Recipes for neighborhood development: A machine learning approach toward understanding the impact of mixing in neighborhoods

John R. Hipp*

Kevin Kane

Jae Hong Kim

March 16, 2017

Post-print. Published in Landscape and Urban Planning 164 (2017)

Word count: 7,993

Word count (including references): 9,132

Running Head: “Machine Learning and household income appreciation”

* Department of Criminology, Law and Society and Department of Sociology, University of California, Irvine. Address correspondence to John R. Hipp, Department of Criminology, Law and Society, University of California, Irvine, 3311 Social Ecology II, Irvine, CA 92697; email: john.hipp@UCI.edu. This research is supported in part by the Metropolitan Futures Initiative

Recipes for neighborhood development: A machine learning approach toward understanding the impact of mixing in neighborhoods

Abstract

Scholars of New Urbanism have suggested that mixing along various dimensions in neighborhoods (e.g., income, race/ethnicity, land use) may have positive consequences for neighborhoods, particularly for economic dynamism. A challenge for empirically assessing this hypothesis is that the impact of mixing may depend on various socio-demographic characteristics of the neighborhood and take place in a complex fashion that cannot be appropriately handled by traditional statistical analytical approaches. We utilize a rarely used, innovative estimation technique—kernel regularized least squares—that allows for nonparametric estimation of the relationship between various neighborhood characteristics in 2000 and the change in average household income in the neighborhood from 2000 to 2010. The results demonstrate that the relationships between average income growth and both income mixing and racial/ethnic mixing are contingent upon several neighborhood socio-demographic “ingredients.” Racial mixing, for example, is found to be positively associated with average income over time when it occurs in neighborhoods with a high percentage of Latinos or immigrants, high population density, or high housing age mixing. Income mixing is associated with worsening average household income in neighborhoods with more poverty, unemployment, immigrants or population density. It appears that considering the broader characteristics of the neighborhood are important for understanding economic dynamism.

Keywords: neighborhoods, household incomes, data mining.

Bio

John R. Hipp is a Professor in the departments of Criminology, Law and Society, and Sociology, at the University of California Irvine. His research interests focus on how neighborhoods change over time, how that change both affects and is affected by neighborhood crime, and the role networks and institutions play in that change. He approaches these questions using quantitative methods as well as social network analysis. He has published substantive work in such journals as *American Sociological Review, Criminology, Social Forces*, *Social Problems, Mobilization, City & Community, Urban Studies* and *Journal of Urban Affairs*. He has published methodological work in such journals as *Sociological Methodology*, *Psychological Methods*, and *Structural Equation Modeling*.

Jae Hong Kim is an Assistant Professor in the Department of Planning, Policy and Design at the University of California, Irvine. His research focuses on urban land use change, economic development, and dynamic system modeling. He has published his work in *Journal of Planning Literature*, *Urban Studies*, and other planning-related journals.

Kevin Kane, Ph.D., is a Postdoctoral Research Fellow in the Department of Planning, Policy and Design at the University of California, Irvine. He is an economic geographer interested in the quantitative spatial analysis of urban land-use change and urban development patterns, municipal governance, institutions, and economic development. His research uses land change as an outcome measure – in the form of changes to the built environment, shifting patterns of employment, or the socioeconomic composition of places – and links these to drivers of change including policy, structural economic shifts, or preferences for how we use and travel across urban space.

Recipes for neighborhood development: A machine learning approach toward understanding the impact of mixing in neighborhoods

There is a long-standing interest in understanding the economic dynamism of neighborhoods (Galster, Hayes, and Johnson 2005; Temkin and Rohe 1996). Scholars have noted that although many neighborhoods maintain relative economic stability over time as measured by the average income of residents, smaller numbers of neighborhoods either experience economic declines over time or exceptional growth. Various theories have also been proposed to explain changes in neighborhoods, particularly as measured by the average level of income of residents. Among others, recently the New Urbanism perspective has emphasized the possible positive role of mixing along various dimensions for bringing about economic dynamism (Calthorpe 1993; Calthorpe and Fulton 2001). Specifically, it has been suggested that mixing based on land use or building age, or mixing based on such socio-demographic characteristics of residents as income or race/ethnicity, can have positive consequences for neighborhoods (Knaap 2005).

A significant challenge, both theoretically and empirically, for studies in the New Urbanism tradition is that mixing along various dimensions may not have uniform consequences for neighborhoods depending on the particular context. For example, it is unclear whether combining different types of mixing (such as land use mixing, income mixing, etc.) in the same neighborhood will have similar consequences as when just one of these dimensions of mixing is present. Some language in the New Urbanism literature implies that there may be synergistic qualities from combining different types of mixing (Knaap 2005; Roberts 2007), however, some studies have found cautionary evidence calling this into question (Chapple and Jacobus 2009).

Furthermore, mixing based on various dimensions may have different consequences for the neighborhood depending on the socio-economic context, or the socio-demographic context. Certain dimensions of mixing may negatively impact economic dynamism when they occur in economically challenged neighborhoods.

The possibility that the impact of mixing on economic dynamism in a neighborhood can be moderated (or amplified) by various contextual factors or other dimensions of mixing has received limited empirical assessment in the literature, arguably because of the methodological difficulty of addressing such a question. These possible moderating effects of the context for mixing imply the need for an analysis that includes a large number of multiplicative interactions when adopting the traditional modeling strategy. We instead address these questions with an existing machine learning technique that we argue is perfectly suited to these research questions. The Kernel Regularized Least Squares (KRLS) estimation approach, described in more detail below, allows us to flexibly assess nonlinear moderating effects among our variables of interest. We can assess whether the relationship between four dimensions of mixing – income, racial, housing age, and land use mix – and average income appreciation in neighborhoods exhibit nonlinear interaction patterns. We next describe theories of neighborhood change, particularly focusing on the importance of mixing along various dimensions for economic dynamism.

Literature Review

Theories explaining neighborhood change

A body of literature has explored how neighborhoods change over time, specifically how they change regarding their socio-economic resources. Whereas early research focused on human ecology theory in which neighborhoods operate in a larger system (Park, Burgess, and

McKenzie 1925), later research turned to subcultural theory which argued for important non-economic factors in neighborhoods. (Pitken 2001). In the 1970s the political economy approach gained in prominence and focused directly on the social relations of production and accumulation in which elites drove the economic processes (Molotch 1976). Studies have empirically explored the relationship between various neighborhood characteristics and change in neighborhood income (Ellen and O'Regan 2008; Jun 2016; Rosenthal 2008).

More recently, there has been a rise in a perspective broadly characterized as New Urbanism. The New Urbanism perspective can be traced to the founding of the Congress for the New Urbanism in 1993 by a group of architects and planners (Leccese and McCormick 1999). New Urbanist design theory focuses on creating neighborhoods and cities that foster a “sense of community” by organizing neighborhoods with diversity in use and population (Talen 1999; Talen 2013). A primary design element of New Urbanism is high density, mixed use development to create vibrant public spaces (Calthorpe 1993; Calthorpe and Fulton 2001). A challenge is that density can come in different forms (Campoli 2012; Campoli and MacLean 2007). In particular, mixing land uses, such as “jobs, housing, and food outlets, cross walks, bike racks” (Campoli 2012) has been advocated as an effective means to promote social interaction, neighborhood vibrancy, and thus scholars have concluded that communities with a high density of population *and* a mix of several land uses can help bring about this vibrancy. This implies considering the simultaneous impact of different types of mixing, an issue to which we turn next.

How mixing can help neighborhood dynamism

The desire for and emphasis on mixed neighborhoods, arguably, was born from the failure of public housing projects and the thinking that mixing might help the recipients of public housing (overwhelmingly low-income, poorly-educated urban minorities) to avoid the pitfalls of concentrated poverty and socioeconomic disadvantage. Socioeconomic mixing – particularly along income lines – is thought to promote social and economic integration as well as increased opportunities for low-income residents (Wilson 1987). The positive idea of mixing is also linked to the more recent demographic trend of urban inversion and downtown renewal, whereby larger populations (most notably young adults or retirees) are moving "back" to central city neighborhoods (Ehrenhalt 2012).

