

**Climate Change, Vegetation, and Welfare: Estimating the Welfare Loss to
Landowners of Marginal Shifts in Blue Oak Habitat**

Peter H. Howard
Agricultural and Resource Economics
University of California, Davis
howard@primal.ucdavis.edu

*Selected Paper prepared for presentation at the Agricultural & Applied Economics
Association's 2012 AAEA Annual Meeting, Seattle, Washington, August 12-14, 2012*

*Copyright 2012 by Peter H. Howard. All rights reserved. Readers may make verbatim copies of
this document for non-commercial purposes by any means, provided that this copyright
notice appears on all such copies.*

Abstract

Scientists predict that climate change will cause suitable habitat ranges to shift for many plant species. To the extent that proximity to particular vegetation types increases residents' utility and/or these shifts affect services valued by all of society, such geographic shifts in ecosystems may significantly affect societal welfare. In this paper, I estimate the possible welfare change from the marginal loss of blue oak due to development and climate change in the Tulare Lake Basin (Fresno, Kern, and Tulare Counties) in California. Using a hedonic pricing model, the marginal values of blue oaks and the land cover types most likely to replace them (herbaceous, urban, and crop land) are estimated at multiple spatial scales, using 1997-2003 sales of single family residences for the Tulare Lake Basin. In addition to the common identification problems of specification error, omitted variable bias, and multicollinearity, the variables measuring the degree of proximity of a property to land cover types are endogenous. To identify the marginal values of land cover types at multiple spatial scales using two-stage least squares, instrumental variables are developed using soil data. Cluster robust standard errors are calculated due to spatial autocorrelation within neighborhoods. Results indicate that households do not differentiate between vegetation land cover types; there is no indirect cost of climate change resulting from marginal shifts in land cover types. The results also indicate that Tulare Lake Basin households are unlikely to be negatively affected by, and may actually benefit from, marginal losses of blue oak woodlands to agriculture and urban land use. These results highlight the importance of non-use and ecosystem services values, and the importance of coordinating land use policies at spatial scales above the municipality level.

Introduction

Many scientists predict that global warming will cause suitable habitat ranges to shift for many plant species around the globe. To the extent that proximity to particular vegetation types enhances local residents' welfare and/or these shifts affect services valued by all of society, such geographic shifts in ecosystems may significantly affect human welfare. In California, climate change will cause significant shifts in many vegetation types over the next century. Because California is a biologically diverse area with many unique habitats, the welfare changes from these vegetation movements may be substantial. Blue and valley oak habitats, two important ecosystems, are predicted to shrink and move north and upslope (Kueppers et al. 2005; Hannah et al. 2008). These habitats will most likely be replaced by herbaceous vegetation (Ritter 1988; Lenihan et al. 2003). This paper aims to measure the local welfare change from marginal shifts in blue oak habitat in California's Tulare Lake Basin (Fresno, Kern, and Tulare counties) due to climate change as measured by marginal willingness to pay.

Valuing regional welfare losses of future shifts in the suitable habitat range of blue oaks is necessary to identify the magnitude of these possible future losses (or gains). The current literature recognizes the direct welfare effects of climate change through effects on agricultural production and willingness to pay to live in a location with a particular climate as described by temperature and precipitation; see, for example, Howitt, Medellín-Azuara, and MacEwan (2009) and Timmins (2003). However, these welfare measurements ignore the indirect effects of climate change on willingness to pay to live in a location (Howard 2011). This paper estimates the marginal value of the amenities associated with several land cover types, including blue oak habitat. These values represent additional costs that should be added to the current literature's marginal cost estimates of carbon emissions.

Regardless of whether the indirect cost of climate change is significant or insignificant, the results are important in terms of their implications for future analysis. Significant differences between the marginal implicit prices of blue oak habitat and other vegetation types, particularly those that are likely to replace it, indicates the need to use structural models, such as a semiparametric hedonic model (Bajari and Benkard 2005) or a Tiebout sorting model (Klaiber and Phaneuf 2009; Walsh 2007), to explore the effects of non-marginal changes in suitable habitat ranges. Insignificant differences indicate the need to focus on non-use values, such as existence value, and location-independent use values, such as the value of ecosystem services that decay slowly over space, that are only affected by non-marginal changes.¹

