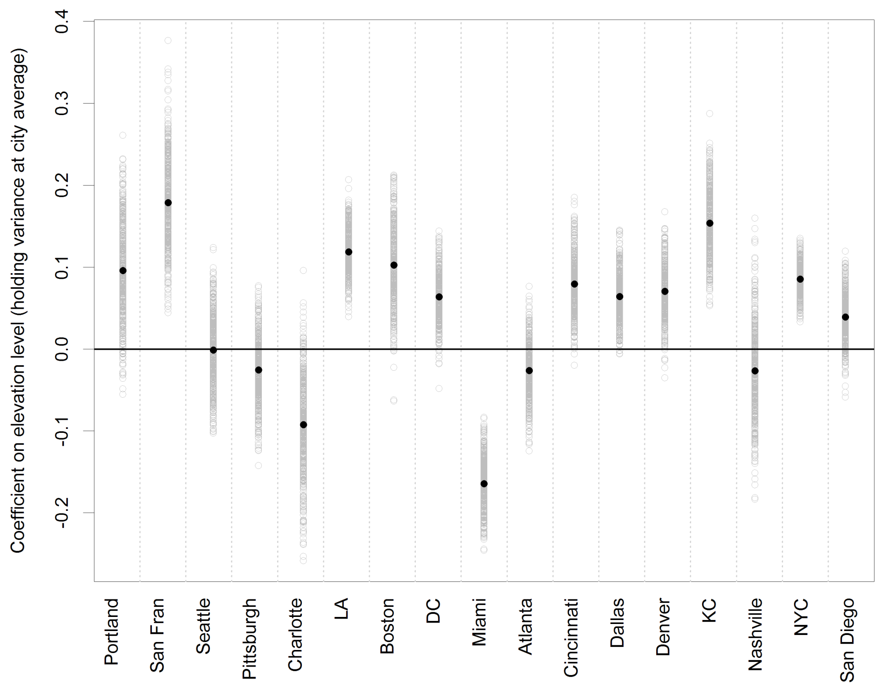


Figure A6: Distributions of random effect coefficient estimates of log(elv) by city, median household income model

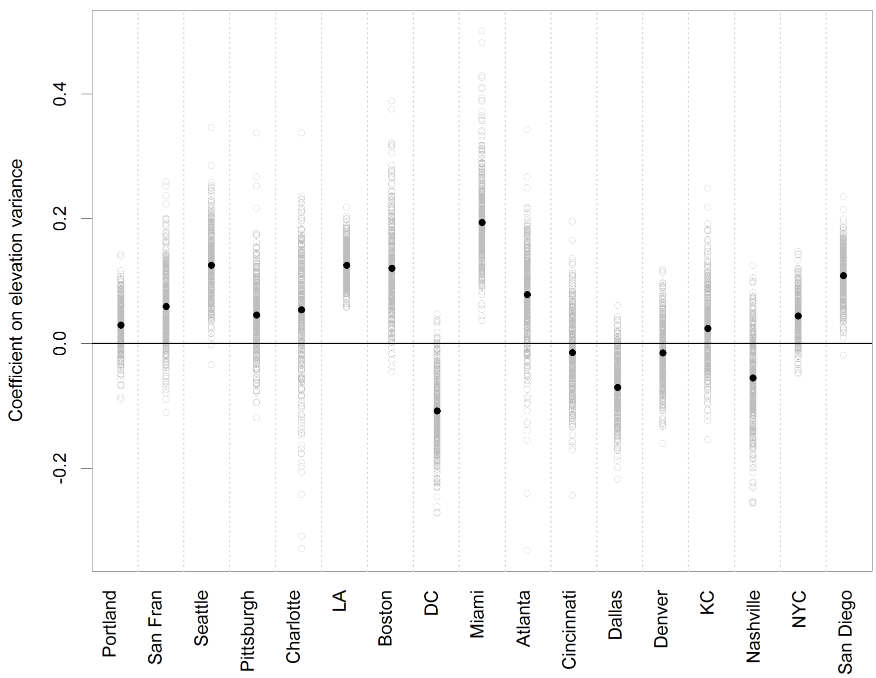


Figure A7: Distributions of random effect coefficient estimates of log(elv) by city, gini coefficient model