There is evidence that mixing income of residents may have positive consequences for neighborhoods. A body of research has focused on how mixed income areas can have various positive consequences for the lower income households living in such neighborhoods, including possible improved social networks for job contacts leading to better employment outcomes, mental health benefits, increased self-esteem, and behavioral and health improvements for children (for a review of this literature see Levy, McDade, and Dumlao 2010). There are also proposed advantages for the neighborhood as a whole, including improved social control to address safety issues given that higher income residents might provide particular norms to increase safety (Fraser and Nelson 2008) or economic advantages by increasing market demand for higher-quality goods and services that can then be enjoyed by all residents (Levy, McDade, and Dumlao 2010). Nonetheless, there is also a possible long-term side effect in which income mixing brings about gentrification, which then can lead to increased income segregation over time, as was found in a study of rural settings (Golding 2015).

The mixing of land uses, namely the accessibility of workplaces, schools, retail, and other services to residential areas follows a similarly-renewed emphasis on walkability. Much of this comes from the New Urbanist and Smart Growth movements that began in earnest during the 1990s (Knaap 2005). A mixing of land uses can increase social interaction and decrease the need for long-distance transportation and thus cut carbon emissions. By putting jobs and housing close to each other, mixing land uses can also lead to better job outcomes, and hence economic dynamism; indeed, a study of Chicago found that a greater number of jobs within two miles of neighborhoods led to higher employment and lower unemployment rates for residents (Immergluck 1998).

Mixing is also related to gentrification, or the inflow of capital into a neighborhood. While increasing property values and vibrant communities are generally seen as positive outcomes, gentrification can also displace an area's original resident – and business – populations, raising the question of who is the recipient of neighborhood improvements (Newman and Wyly 2006). Some believe social mixing policies to be veiled attempts at gentrification with minimal impact on upward mobility of struggling communities (Bridge, Butler, and Lees 2012). Thus, although we will focus on average income appreciation in neighborhoods in this study, a caution to be heeded in all such studies is that it sidesteps the question of residential displacement. Similar to land-use mixing, urbanist Jane Jacobs (1961) was a strong advocate for a mixing of ages of buildings in a neighborhood. She argued that older buildings, being less expensive to rent, present a point of entry into a community for residents or businesses and allow for them to co-exist with the tenants and owners of newer, expensive buildings. A number of cities who are keen to promote downtown renewal (Charlotte, NC being one example – see Ehrenhalt (2012)) have found their lack of a historic building stock challenging since newer space is more expensive, and less flexible in terms of use, leasing, and ownership. Although there is some evidence that older housing has a discount rate, perhaps due to being a proxy for the quality of housing (Rubin 1993), the mix of housing age may allow for income mixing and the proposed positive consequences.

Recent scholarship has posited that racial/ethnic mixing in neighborhoods might signal a multi-cultural environment that is desirable to certain segments of the population. Florida (2002) in particular emphasizes longer-term benefits of such openness, arguing that creative places "were open, diverse, and culturally creative first. Then they *became* technologically creative and subsequently gave rise to new high-tech firms and industries" (p. 207). Cultural amenities, a vibe, and a buzz, in his view, often flow from an area's original openness to diversity, mixing, and ultimately new ideas, whether at the metropolitan or neighborhood level. For example, the presence of a multi-cultural population, along with an accompanying wide variety of ethnic restaurants may be highly desirable for certain demographic groups. Such areas may also foster a vibrant music or arts scene, as well as multicultural festivals and events that appeal to "hipsters" and lead to more economic dynamism in such neighborhoods. As evidence of the economic stagnation of neighborhoods without such characteristics, a study of Baltimore inner-ring suburbs pointed to racial segregation, as well as labor market restructuring and income segregation, as important drivers of neighborhood decline (Hanlon and Vicino 2007).

How mixing might hinder neighborhood dynamism

Although advocacy for mixing is largely a reaction to the perceived negative outcomes of homogeneity or segregation, there can be benefits to certain types of segregation in cities. Zoning codes largely exist to guard homes against the noxious fumes of industry or late-night noise of restaurants and bars, for example. The 1916 U.S. Supreme Court case which is credited with legalizing municipal police power (*Euclid v. Ambler*) makes explicit that land uses in conflict can be separated through zoning to avoid possible negative externalities and decreased property values of individual homeowners (Hall 2007).

Furthermore, there are reasons that neighborhoods may not necessarily thrive economically due to a mix of residential and business land uses. Retail businesses – especially national chains with well-developed and clearly defined product types and target markets – may have difficulty thriving in mixed areas. The local customer base is too varied, while customers are generally drawn to well-known areas which offer scale economies and a variety of retail options (Chapple and Jacobus 2009). Property crime rates may even be higher in mixed areas (Hipp 2007), and this crime, or the perception of it, can be a deterrent for both retailers and their potential customers (Hipp 2010a).

Whereas a growing number of studies in the literature presume that racial/ethnic mixing will be desirable for reasons already discussed, there are countervailing reasons why that may not be the case. For example, the presumption that there will be social ties spanning racial/ethnic groups is questionable, as studies have found that there are fewer social ties in general in such neighborhoods (Lowenkamp, Cullen, and Pratt 2003; Warner and Rountree 1997), less neighborhood attachment (Sampson 1991), and less neighborhood satisfaction (Hipp 2009; Sampson 1991). Given the consistent evidence that neighborhoods with higher levels of racial/ethnic heterogeneity have higher levels of crime (Hipp 2007; Roncek and Maier 1991; Rountree and Warner 1999; Sampson and Groves 1989), this provides additional evidence that such neighborhoods may not always exhibit economic vibrancy as expected. Indeed, studies have found that racial change is related to decreasing household income (Baxter and Lauria 2000). A recent investigation of the 100 largest US metropolitan areas by Jun (2016) also reported a strong negative association between the share of non-White population and the change in neighborhood per capita income.

There are also reasons to suspect that income mixing will not necessarily lead to positive neighborhood outcomes. For example, there is evidence that social ties do not necessarily cross income levels in mixed income neighborhoods. A study of a Hope VI site in Seattle found that social ties tended to not cross income differences, even in an award-winning mixed income development (Kleit 2005). A study of a New Urbanist mixed income community in North Carolina also found that income differences reduced the probability of forming a social tie, even controlling for the spatial distance between housing units (Hipp and Perrin 2009). And the evidence that mixed income neighborhoods tend to have higher levels of crime also calls into question the presumption that they will have long-term beneficial consequences (Hipp 2007; Hipp and Boessen 2013; Messner and Tardiff 1986). One review of existing mixed-income developments concluded that there is a need for a land use design that encourages the actual social mixing of residents of different income levels, implying that it is a *combination* of income mixing along with land use mixing that may be important for neighborhood outcomes (Roberts 2007). We therefore next turn to a discussion of the need to consider some of these measures of mixing in combination, rather than as distinct measures.

Considering the interdependence of mixing dimensions

The challenges for studies of neighborhood change are twofold. First, whereas theories posit that certain structural characteristics will have either positive or negative impacts on the socio-economic change in a neighborhood over time, they rarely specify the functional form of the true relationship that should be expected. As a consequence, studies typically only test for possible linear (or linearized) relationships between posited important structural characteristics and the socio-economic dynamics of the neighborhood. There are theoretical reasons to posit that some of these processes may not play out in a linear fashion, but rather exhibit threshold effects (Schelling 1971). There is also evidence that neighborhoods do not simply respond to exogenous shocks in a consistent, linear fashion (Galster, Cutsinger, and Lim 2007). For these reasons, there is a need to assess possible nonlinear or threshold functions that might characterize the relationship between these measures and neighborhood economic dynamism.

Second, a challenge is that the structural characteristics of neighborhoods are likely not independent of one another, but rather highly interdependent. Thus, the typical assumption of linear statistical modeling that we can "hold constant" one measure while manipulating another is fine in principle, but it is likely not reasonable in practice when studying neighborhood dynamism. To understand how neighborhoods can change over time, it is likely that we need to understand how various structural characteristics of neighborhoods might operate in tandem to impact neighborhood change trajectories. For example, a study of neighborhoods in Canada concluded that a number of factors were important for explaining neighborhood economic dynamics, ranging from local conditions to wider economic and policy shifts (Kitchen and Williams 2009).

The machine learning technique that this paper adopts, Kernel Regularized Least Squares (KRLS), directly addresses these two challenges. KRLS' nonparametric estimation of covariate effects helps isolate the structural measures impacting neighborhood change, while providing the marginal effects of each independent variable across the covariate space allows for a better identification of threshold effects than a pointwise, linear estimate. Most importantly, the marginal effects can be regressed upon the other variables in the model, allowing for us to determine which "ingredients" of mixing result in greater economic dynamism in neighborhoods.