In addition to contributing to the economic literature regarding the value of habitat preservation, this paper provides information to policymakers making decisions regarding land use and habitat preservation in the study area. By valuing land use related amenities embedded in property prices, this paper informs the tradeoffs among urban development, agricultural use, and the preservation of natural landscapes. Significant differences between the marginal implicit prices of blue oak land cover and agricultural and urban uses indicate a possible welfare benefit from preserving oaks. Because 76% of blue oaks are on private property in the Tulare Lake Basin, this preservation is likely to require market instruments, e.g. development fees and preservation payments, whose values should be based on non-market valuation studies such as this one. Alternatively, insignificant differences between the marginal implicit prices of blue oaks and alternative land uses indicate that conservation resources should potentially be focused on other at-risk habitats if policymakers believe that the recreational and non-use values of blue oaks are also relatively insignificant.

¹ Klaiber and Phaneuf (2009) and Walsh (2007) estimate the willingness to pay for open space, while Bajari and Benkard (2005) estimates the welfare loss resulting from housing price appreciation.

This paper makes three methodological contributions. First, it captures the multi-scale capitalization of land cover types into housing prices. Land cover types, including vegetation types, produce a variety of spatial amenities which may dissipate over different distances. As a consequence, the capitalization of land cover types into properties may occur at a variety of scales of analysis. This paper captures the marginal implicit prices that households place on these land cover types by carefully defining several variables that measure different aspects of a house's proximity to these land covers. Second, this paper obtains asymptotically unbiased coefficients regarding the effect of land cover types, including vegetation, on housing prices. I develop instrumental variables based on average soil characteristics at various scales of analysis. By utilizing a two-stage least squares approach, I avoid the use of spatial fixed effects which bias the estimate of overall land cover capitalization by looking at only within-neighborhood variation of amenities (Abbott and Klaiber 2010; 2011).

Third, this paper uses cluster robust standard errors to demonstrate that the use of heteroskedasticity robust standard errors may lead to overstating the statistical significance of coefficients on land cover variables when estimating the determinants of the price of housing. Many papers in this literature impose neighborhood-level data, including neighborhood vegetation and open space data, at the housing level and assume that regression error terms are independently distributed. This can potentially result in standard errors that are biased downwards, particularly in the presence of spatial autocorrelation. As a consequence, some of the findings in the open space and urban-forestry literatures may be due to imposing an incorrect assumption. While the potential problems of imposing macro data at a micro-scale are not discussed in these literatures, some papers utilize approaches that may mitigate the effects with additional data or assumptions. Unfortunately both strategies have weaknesses of their own.

First, they construct proxies for land cover amenities at the micro-level scale. However, this method does not address the likely presence of spatial autocorrelation at the neighborhood scale. Second, they adjust estimates for spatial autocorrelation through the use of spatial fixed effects or the Haining (1993) method; the former method results in biased estimates of the overall value of land cover types and the latter method suffers from the difficulty of defining the nearest neighbor. I implement an alternative approach which does not suffer from either of these weaknesses. Specifically, the use of cluster robust standard errors results in unbiased estimates of overall capitalization, unbiased standard errors, and the definition of neighborhood as the scale of spatial aggregation chosen for variables.

Following Irwin and Bockstael (2001) and Irwin (2002), I use a reduced form hedonic model to estimate the first stage of the Rosen (1974) two-stage procedure. In general, the identification problems facing a first stage hedonic analysis are specification error and omitted variable bias; multicollinearity is also a frequent problem. An additional problem facing analyses that address land use issues, including this one, is the endogeneity of land cover types that are predominately privately owned (Irwin and Bockstael 2001). Before estimating the various models, I utilize variance inflation factors (VIFs), common indexes, and the corresponding variance-decomposition proportion matrices to demonstrate that the coefficients corresponding to the proxy variables for land cover amenities are valid. To address the problem of specification error, I choose a log-log specification for the hedonic price function using a linear Box-Cox transformation, the Ramsey reset test, and the link test. To address omitted variable bias and endogeneity of land cover types, I use a two-stage least squares (2SLS) estimator where a parcel's soil properties and the average of these properties at the census block and census block group levels are utilized as instruments for the various proxies of land cover amenities. Last, I

calculate cluster robust standard errors to adjust downwards previously upward-biased t-tests resulting from the use of neighborhood-level variables at the property level when spatial autocorrelation is present.