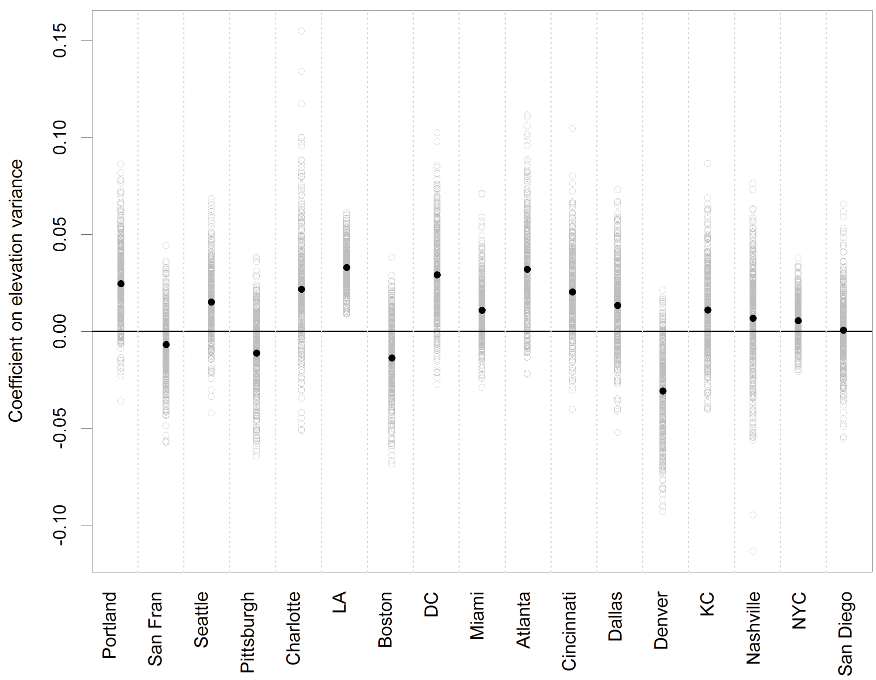


Figure A8: elv-population density and ral-population density curves, holding ral and elv respectively constant

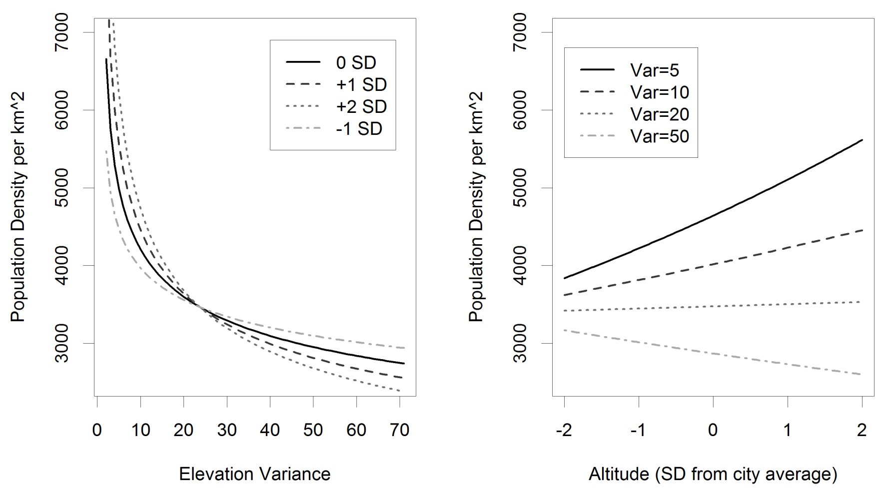


Figure A9: DEM contour density plots of select dataset cities