We focus on several factors that may moderate the relationship between mixing and average income growth (the measure of neighborhood economic dynamism used in this study).

First, mixing may have differential consequences depending on the socio-economic status (SES) of the neighborhood at the beginning of the decade. Mixing that occurs in the context of more economically disadvantaged neighborhoods may be less likely to have the anticipated positive consequences for income growth. Second, high population density locations are more in the spirit of New Urbanist principles, and therefore mixing that occurs in these contexts may be more beneficial for income growth. Third, given that residential instability may be a sign of a neighborhood in flux, mixing in such contexts may have negative consequences. Fourth, if the presence of more racial minorities or immigrants brings more potential vitality to a neighborhood, the presence of more mixing in such contexts may have stronger positive consequences on income growth. Finally, mixing may be most beneficial for the dynamism of New Urbanist neighborhoods when it occurs in a context in which the age structure contains a relatively smaller number of households with children. We describe our statistical approach next.

Data and methods

Data

The study area is the 5-county area comprising Southern California, a large region with a population of about 17 million. The Southern California region is an ideal setting for this study because: a) it is the prototypical example of a booming Sunbelt area that is characterized by rapid population growth and a sprawled pattern of urban development; b) it nonetheless contains numerous highly concentrated, historically-embedded neighborhoods where compact growth is increasingly popular; and c) it is a racially and ethnically heterogeneous area with considerable racial/ethnic mixing.

The socio-demographic data come from the 2000 U.S. Census and the American Community Survey (2010-2014 5-year estimates). Land use data come from the Southern California Association of Governments, a regional planning authority. We used census tracts to represent neighborhoods. Our outcome variable captures the change in average household incomes from 2000 to 2012, and our independent variables are all measured in 2000 (land use data is measured in 2001). Thus, we are asking what neighborhood measures in 2000 explain greater increases in reported household incomes over the subsequent 12 years. In this study we focus on this relatively shorter period of neighborhood change over a single decade; a longer period is outside the scope of the present study and will instead be the focus of our future work.

Dependent variable

The outcome variable is the change in the reported household incomes (logged) between 2000 and 2012 (based on the 2010-14 ACS 5-year estimates). We harmonized the data to 2010 tract boundaries (apportioning the 2000 data based on the population-weighted overlap with 2010 boundaries), log transformed the average household incomes at each time point, and then computed the difference over the decade. We use average income rather than median income since the need to harmonize aggregated data to 2010 tract boundaries makes it impossible to calculate the median income in 2000. Thus, we are capturing the percentage change in average household incomes over the decade for each tract.

Independent variables

Our key measures of interest capture different types of mixing. We used the entropy index to measure the relative level of mixing for most of our dimensions of mixing; this captures the relative proportion of each category (Massey and Denton 1988). Entropy has been widely adopted as a mixing measure—for example, using it to assess the relationship between land use mixing and housing values (Song and Knaap 2004). Values range from 0 to 1, and a higher value indicates higher mixing.

We constructed measures of *race entropy*, *housing age entropy*, *land use entropy*, and *household income inequality*. Table 1 describes the categories used in the three entropy measures. Given that income inequality is a continuous measure we constructed it as a Gini coefficient based on the household income category bins reported to the U.S. Census. The Gini coefficient is a common measure of income inequality (i.e., a proxy of income mixing within a geographic area) that uses cumulative earnings at each percentile of the income distribution to develop a continuous measure of income inequality by area. This was computed with the prln.exe software program developed by Francois Nielsen (available at http://www.unc.edu/~nielsen/data/data.htm). We refer to these measures as “mixing” throughout the results section.

<<<Table 1 about here>>>

We also included several socio-demographic variables that likely impact the change in household incomes in a neighborhood over the subsequent decade. We account for the *average household incomes* at the beginning of the decade, log transformed. Given that a higher concentration of owner-occupied units may increase household incomes in a neighborhood, we included a measure of the *percent homeowner*s. We account for the racial/ethnic composition of the neighborhood with measures of *percent black* and *percent Latino*. We included a measure of *percent immigrants* to account for the possibility that this group may have a negative or positive impact on household income appreciation. The presence of more residential stability in a neighborhood might reflect greater satisfaction and cohesion, and we therefore included a measure of the *average length of residence*. The economic vibrancy of an area can impact the trajectory of household incomes, and we capture this with a measure of the *unemployment rate*. Likewise, neighborhoods with higher vacancy rates will likely depress household incomes, and we therefore included a measure of the *percent occupied units*. We account for the age composition of the neighborhood with two measures of retirees and children: *percent aged 65 and above*, and *percent less than 20 years of age*. We included a measure of *population density* to account for the competing views of whether this measure has a positive or negative impact on household income growth. We also control for the *percent residential land.* Residential land includes single-family and multi-family housing as a proportion of all urbanized land. Finally, we accounted for the *percent open land*. Open land includes the share of land area that is in urban recreational use such as parks and golf courses as well as non-urbanized uses such as natural areas and vacant space which indicate the share of unbuilt area in a tract.

We also account for possible effects from nearby neighborhoods. For each census tract, we used a GIS to identify all other tracts whose centroids lie within five miles. Characteristics of each tract's surrounding neighborhood were calculated using an inverse distance decay function that weights nearby tracts heavily, while those further away (up to five miles) were weighted less. The summary statistics for the variables used in the analyses are shown in Table 2.

<<<Table 2 about here>>>

Methods

To capture possible nonlinearities and nonlinear interactions among the covariates explaining the change in household incomes over the subsequent decade, we used a relatively new analytic technique: Kernel-based regularized least squares (KRLS) described in (Hainmueller and Hazlett 2014) and implemented for Stata in (Ferwerda, Hainmueller, and Hazlett 2013). KRLS comes out of the machine learning literature, and builds on techniques developed in the 1990s. The KRLS approach provides estimates of the marginal effects of each independent variable at each data point in the covariate space and provides closed-form estimates of the pointwise partial derivatives. To avoid over-fitting, the function minimizes squared loss, and prefers smoother functions (by reducing complexity in the optimal solution). KRLS enables us to nonparametrically estimate the relationship between all of our covariates and the outcome variable, and considers their (nonparametric) interactions in the analysis.

KRLS analyses provide estimates of the marginal coefficient for each case in the sample (that is, the derivatives of this relationship). We can then assess whether these derivative estimates are systematically related to other variables in the model. We accomplished this by regressing these derivative estimates for each variable on each other variable in the model (the original variable, a squared version, and a cubic version to capture nonlinearities) one at a time and assessed the amount of variance explained. The R-square of these regressions captures the degree to which the effect of a measure on the outcome differs based on values of the explanatory variable (i.e., interaction effects), and we found that R-squared values of at least .10 typically captured relationships of substantive interest, and we explore these in the results section. Note that when these derivatives are strongly related to other variables in the model (as captured by a high R-square), these are implied interaction effects. Most relationships were suitably captured by a quadratic specification, although a few were substantially improved by the cubic specification; Table A1 in the Appendix displays the R-square values for all interactions. We then plotted these interactions between the derivatives and a variable that exhibited a substantial relationship using Lowess regression to capture any and all nonlinearities—which groups observations with similar covariate values (Cleveland 1979)—and we plot the predicted values from these in the figures. As can be seen in these figures, another advantage of KRLS over OLS is that it is not constrained to linear or linearized interactions, but rather can capture nonlinear interactions that need not have a parametric form. Nonetheless, we also estimated an OLS model using Stata 13.1 as a comparison to the KRLS estimates. Finally, there is little evidence of spatial correlation in our residuals: whereas the Moran's I for the outcome variable is .09, the value for the residuals is just .03, implying that our model explains nearly all of the spatial patterning.

Results

Table 3 presents the results from the KRLS analyses and the OLS analyses (for comparison) for the relationship between the neighborhood characteristics in 2000 and the change in logged household incomes from 2000 to 2012. Note that the first column shows the averages of the pointwise derivatives of a variable on the change in household incomes over the decade for the KRLS results. However, this effect can vary over each observation, and this is shown in three subsequent columns of the KRLS results that list the 25^{th}, 50^{th}, and 75^{th} percentile values for this marginal effect. For example, we see that whereas racial mixing has an average positive coefficient of .022, there is much variability in this coefficient ranging from negative (-.018) at the 25^{th} percentile of the coefficient to positive (.062) at the 75^{th} percentile.