The results demonstrate that households in the Tulare Lake Basin do not differentiate between vegetation types (land cover types primarily characterized by the presence of vegetation), regardless of whether vegetation is disaggregated by habitat types (conifers, oak forest, blue oak woodland, other oak woodland, herbaceous, shrubs, and wetlands) or tree density (forest, woodland, grass/shrub lands). These results imply that there is no indirect welfare effect of climate change through property prices on households. As consequence, any additional effort to conserve blue oaks over other natural habitats is justifiable based on non-use value and ecosystem service value criteria only.

The results also demonstrate that households perceive a difference between vegetation and urban land covers at the neighborhood scale, but they do not at the within neighborhood scale. There is some evidence that households differentiate between vegetation and other non-urban land cover types (agriculture, barren land, and water) at the neighborhood and within neighborhood scales. These results indicate that households prefer to live in neighborhoods with more urban and agricultural land, and less vegetation. Households prefer to live adjacent to more vegetation and urban land, and less agricultural land. Of course, as is always the case for land use studies, the exact welfare effects depend on the spatial distribution of the Tulare Lake Basin population and its relationship with the spatial distribution of blue oaks.

These results differ from previous hedonic studies in the urban forestry and open space literatures, which in general find that access to forests and adjacent open space increase property prices. In the urban forestry literature, Powe et al (1997) finds that housing price increases with

forest access, and Tyrväinen and Miettinen (2000) find that property prices decrease with distance to forested areas and increase with views of forests. In the open space hedonic literature, Irwin and Bockstael (2001) find that property price increases with the amount of surrounding open space, regardless of whether it is developable or privately owned. Irwin (2002) finds that households value open space differently by type, e.g. households value pasture more than forest; however, Irwin (2002) concludes that the majority of open space value is derived from land not being developed. Though not emphasized in the paper, Irwin (2002) finds the cost of developing privately owned forests to surrounding landowners may be negligible in the case of low density development, and actually a benefit in the case of high density development. In San Joaquin County, Kuminoff (2009) finds a quadratic relationship for the MWTP for surrounding cropland, such that households with little surrounding cropland are willing to pay more for crop land than those with an abundance of it. Finally, Standiford and Scott (2001), the only existing hedonic study focusing on the valuation of California oaks, finds evidence that California native oaks increase property values in Southern Riverside County.² Possible reasons for the difference between the results of this paper and the current literature are discussed in Section VI.

The paper is structured as follows. Section II reviews the key literature on hedonic methods. Section III reviews valuation studies that utilize these methods for the purpose of valuing vegetation, open space, and climate with an emphasis on variable choice. Section IV discusses the choice of model and derives the estimator. Section V discusses the data. Section VI summarizes the key findings. Section VII concludes with a discussion of the broader implications of these results and the direction of future work.

² The Standiford and Scott (2001) study has several drawbacks: (1) the authors use the assessed value of houses instead of sales prices, (2) the study uses data for houses sold over a twenty-one year period ignoring structural changes in the California housing market, (3) the authors only control for distance to native oaks, which only partially captures oak amenities, and (4) the authors do no instrument for distance to oaks even though privately owned open space is endogenous.

The Reduced Form Hedonic Method

This paper utilizes hedonic regression, a revealed preference technique, to estimate household preferences for vegetation types. While revealed preference methods have recognized weaknesses, the available data allow me to address these problems in the context of my specific empirical analysis. The main drawbacks of these methods are that data are not always available, market distortions, such as market power and government policies, can affect the market outcome, and the resulting value estimates do not fully capture the value of a habitat. However, these drawbacks are less problematic in this analysis because this paper analyzes sales of single family residences from 1997 to 2003. There is little market power in the housing market, and I am able to control for government policies such as zoning.

The hedonic regressions used in revealed preference analysis only capture location-dependent use values of habitats, and do not capture non-use values, recreational values of non-residents, and location-independent use values.³ These omissions are unlikely to be significant in my empirical context. Because individual vegetation types in Kern County are unlikely to disappear completely within the next century and non-use values of vegetation are not specific to the Tulare Lake Basin, non-use values are unlikely to be significantly affected by shifting suitable habitat ranges. In addition, recreational values of non-residents are likely to be relatively small because 62% of the Tulare Lake Basin was privately owned in 2000 and unavailable for public recreation use. Seventy-six percent of blue oak habitat is privately owned and the government leases portions of publicly owned oak woodlands for private use. Last, while the hedonic method fails to capture the value of location-independent amenities or amenities that

³ Direct use value is the value that economic agents gain by consuming consumptive (e.g. timber and crops) and non-consumptive (e.g. recreation and aesthetics) habitat services and indirect value is the value of regulation services (e.g. erosion prevention, pest control, and water purification) services. Non-use values include bequest, altruist, and existence values, which are the values of preserving habitat for future generations, for others in the current generation, and for the knowledge of its existence (Pascual et al., 2010).

decay slowly over distance, e.g. water purification or carbon sequestration, the hedonic model estimates are good approximations of the full welfare change if the value of or change in value of location-independent services are relatively small. This is true for the marginal changes in land cover types analyzed here.