<<<Table 3 about here>>>

To get a sense of the magnitude of these effects, the "std" column shows the change in average income over the subsequent decade for a one standard deviation change in the covariate of interest. Given that the outcome is the change in logged income, these coefficients can be interpreted as percentage change in income. Thus, a neighborhood with one standard deviation higher income mixing is expected to have 2.7% lower average income appreciation over the subsequent decade than an otherwise similar neighborhood (-.027). And we see that whereas a neighborhood with one standard deviation higher land use entropy experiences 1.6% lower average income appreciation over the subsequent decade, one with high housing age entropy experiences 0.8% higher average income appreciation.

In this same table we present the results for the more conventional OLS analysis for comparison purposes. One thing to note is that whereas the OLS model explains 23% of the variance, the KRLS model explains 37% of the variance. Nonetheless, it is worth acknowledging that there is additional variance to explain even in the KRLS model, as 63% of the variance remains unexplained. This improvement highlights the advantage of this alternative approach, which captures a larger extent of variation by considering nonlinearities and interaction effects that are not apparent in the traditional OLS approach. There are some differences in parameter estimates across the OLS and KRLS models. For example, in the OLS model it appears that higher percent black residents at the beginning of the decade are negatively associated with the change in average household income over the subsequent decade, but the parameter is close to zero in the KRLS model. And whereas the percent black in the surrounding area is not related to income change in the OLS model, it shows an average positive relationship in the KRLS model. Given these differences, it is useful to explore whether these coefficient estimates systematically vary based on values of other variables in the model, and we do this next. While it is possible to examine nonlinearities by parameterizing an OLS model using interaction terms (e.g. the joint effect of racial mixing and household vacancy on income growth), this would require dozens of additional covariates whose joint effects must be individually interpreted relative to their marginal effects. This is a cumbersome process and it is often challenging to isolate key interactions; furthermore it would only approximate the more flexible and nonparametric KRLS results (Hainmueller and Hazlett 2014).

How mixing is moderated by other types of mixing

To assess whether these coefficient effects depend on other variables in the model, we next plot the predicted values from Lowess regressions of the derivatives on a specific covariate (and its quadratic term) that showed R-squares of at least .10 (all of these relationships were also statistically significant). Each instance with such a notable moderating effect is summarized in Table 4 for each of our mixing measures. In this table, "high" refers to the upper part of the distribution of a variable; "moderate" refers to the middle range (typically the 40th to 60th percentile), and "low" refers to the bottom part of the range of a variable.

We find that income mixing has a stronger positive relationship with the change in household incomes when there are low levels of racial and housing age mixing. Figure 1a shows how income mixing is conditioned by the level of racial mixing in the neighborhood. In Figure 1a, the x-axis represents various values of the moderating variable (in this case, racial mixing) whereas the y-axis is the estimated derivative for the moderated variable (in this case, income mixing) on the outcome variable of change in logged income (this can be thought of as the coefficient value at a particular value of the x-axis variable). For example, an increase in income mixing in neighborhoods with high racial mixing (the right side of the graph) is expected to result in a decrease in average income in the subsequent decade (given that the y-axis values are below zero). If, instead, the relationship between income mixing and the change in average income did not differ based on the racial mixing of the neighborhood, this plot would be approximately the flat dotted line depicting the median marginal effect. Instead, increasing income mixing one standard deviation results in about 5% *lower* average income appreciation in neighborhoods with very high levels of racial mixing, but increasing income mixing is associated with about 1% *greater* average income appreciation in neighborhoods with very low racial mixing—seen in the positive y-axis values on the left side of the graph (all interpretations are based on a one standard deviation change). Given that the average effect of a one standard deviation increase in income mixing in this model was a 2.8% decrease in average income appreciation, we can see that a substantial amount of this effect is determined by the level of racial mixing. In other words, while in general mixed income areas show lower levels of household income growth, income mixing does *not* have a detrimental impact on household income growth in racially homogenous neighborhoods. Likewise, the negative relationship between income mixing and average income appreciation is weaker in neighborhoods with low housing age mixing, as seen in Figure 1b. A one standard deviation increase in income mixing reduces average income appreciation about 4% in neighborhoods with very high housing age mixing, whereas the negative impact is about 2% in neighborhoods with low housing age mixing.

<<<Table 4 about here>>>

<<<Figure 1 about here>>>

We find that racial mixing has a stronger positive relationship with household income appreciation in neighborhoods with high levels of housing age mixing. Racial mixing has a positive relationship with average income appreciation in neighborhoods with very high levels of housing age mixing, but effectively no relationship in neighborhoods at or below the mean in housing age mixing in a pattern somewhat similar to Figure 1c. Racial mixing also exhibits a nonlinear relationship itself, in that changes in racial mixing have a negative effect at low values of racial mixing, but changes have a positive effect at high values of racial mixing, also similar to Figure 1c.

How income mixing is moderated by neighborhood conditions

We find that income mixing has a stronger positive relationship with average income appreciation in high socio-economic status neighborhoods. As seen in Figure 1c, whereas income mixing has a strong negative effect on average income appreciation in relatively poor neighborhoods—income mixing reduces average income appreciation about 5% in neighborhoods with low average income—income mixing actually is associated with increasing average income in more advantaged neighborhoods—higher income mixing results in about 1% greater average income gains in very high income neighborhoods. The same pattern was found based on the average income in the surrounding area, as well as the unemployment rate of the neighborhood or surrounding area. In other words, income mixing is not detrimental to income growth rates so long as the area is fairly wealthy on average.

Income mixing is associated with lower average income appreciation neighborhoods with high population density or residential instability. Thus, income mixing has its strongest negative effect on average income appreciation in neighborhoods with relatively high population density, surrounded by high density, or in which the vacancy rate is decreasing (implying higher density). The result is similar in neighborhoods with high residential instability or surrounded by high instability (measured as low average length of residence or a high proportion of renters), but shows a modest positive effect in very low population density neighborhoods. For the vacancy rate, it is only at the highest levels (which typically are a sign of dysfunction in a neighborhood) that this effect reverses.

Income mixing also has a stronger negative relationship with average income appreciation in neighborhoods with more Latinos or immigrants, or surrounded by areas with more members of these groups. Income mixing has effectively no relationship with average income appreciation in neighborhoods with no Latinos as seen in Figure 1e, but an increasingly stronger negative relationship as the percent Latino in the neighborhood increases. Likewise, increasing income mixing results in about 5% lower average income over the subsequent decade in neighborhoods with 60% immigrants.

The age structure of the neighborhood also matters, as income mixing has a stronger negative relationship with average income appreciation in neighborhoods with fewer retirees or more persons under 20. Whereas income mixing has a modest negative effect on average income appreciation in neighborhoods with a higher percentage over 65 (the right side of Figure 1f), this is a strong negative relationship in neighborhoods with a low proportion of retirement-age individuals (the left side of the graph).

How racial mixing is moderated by neighborhood conditions

It appears that racial mixing has more positive consequences when it occurs in neighborhoods that are more disadvantaged economically. In neighborhoods with very low average income, higher levels of racial mixing actually are associated with larger increases in average income over the subsequent decade (Figure 1h). In contrast, racial mixing in high income tracts is associated with negative average income appreciation. The pattern is similar when measuring economic disadvantage based on the unemployment rate, or when the neighborhood is surrounded by low income areas.

Racial mixing appears to have a more positive impact on average household income appreciation in neighborhoods with higher levels of racial minorities or surrounded by such groups (measured as percent Latino, or percent immigrants). For example, racial mixing has a positive relationship with average income appreciation when it occurs in neighborhoods with a high percentage Latino, as shown in Figure 1g. Likewise, racial mixing has a stronger positive effect when it occurs in neighborhoods with high immigrant concentrations (average income increases about 1 to 1.5% more in such neighborhoods).

Racial mixing has a stronger positive relationship with average income appreciation in neighborhoods with high population density or more renters. This was also the case in neighborhoods with very low percent open land (and therefore higher density), or surrounded by high density. Racial mixing has positive consequences in neighborhoods dominated by renters, but less so in neighborhoods with more owners (similar to Figure 1e). The effect of renters in the surrounding area was similar, except that racial mixing actually has negative consequences when the neighborhood is surrounded by high homeownership areas.

How housing age mixing is moderated by neighborhood conditions

Housing age mixing has a stronger positive relationship with average income appreciation in neighborhoods surrounded by a mix of owners or renters, or low population density. Housing age mixing has its strongest positive impact on household income appreciation in neighborhoods surrounded by 40-70% homeowners, but weaker effects in neighborhoods surrounded by either a very low proportion or very high proportion of homeowners. Housing age mixing has a positive relationship with average income appreciation in neighborhoods surrounded by low population density, similar to Figure 1e. Housing age mixing also exhibits a nonlinear relationship itself, as it has a negative relationship with average income appreciation in neighborhoods with low housing age mixing, but a positive relationship in neighborhoods with high housing age mixing.

How land use mixing is moderated by neighborhood conditions

Land use mixing has its strongest negative relationship with average income appreciation in neighborhoods with a moderate percentage black, or surrounded by low to average residential stability. In neighborhoods with about 5-15% black in the neighborhood itself or the surrounding area the relationship is at its strongest negative, but it is less negative when there is a very small or very large percentage black. And, similar to Figure 1d, neighborhoods with increasing land use mixing that are surrounded by low residential stability experience a stronger negative relationship.

Ancillary models

In KRLS models, as in all models, there is a concern of omitted variables that can bias the results. We have adopted an approach in which we use measures at the beginning of the decade to explain changes in average income over the subsequent decade. The advantage of this approach is that it minimizes the potential of endogeneity that can occur by including measures of change in the neighborhood at the same time as the change in our outcome measure. Nonetheless, there may be concern that neighborhoods that are experiencing increasing average income are also experiencing an increase in population and housing units given that they may be desirable locations. We assessed this by estimating ancillary models that included the change in population density during the decade as a covariate. It is encouraging to note that although this population density measure demonstrated a significant relationship (although it was in fact a negative one) the results of our other variables in the model were very similar to those in the presented models when including this change variable (results available upon request).

Discussion and Conclusions

This study has explored the relationship between the level of mixing in neighborhoods based on four dimensions and the consequences for average income appreciation over the subsequent decade for neighborhoods in the southern California region. We have highlighted that the existing literature often points to the importance of considering how mixing based on certain dimensions may have different consequences for neighborhoods based on other neighborhood characteristics. There are also non-specific theoretical predictions regarding the functional form of the relationship between these mixing dimensions and economic dynamism in neighborhoods. For these reasons, we utilized a relatively new statistical strategy—kernel regularized least squares—a machine learning approach that is flexible enough to estimate various functional forms of these relationships, as well as estimate the possible interdependencies of different neighborhood structural measures when assessing their relationship with the change in neighborhood average household income over the subsequent decade. The results highlight important interdependencies between certain dimensions of mixing and key neighborhood structural characteristics. These interdependencies were particularly important for assessing the relationship between income mixing and neighborhood dynamism and refining existing urban policy instruments.

We can think of these neighborhood characteristics that moderate the relationship between dimensions of mixing and economic dynamism as “ingredients” that are important for fostering dynamism. Whereas income mixing on average showed a negative relationship with average income appreciation, income mixing in the context of certain neighborhood ingredients did not reduce average household income over time as much. Thus, in our study income mixing is associated with greater income increases for a neighborhood with 1) low mixing on other dimensions (racial and housing age); 2) higher SES (average income or unemployment rate); 3) high population density (and few vacancies); 4) high residential stability (owners and average length of residence); 5) fewer racial minorities (Latinos or immigrants); 6) an older age structure (more retirees, fewer children). Thus, income mixing when combined with other types of mixing —specifically, racial mixing and housing age mixing—is associated with lower average income appreciation over the subsequent decade. This may suggest that the combination of racial mixing and income mixing often indicates polarization—or what Peter Blau referred to as consolidated inequality (Blau 1987)—and leads to negative outcomes rather than economic benefits to the residents. This is the general idea of social distance based on various social dimensions, and one study found that micro-neighborhoods with higher levels of social distance reported higher levels of disorder and crime (Hipp 2010b). Whereas housing age mixing might promote mixed-income neighborhoods in a process similar to Jacobs' (Jacobs 1961) suggestion that building age mixing promotes a wider variety of local retail establishments, in our study of Southern California housing age mixing actually has negative consequences for neighborhoods when combined with income mixing.

Income mixing demonstrated better consequences when it occurs in more economically advantaged neighborhoods than in disadvantaged neighborhoods. This may imply that more disadvantaged neighborhoods are more fragile and vulnerable. One possibility is that a mix of income groups at the low end of the income scale may occur during the process of neighborhood decline or induce a lowered sense of cohesion and sense of attachment to the neighborhood. This may make the neighborhood appear less desirable to other potential in-migrants. While this is speculative, our results highlight the need for future research to explore more closely what it is about income mixing for more disadvantaged neighborhoods that may lead to negative outcomes.

It is interesting to note that income mixing and racial/ethnic mixing had different consequences for average income appreciation when they occurred within the context of neighborhoods containing other New Urbanist principles. Thus, whereas increasing income mixing in a context of high housing age mixing had negative consequences for average income appreciation, increasing racial mixing in a context of high housing age mixing actually had *positive* consequences for average income appreciation. Racial mixing in the context of mixed-age housing may capture the quintessential multicultural transition area that is desirable to young adults. Similarly, racial/ethnic mixing had a stronger positive impact on income appreciation in the context of high population density in the tract and nearby, whereas income mixing in such a context had negative consequences. The higher density may reflect more opportunities for different racial/ethnic groups to interact following the insights of contact theory (Allport 1958 [1954]), resulting in more cohesion in the neighborhood. This could then possibly lead to a more economically vibrant neighborhood, although further research would be necessary to assess if this indeed occurs in such neighborhoods. As to why income mixing does not yield such positive benefits in the context of high population density is not entirely clear. One possibility is that the typical preference for low density housing among higher income residents results in income mixing being less effective in high density locations.

We found that racial mixing can have a positive impact on average income over time when it is accompanied by the following ingredients: 1) high housing age mixing; 2) low socio-economic status (average income, unemployment rate); 3) more racial minorities (Latinos or immigrants); 4) more population density (and low percentage of open land); 5) more renters. It appears that racial mixing may capture more multicultural neighborhoods with more interesting amenities. Thus, the positive relationship of racial mixing was accentuated by the presence of more immigrants, which may directly translate into diverse and multicultural food options for residents. It may also be that neighborhoods with more immigrants provide a signal that an area is more amenable to diversity (Florida 2002). Likewise, the fact the positive relationship of racial mixing was accentuated by the presence of many renters may also be consistent with the notion that these are transitional neighborhoods dominated by younger, less established persons interested in diverse neighborhoods.

There was more modest evidence that housing age and land use mixing impacted neighborhood dynamism. Housing age mixing, which typically occurs in older, more established areas which have experienced some new housing construction through infill, exhibited a positive relationship with the change in average income when it is accompanied by two ingredients: 1) low population density in the surrounding area; 2) a relatively mixed percentage of owners and renters at a broader scale (in the surrounding area). Housing age mixing and owner/renter mixing in conjunction result in a more economically dynamic neighborhood. Thus, housing age mixing operates in tandem in a negative fashion with income mixing, and in a positive fashion with racial mixing and owner/renter mixing, to impact economic dynamism. This highlights that simultaneously accounting for different dimensions of mixing is important for understanding how neighborhoods evolve over time. It is interesting to note that in our study housing age mixing impacted neighborhood dynamism more than did land use mixing, despite the latter's more prominent feature in much research. In fact, land use mixing had an overall negative relationship with economic dynamism, and only had a positive relationship when accompanied by a relatively small proportion of residential units; this implies that land use mixing needs to be quite pronounced—and not simply a small mix of other land use with residential units—to be effective.

We acknowledge some limitations to this study. First, we have focused on a single decade of average income growth in neighborhoods, and therefore cannot address longer-term effects. Second, although tracts are not necessarily an ideal measure of "neighborhood", our reliance on census-generated data required us to use this particular unit of analysis. Third, we have focused on mixing within tracts and have therefore not viewed mixing at larger spatial scales. This was done to maintain proper scope of the study, but nonetheless suggests a need for future research that accounts for mixing at larger scales. Fourth, we have focused on neighborhoods in a single region. Despite Southern California's large size, there is a need for similar studies in other regions to assess the generalizability of these results. Fifth, there is always a concern with omitted variables that can bias results. Although this is a concern with all studies, it is worth emphasizing that despite the flexibility of the KRLS approach, it does not solve this potential problem. Finally, the focus on average income growth rather than median growth – necessitated due to the use of interpolated census geographies – does not reflect as accurately the experience of a neighborhood's typical resident and can be inflated by a small number of very wealthy entrants. Given concerns over the potential displacing effects of gentrification (Newman and Wyly 2006), average income growth may not be an ideal indicator of neighborhood well-being at all – a future study that takes moving into account may be better suited to address this issue though such an analysis is outside the scope of this paper.