The basic argument underlying hedonic models is that the price of a property will reflect productivity differentials in a competitive land market. The environmental characteristics of a property and its surrounding areas should be reflected in property prices because they affect consumer and producer productivity. If the study area is one market and perfect information and mobility hold then the price of a property j , p_j , is a function of its structural housing, neighborhood, and environmental characteristics, i.e. $p_j = p(\vec{Z}_j)$ where \vec{Z}_j is the vector of K characteristics associated with the composite good, i.e. housing. The hedonic price function for housing represents the market equilibrium where the market price for each quantity of characteristic equates demand and supply. As a consequence, the marginal implicit price of a characteristic is the derivative of the price function with respect to that characteristic and is equal to the marginal willingness to pay (MWTP) for that characteristic (Palmquist 1999; Freeman 1996).

Rosen (1974) presents a two-step methodology for estimating the supply and demand of characteristics using the hedonic method. Assuming that consumers are price takers in the housing market, the supply side can be ignored and the Rosen two-stage procedure simplifies to first estimating the hedonic price function by regressing property price on housing characteristics and obtaining households' marginal willingness to pay for each characteristic, and then estimating the inverse demand function for a characteristic by regressing the implicit price of that

characteristic obtained from the first stage on the factors that influence demand in order to estimate the willingness to pay for non-marginal changes.

Potential empirical problems

Several problems may arise when implementing this procedure. In the first stage, specification error and omitted variable bias are problems. Specification error arises because economic theory does little to restrict the possible shape of the hedonic price function. While there is still little consensus in the literature about the best functional form to use, many authors choose to utilize simple functional forms based on the results of Cropper, Deck, and McConnell (1988) that find that the linear Box-Cox and linear functional forms produce the smallest errors compared to quadratic Box-Cox and other common distributions (semi-log, double-log, quadratic) when important variables are omitted. Alternatively, Bajari and Benkard (2005), Bajari and Kahn (2005), and Heckman, Matzkin, and Nesheim (2003) avoid this problem altogether through the use of non-parametric estimators.

Omitted variable bias in the first stage regression can be addressed in several ways: choice of functional form, instrumental variables, and spatial dummies. As mentioned above, Cropper, Deck, and McConnell (1988) determine which functional forms are the most robust to omitted variable bias. An alternative strategy is to instrument explanatory variables that are most likely to be correlated with omitted variables (e.g. Irwin 2002). Another strategy is to use spatial dummies to represent unobserved variables (e.g. Chattopadhyay 1999). However, Abbott and Klaiber (2010; 2011) argue that these spatial fixed effects result in biased overall estimates of capitalization if the good of interest capitalizes at a scale equal or greater than the scale of the spatial fixed effects. While spatial fixed effects may result in unbiased estimates for smaller scale capitalization of the good of interest by looking at within neighborhood variation exclusively, the

discarding of between neighborhood variation results in biased overall estimates of value. Thus, there is a tradeoff between omitted variable and excluded capitalization biases.

In the second stage, identification and endogeneity problems can arise. The identification problem is the result of many, if not all, of the explanatory variables in the second stage regression being explanatory variables in the first stage regression, while the endogeneity problem arises because consumers simultaneously choose the implicit price and quantity of a characteristic (Bishop and Timmins 2008; Palmquist 1999). While several solutions have been proposed for these problems (Brown and Rosen 1982; Ekeland, Heckman, and Nesheim 2002; 2004), other papers avoid the Rosen second-stage altogether by either replacing it with a preference inversion procedure (Bajari and Benkard 2005; Bajari and Kahn 2005) or only estimating the first stage of Rosen's procedure. I follow the latter strategy and only estimate the MWTP for land cover types. While estimating MWTP provides valuable insights, the results should not be used to measure the welfare change from non-marginal movements of land cover types.