In conclusion, this study has highlighted that whereas various forms of mixing can have important implications for economic dynamism in neighborhoods, this mixing is not independent of other neighborhood characteristics. By utilizing a statistical analysis technique that explicitly accounts for nonlinearities in these relationships, and explicitly accounts for possible nonlinear interactions with other measures, we have demonstrated that the neighborhood context as a whole should be considered in understanding which neighborhoods will exhibit greater average income appreciation over the subsequent decade. Our results suggest that any theory presuming a linear marginal relationship between a particular neighborhood structural measure and economic growth is not entirely reasonable. Instead, there appear to be nonlinearities and possible threshold points for some of these relationships that deserve more attention and that considering the broader "ingredients" of the neighborhood are important for better understanding outcomes of mixing.

References

Allport, Gordon W. 1958 [1954]. *The Nature of Prejudice*. Garden City, NY: Doubleday.

Baxter, Vern and Mickey Lauria. 2000. "Residential mortgage foreclosure and neighborhood change." *Housing Policy Debate* 11:675-699.

Blau, Peter M. 1987. *Inequality and Heterogeneity : A Primitive Theory of Social Structure*. New York: Free Press.

Bridge, Gary, Tim Butler, and Loretta Lees. 2012. *Mixed Communities: Gentrification by Stealth?* Bristol: Policy Press.

Calthorpe, Peter. 1993. *The Next American Metropolis: Ecology, Community, and the American Dream*. New York: Princeton Architectural Press.

Calthorpe, Peter and William Fulton. 2001. *The Regional City: Planning for the End of Sprawl*. Washington, DC: Island Press.

Campoli, Julie. 2012. *Made for Walking: Density and Neighborhood Form*. Cambridge, MA: Lincoln Institute of Land Policy.

Campoli, Julie and Alex S. MacLean. 2007. *Visualizing Density*. Cambridge, MA: Lincoln Institute of Land Policy.

Chapple, Karen and Rick Jacobus. 2009. "Retail trade as a route to neighborhood revitalization." in *Urban and Regional Policy and Its Effects*, vol. 1, edited by N. Pindus, H. Wial, and H. Wolman: Brookings Institution Press.

Cleveland, William S. 1979. "Robust locally weighted regression and smoothing scatterplots." *Journal of the American Statistical Association* 74:829-836.

Ehrenhalt, Alan. 2012. *The Great Inversion and the Future of the American City*. New York: Random House LLC.

Ellen, Ingrid Gould and Katherine O'Regan. 2008. "Reversal of Fortunes? Lower-income Urban Neighbourhoods in the US in the 1990s." *Urban Studies* 45:845-869.

Ferwerda, Jeremy, Jens Hainmueller, and Chad J. Hazlett. 2013. "krls: A Stata Package for Kernel-Based Regularized Least Squares." *Journal of Statistical Software*.

Florida, Richard. 2002. *The Rise of the Creative Class*. New York: Basic Books.

Fraser, James and Michael H. Nelson. 2008. "Can Mixed-Income Housing Ameliorate Concentrated Poverty? The Significance of a Geographically Informed Sense of Community." *Geography Compass* 2:2127-2144.

Galster, George C., Jackie Cutsinger, and Up Lim. 2007. "Are Neighbourhoods Self-stabilising? Exploring Endogenous Dynamics." *Urban Studies* 44:167-185.

Galster, George C., Chris Hayes, and Jennifer Johnson. 2005. "Identifying robust, parsimonious neighborhood indicators." *Journal of Planning Education and Research* 24:265-280.

Golding, Shaun A. 2015. "Gentrification and Segregated Wealth in Rural America: Home Value Sorting in Destination Counties." *Population Research and Policy Review* 35:127-146.

Hainmueller, Jens and Chad J. Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis* 22:143-168.

Hall, Eliza. 2007. "Divide and sprawl, decline and fall: A comparative critique of Euclidean zoning." *University of Pittsburgh Law Review* 68:915-962.

Hanlon, Bernadette and Thomas J. Vicino. 2007. "The Fate of Inner Suburbs: Evidence From Metropolitan Baltimore." *Urban Geography* 28:249-275.

Hipp, John R. 2007. "Income Inequality, Race, and Place: Does the Distribution of Race and Class within Neighborhoods affect Crime Rates?" *Criminology* 45:665-697.

—. 2009. "Specifying the Determinants of Neighborhood Satisfaction: A Robust Assessment in 24 Metropolitan Areas over Four Time Points." *Social Forces* 88:395-424.

—. 2010a. "A dynamic view of neighborhoods: The reciprocal relationship between crime and neighborhood structural characteristics." *Social Problems* 57:205-230.

—. 2010b. "Micro-structure in Micro-Neighborhoods: A New Social Distance Measure, and its Effect on Individual and Aggregated Perceptions of Crime and Disorder." *Social Networks* 32:148-159.

Hipp, John R. and Adam Boessen. 2013. "Egohoods as waves washing across the city: A new measure of “neighborhoods”." *Criminology* 51:287-327.

Hipp, John R. and Andrew J. Perrin. 2009. "The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties." *City & Community* 8:5-25.

Immergluck, Daniel. 1998. "Job Proximity and the Urban Employment Problem: Do Suitable Nearby Jobs Improve Neighbourhood Employment Rates?" *Urban Studies* 35:7-23.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Random House.

Jun, Hee-Jung. 2016. "The Effect of Racial and Ethnic Composition on Neighborhood Economic Change: A Multilevel and Longitudinal Look." *Housing Studies* 31:102-125.

Kitchen, Peter and Allison Williams. 2009. "Measuring Neighborhood Social Change in Saskatoon, Canada: A Geographic Analysis." *Urban Geography* 30:261-288.

Kleit, Rachel Garshick. 2005. "HOPE VI new communities: neighborhood relationships in mixed-income housing." *Environment and Planning A* 37:1413-1441.

Knaap, G. 2005. "New Urbanism and Smart Growth: A Few Words from the Academy." *International Regional Science Review* 28:107-118.

Leccese, Michael and Kathleen McCormick. 1999. "Charter of the New Urbanism." Pp. 194. New York: McGraw Hill.

Levy, Diane K., Zach McDade, and Kassie Dumlao. 2010. "Effects from Living in Mixed-Income Communities for Low-Income Families." Urban Institute.

Lowenkamp, Christopher T., Francis T. Cullen, and Travis C. Pratt. 2003. "Replicating Sampson and Groves's Test of Social Disorganization Theory: Revisiting a Criminological Classic." *Journal of Research in Crime and Delinquency* 40:351-373.

Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67:281-315.

Messner, Steven F. and Kenneth Tardiff. 1986. "Economic Inequality and Levels of Homicide: An Analysis of Urban Neighborhoods." *Criminology* 24:297-317.

Molotch, Harvey. 1976. "The City as a Growth Machine: Toward a Political Economy of Place." *American Journal of Sociology* 82:309-332.

Newman, Kathe and Elvin Wyly. 2006. "The right to stay put, revisited: Gentrification and resistance to displacement in New York city." *Urban Studies* 43:23-57.

Park, Robert E., Ernest W. Burgess, and Roderick D. McKenzie. 1925. *The City*. Chicago: University of Chicago.

Pitken, Bill. 2001. "Theories of Neighborhood Change: Implications for Community Development Policy and Practice." Pp. 28 in *UCLA Advanced Policy Institute*. Los Angeles, CA.

Roberts, Marion. 2007. "Sharing Space: Urban Design and Social Mixing in Mixed Income New Communities." *Planning Theory & Practice* 8:183-204.

Roncek, Dennis W. and Pamela A. Maier. 1991. "Bars, Blocks, and Crimes Revisited: Linking the Theory of Routine Activities to the Empiricism of 'Hot Spots'." *Criminology* 29:725-753.

Rosenthal, Stuart S. 2008. "Old homes, externalities, and poor neighborhoods. A model of urban decline and renewal." *Journal of Urban Economics* 63:816-840.

Rountree, Pamela Wilcox and Barbara D Warner. 1999. "Social ties and crime: Is the relationship gendered?" *Criminology* 37:789-813.

Rubin, Geoffrey M. 1993. "Is Housing Age a Commodity? Hedonic Price Estimation of Unit Age." *Journal of Housing Research* 4:165-184.

Sampson, Robert J. 1991. "Linking the Micro- and Macrolevel Dimensions of Community Social Organization." *Social Forces* 70:43-64.

Sampson, Robert J. and W. Byron Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 94:774-802.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143-186.

Song, Yan and Gerrit-Jan Knaap. 2004. "Measuring the effects of mixed land uses on housing values." *Regional Science and Urban Economics* 34:663-680.

Talen, Emily. 1999. "Sense of Community and Neighbourhood Form: An Assessment of the Social Doctrine of New Urbanism." *Urban Studies* 36:1361-1379.

—. 2013. "Charter of New Urbanism: Congress for the New Urbanism." Pp. 302. New York: McGraw-Hill.

Temkin, Kenneth and William M. Rohe. 1996. "Neighborhood Change and Urban Policy." *Journal of Planning Education and Research* 15:159-170.

Warner, Barbara D. and Pamela Wilcox Rountree. 1997. "Local Social Ties in a Community and Crime Model: Questioning the Systemic Nature of Informal Social Control." *Social Problems* 44:520-536.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Tables and Figures

Table 1

Race	Housing age	Land Use Category
White only (non-hisp)	Pre 1939	Single-Family Residential
Black only (non-hisp)	1940s and 1950s	Multifamily Residential
Asian only (non-hisp)	1960s and 1970s	Commercial
Hispanic - 1 race	1980s and 1990s	Industrial
Other/Mixed Race	2000s and 2010s	Open Space

Table 2. Summary statistics of variables used in analyses

	Tract		Surrounding	
	Mean	SD	Mean	SD
Change in average household income	0.26	0.24		
Mixing variables				
Household income Gini Coefficient	40.04	5.93		
Race entropy	0.67	0.19		
Housing age entropy	0.70	0.23		
Land use entropy	0.47	0.26		
Tract-level variables				
Average household income (logged)	10.94	0.46	4.16	0.30
Percent open land	3.64	9.93		
Percent residential	37.92	24.47		
Percent over 65	10.58	7.53		
Percent under 20	30.39	8.82		
Percent black	6.91	11.90	7.13	7.94
Percent Latino	37.52	27.24	39.88	19.20
Percent immigrants	28.97	16.39		
Population density	8.18	8.51	6.21	4.36
Unemployment rate	7.72	5.06	7.70	2.67
Percent owners	55.69	25.93	53.46	15.96
Percent occupied units	94.61	7.23	95.03	4.75
Average length of residence	9.42	3.14	9.25	1.39
N = 4,441 tracts in Southern California				

Table 3. Results of Kernel Regularized Least Squares model and OLS model predicting change in average household income from 2000 to 2010

	KRLS model								OLS model		
	Avg	t-value	Sig	P25	P50	P75	Std		Coef	t-value	Sig
Mixing variables											
Household income Gini Coefficient	-0.005	-(7.77)	**	-0.007	-0.005	-0.003	-0.027		-0.004	-(6.89)	**
Race entropy	0.022	(0.99)		-0.018	0.022	0.062	0.004		0.005	(0.23)	
Housing age entropy	0.034	(2.15)	*	0.006	0.033	0.061	0.008		0.049	(2.80)	**
Land use entropy	-0.062	-(4.39)	**	-0.102	-0.056	-0.017	-0.016		-0.088	-(5.99)	**
Tract-level variables											
Average household income (logged)	-0.279	-(25.26)	**	-0.353	-0.285	-0.218	-0.128		-0.403	-(25.92)	**
Percent open land	0.080	(5.72)	**	0.035	0.069	0.110	0.008		0.191	(11.50)	**
Percent residential	0.069	(5.54)	**	0.034	0.069	0.105	0.017		0.093	(6.16)	**
Percent over 65	-0.116	-(1.92)	†	-0.176	-0.117	-0.054	-0.009		-0.319	-(5.57)	**
Percent under 20	-0.122	-(2.35)	*	-0.238	-0.114	0.000	-0.011		-0.226	-(3.33)	**
Percent black	0.000	(0.15)		0.000	0.000	0.000	0.001		-0.002	-(4.41)	**
Percent Latino	-0.001	-(5.83)	**	-0.001	-0.001	-0.001	-0.025		-0.002	-(7.22)	**
Percent immigrants	-0.001	-(5.84)	**	-0.002	-0.001	-0.001	-0.021		-0.002	-(5.41)	**
Population density	-0.002	-(3.40)	**	-0.002	-0.002	-0.001	-0.014		-0.003	-(5.84)	**
Unemployment rate	-0.001	-(1.06)		-0.002	-0.001	0.000	-0.005		-0.003	-(3.11)	**
Percent owners	0.000	(2.20)	*	0.000	0.000	0.001	0.010		0.002	(5.10)	**
Percent occupied units	0.000	(0.55)		-0.001	0.000	0.002	0.003		0.000	-(0.73)	
Average length of residence	0.001	(0.57)		-0.001	0.001	0.003	0.002		-0.002	-(1.42)	
Surrounding 5 miles (inverse distance decay)											
Average household income (logged)	0.137	(9.22)	**	0.095	0.146	0.181	0.041		0.328	(12.54)	**
Percent black	0.001	(2.26)	*	0.000	0.001	0.003	0.009		0.001	(1.31)	
Percent Latino	0.000	(1.83)	†	0.000	0.000	0.001	0.008		0.002	(4.73)	**
Population density	0.005	(4.70)	**	0.003	0.005	0.007	0.021		0.008	(5.51)	**
Unemployment rate	-0.009	-(5.11)	**	-0.012	-0.009	-0.006	-0.023		-0.008	-(2.94)	**
Percent owners	-0.002	-(7.58)	**	-0.003	-0.002	-0.001	-0.032		-0.004	-(9.72)	**
Percent occupied units	-0.002	-(1.85)	†	-0.003	-0.002	0.000	-0.008		0.002	(2.55)	*
Average length of residence	0.011	(4.76)	**	0.004	0.012	0.019	0.015		0.012	(4.38)	**
R-square	0.368								0.232		

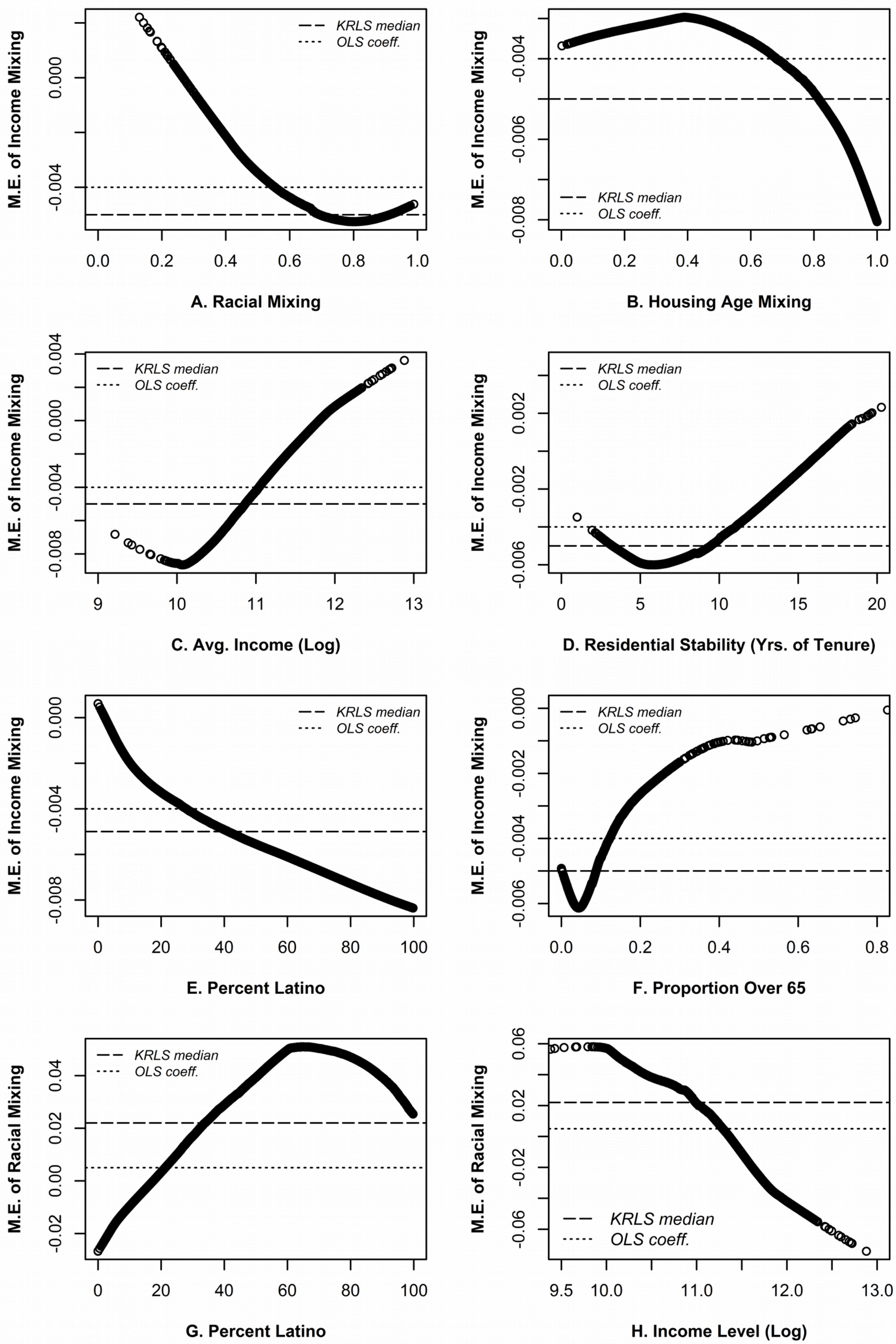
*Note: ** p < .01(two-tail test), * p < .05 (two-tail test), † p < .05 (one-tail test). Avg= average coefficient estimate over the covariate space; t-value= the t-value of the average coefficient estimate; P25= the coefficient estimate at the 25th percentile; P50= the coefficient estimate at the median; P75= the coefficient estimate at the 75th percentile; Std = change in Y for a one standard deviation change in the covariate.*