Several conditions must hold in order for the first-stage hedonic estimates of marginal willingness to pay to be unbiased. First, valid instruments must exist for all relevant endogenous variables. Second, households must choose from a continuous choice set. Violations of this assumption may bias MWTP estimates because the equilibrium implicit price and the marginal willingness to pay for a characteristic will not in general be equal. The direction and magnitude of this bias is unknown ex-ante, though in the aggregate, it may be small because some households will choose properties with more of the discrete characteristic than is optimal under continuity and others will choose properties with less.⁴ Third, there is no sticky decision making because the existence of moving costs bias the willingness to pay estimates downward. Kuminoff

⁴ This bias also decreases as housing characteristics become approximately more continuous.

(2009) argues that the assumption that mobility costs are zero is justifiable if the study region is sufficiently small to have insignificant moving costs within its boundaries and sufficiently isolated such that moving costs increase substantially by leaving the region.

In addition to assuming that the above conditions hold, I assume that all households have already optimized by choosing the Tulare Lake Basin to live and that wages are constant within the region.⁵ In a full wage hedonic model, wages are allowed to vary by location because households are willing to accept lower or higher wages to live in more or less desirable locations. Failure to account for wage changes results in the marginal implicit price of a good being an inaccurate measurement of the marginal willingness to pay for that good. Because the majority of hedonic studies focus solely on the housing market, these assumptions are implicit in most hedonic studies.⁶

Choice of Variables: Explanatory and Instruments

Three strands of the valuation literature inform my econometric model of property price in the Tulare Lake Basin: urban forestry, open space, and climate valuation. Chiefly, they provide guidance for my choice of explanatory variables and instrumental variables. The variables of interest when valuing marginal shifts in land covers are the non-market and amenity services produced by each land cover type. The underlying production function of these services cannot be estimated because whether or how much a household consumes a service and how they value each service is unobservable. Instead, each of these literatures develops proxy variables and assumes that the level of these services change with these proxies (Klaiber and Smith 2009). Following this approach, this paper uses land use, temperature, and precipitation variables as

⁵ This assumption is necessary because I do not observe wages.

⁶ In essence, this assumption is that households take wages and housing prices into account when making inter-regional sorting decisions, while households only account for housing prices when making the intra-regional sorting decision. While vegetation may affect inter-regional sorting, the magnitude of the capitalized values from inter-regional sorting is likely small relative to the capitalized values from intra-regional sorting.

proxies for the amenities that land covers and climate produce. The urban forestry and open space literatures aid me in the selection of proxy variables for land cover amenities. The open space literature provides guidance in the choice of instrumental variables to control for the endogeneity of these proxy variables. Last, the climate change literature helps in the selection of proxy variables for climate amenities that affect household welfare and habitat location.

Urban Forestry

In the urban forestry literature, the definition of forest and other vegetation types and specification error are common problems. Defining forest types is problematic for several reasons. First, forest types, and vegetation types in general, are highly collinear (Garrod and Willis 1992). Because dropping variables is the primary solution to multicollinearity, information loss and omitted variable bias become potential problems. Second, choosing the appropriate degree of specificity of proxy variables for forest types also creates a tradeoff between multicollinearity and omitted variable bias. Aggregate forest types may fail to capture unobserved services that are particular to a sub-group of trees and their relationship to the landscape. Thus, failure to disaggregate forest type may result in omitted variable bias. On the other hand, disaggregating forest types sufficiently may be extremely difficult or impossible due to multicollinearity and lack of data. The problem of specification error arises because aesthetics are a complicated mix of landscape and vegetation characteristics whose functional relationship is unknown (Price 2003). Because the relationship of characteristics that make up aesthetic value are too complex, Price (1995) argues that hedonic price models are not well suited for estimating the monetary value of landscape features.

No solution is specified by Price (1995; 2003) other than to avoid the use of hedonic methods for valuing landscape characteristics unless a subjectively determined index of landscape quality

is used. However, I identify several possible solutions. First, analysis should attempt to reduce some of these issues by focusing on valuing land cover types, rather than focusing on the individual species that make up land cover types and the characteristics of land cover types, such as tree density. By focusing on land cover types more generally and estimating their average marginal value, analysts avoid the complexity of how individual species and landscape characteristics relate to one another to create aesthetic value. Second, particular focus should be placed on household access to and location with respect to each type of landscape (Powe et al. 1997). Third, analysts should attempt to develop estimation strategies that address multicollinearity and specification error, including better definitions of proxy variables to estimate particular vegetation services. These proxy variables for non-market services should include complex indices, which attempt to measure one non-market amenity (e.g. Powe et al. 1997; Geoghegan, Wainger, and Bockstael 1997) and/or proxy variables that capture amenities that capitalize at different spatial scales (e.g. Tyrväinen and Miettinen 2000; Abbott and Klaiber 2010). Last, analysts should conduct robustness checks of their results by varying the specification of the hedonic price equation and the level of aggregation for proxy variables of vegetation. This paper utilizes all four of these strategies.