Table 4. Ingredients that have important implications for the relationship between mixing and average household income appreciations

	Conditions where income mixing increases income growth		Conditions where racial mixing increases income growth		Conditions where housing age mixing increases income growth		Conditions where land use mixing increases income growth	
	Tract	Nearby tracts	Tract	Nearby tracts	Tract	Nearby tracts	Tract	Nearby tracts
Mixing variables								
Household income Gini Coefficient								
Race entropy	low							
Housing age entropy	low		high					
Land use entropy								
Tract-level variables								
Average household income (logged)	high		low	low				
Percent open land			low					
Percent residential land								
Percent over 65	high							
Percent under 20	low							
Percent black							high/low	high/low
Percent Latino	low	low	high	high				
Percent immigrants	low		high					
Population density	low	low	high	high		low		
Unemployment rate	low	low	high	high				
Percent owners	high	high	low	low		modest		
Percent occupied units	high/low	high/low						
Average length of residence	high							high

Note: Noted cases represent interaction effects in which the r-square of regressing the derivates on the variable of interest (and its quadratic) was at least .10. "High" refers to the upper part of the distribution of a variable; "Moderate" refers to the middle range (typically the 40th to 60th percentile), and "low" refers to the bottom part of the range of a variable.

Figure 1: Marginal Effect (M.E.) of Mixing on Income Growth Based on:



Appendix

Table A1. R-square results from regressing derivatives of specific variable on other covariates (both quadratic and cubic form specifications)

	Quadratic	Cubic
Inequality by average income (logged)	0.490	0.510
Inequality by percent Latino	0.393	0.405
Inequality by unemployment rate	0.366	0.399
Racial mixing by percent immigrants	0.389	0.391
Inequality by percent owners	0.365	0.365
Inequality by nearby average income	0.343	0.344
Racial mixing by nearby population density	0.326	0.338
Inequality by nearby unemployment rate	0.293	0.303
Racial mixing by nearby percent owners	0.288	0.288
Inequality by percent immigrants	0.279	0.280
Racial mixing by percent open space	0.269	0.273
Inequality by average length of residence	0.250	0.270
Inequality by nearby percent Latino	0.235	0.237
Racial mixing by nearby percent Latino	0.221	0.230
Inequality by nearby population density	0.215	0.229
Inequality by percent under 20 years of age	0.209	0.221
Inequality by population density	0.208	0.220
Racial mixing by population density	0.200	0.215
Racial mixing by percent Latino	0.194	0.209
Racial mixing by racial mixing	0.190	0.201
Inequality by home type mixing	0.196	0.198
Inequality by percent residential	0.190	0.198
Racial mixing by nearby unemployment rate	0.190	0.190
Inequality by nearby percent owners	0.185	0.188
Inequality by percent 65 and older	0.183	0.187
Racial mixing by percent owners	0.187	0.187
Housing type mixing by home type mixing	0.144	0.155
Racial mixing by home type mixing	0.152	0.152
Land use mixing by nearby average length of residence	0.132	0.135
Housing type mixing by nearby percent owners	0.132	0.134
Land use mixing by nearby percent black	**0.091**	**0.134**
Housing type mixing by nearby population density	0.133	0.133
Racial mixing by average income (logged)	0.129	0.132
Inequality by racial mixing	0.131	0.131
Inequality by nearby percent occupied units	**0.043**	**0.129**
Racial mixing by unemployment rate	0.102	0.126
Racial mixing by nearby average income	0.116	0.124
Land use mixing by percent black	**0.054**	**0.113**
Inequality by percent occupied units	**0.048**	**0.109**
Housing type mixing by percent occupied units	0.070	0.095
Land use mixing by nearby percent Latino	0.087	0.094
Racial mixing by inequality	0.090	0.090
Nonlinear inequality	0.087	0.090
Housing type mixing by nearby average length of residence	0.083	0.086
Inequality by nearby average length of residence	0.081	0.083
Land use mixing by nearby percent occupied units	0.017	0.078
Land use mixing by percent Latino	[illegible]	[illegible]

Nonlinear inequality	0.039	0.077
Inequality by percent black	0.063	0.074
Inequality by land use mixing	0.072	0.073
Land use mixing by racial mixing	0.071	0.072
Land use mixing by percent open space	0.052	0.068
Land use mixing by population density	0.056	0.067
Land use mixing by average length of residence	0.064	0.067
Housing type mixing by percent black	0.044	0.062
Housing type mixing by nearby unemployment rate	0.059	0.059
Housing type mixing by nearby average income	0.015	0.059
Housing type mixing by percent immigrants	0.057	0.058
Racial mixing by percent residential	0.025	0.057
Housing type mixing by percent open space	0.056	0.056
Housing type mixing by percent owners	0.049	0.056
Housing type mixing by nearby percent occupied units	0.041	0.054
Racial mixing by nearby percent black	0.043	0.053
Housing type mixing by population density	0.046	0.052
Racial mixing by land use mixing	0.052	0.052
Land use mixing by percent residential	0.050	0.050
Racial mixing by percent 65 and older	0.050	0.050
Land use mixing by land use mixing	0.046	0.046
Housing type mixing by average length of residence	0.033	0.045
Land use mixing by percent occupied units	0.015	0.044
Housing type mixing by percent Latino	0.036	0.044
Housing type mixing by nearby percent black	0.036	0.043
Inequality by percent open space	0.038	0.043
Housing type mixing by percent residential	0.036	0.042
Housing type mixing by nearby percent Latino	0.040	0.041
Land use mixing by nearby unemployment rate	0.036	0.036
Racial mixing by percent black	0.031	0.036
Housing type mixing by average income (logged)	0.004	0.035
Land use mixing by nearby population density	0.034	0.034
Housing type mixing by land use mixing	0.026	0.028
Housing type mixing by racial mixing	0.025	0.025
Racial mixing by percent under 20 years of age	0.016	0.023
Racial mixing by percent occupied units	0.022	0.022
Land use mixing by percent owners	0.019	0.021
Land use mixing by unemployment rate	0.020	0.021
Racial mixing by nearby percent occupied units	0.020	0.020
Racial mixing by average length of residence	0.017	0.018
Housing type mixing by unemployment rate	0.009	0.017
Land use mixing by home type mixing	0.017	0.017
Nonlinear inequality	0.014	0.015
Land use mixing by nearby percent owners	0.014	0.014
Land use mixing by average income (logged)	0.011	0.013
Land use mixing by nearby average income	0.007	0.013
Land use mixing by percent immigrants	0.003	0.012
Housing type mixing by percent under 20 years of age	0.011	0.011
Land use mixing by percent 65 and older	0.010	0.011
Racial mixing by nearby average length of residence	0.009	0.010