Open Space

The open space literature addresses how to value heterogeneous open space. While open space is often differentiated by type of land use and ownership, land cover type is an alternative means by which to disaggregate open space. As a consequence, many of the estimation issues raised in the open space literature also apply when valuing land cover types. One such issue raised by Irwin and Bockstael (2001) is the endogeneity of privately-owned open space. This

endogeneity arises for two reasons: privately held open space is subject to the same economic forces as residential housing, and spatial autocorrelation exists.

To correct for endogeneity, Irwin and Bockstael (2001) use variables that proxy for the opportunity cost of developing a specific property (parcel slope, soil drainage ability, and soil quality) as instruments for the percentage of open space. While the authors argue that these variables are exogenous to the residential housing market because the hedonic price equation is only estimated for single family homes, they are correlated with the amount of each land use. Irwin (2002) and Kuminoff (2009) use similar approaches. Therefore, this paper uses these opportunity cost variables, along with other soil variables that affect vegetation type, to instrument for land cover types.

Climate

When estimating household welfare from land cover types, the omission of climate variables poses a potential identification problem. This is because precipitation and temperature affect the location of suitable habitat ranges, and directly affect household welfare through their preferences for climate.⁷ Many economists have used hedonic methods to estimate the willingness to pay for climate, including Cragg and Kahn (1997), Maddison and Bigano (2003), Timmins (2003), Rehdanz (2006), and Rehdanz and Maddison (2008). These authors differ in their choice of proxy variables for climatic amenities, however including a large set of these variables in a hedonic regression would likely result in multicollinearity. Thus, the most appropriate proxy variables depend on the precise question of interest. Therefore, I initially

⁷ For example, Kueppers et al. (2005) uses climate variables (mean temperatures of the coldest and warmest months, total annual precipitation, and April–August precipitation) to predict the current and future locations of California blue and valley oak.

utilize average precipitation in the driest and wettest months and mean temperature in the warmest and coldest quarters as explanatory variables for housing price.⁸

Methodology

The goal of this paper is to estimate the marginal implicit price of blue oak woodland and the land cover types that are likely to replace it by estimating the hedonic price function for properties in the Tulare Lake Basin. The marginal price is obtained by differentiating this function with respect to the proxy variables for blue oak amenities and calculating the mean willingness to pay.

The hedonic price equation

To capture the full value of each land cover type and avoid multicollinearity, I construct five variables that are proxies for the potentially spatially distinct services produced by each land cover type. First, I construct a dummy variable for whether a land cover type is within 0.1 km of that parcel to capture its aesthetic and use values to the owner. The land cover data are at a 0.1 km resolution, since this is the most precise measure available. Second, I construct a dummy variable for whether a land cover type is within 0.5 km of a parcel to measure the amenities obtained by having the land cover type within walking distance. Third, I construct a dummy variable for whether a land cover type is within 1.0 km of a parcel to measure amenities obtained by having a land cover type within a neighborhood.⁹ Fourth, I construct the percentage of the house's census block covered by a land cover type to measure both adjacent and walking distance amenities; the mean and median census block size are 0.46 km² and 0.03 km²,

⁸ Due to the high level of collinearity between these climate variable and elevation, I replace the monthly precipitation and quarterly temperature variables with their annual counterparts, mean annual temperature and annual precipitation.

⁹ I construct all three distance proxy variables from the edge of the property to the edge of the land cover type.

respectively.¹⁰ Fifth, I construct the percentage of the house's neighborhood (census block group) that is covered by a land cover type to proxy for amenities from the overall neighborhood's character; the mean and median census block group sizes are 39.58 km² and 1.01 km², respectively.¹¹

The hedonic residential price function is

$$p_j = p(\vec{Z}_j; \beta) + \varepsilon_j$$

where \vec{Z}_j is the vector of house j 's characteristics and β is the corresponding parameter vector. Based on the scale of the variable, housing characteristics are subdivided into household-specific characteristics and neighborhood-specific characteristics. The former group is further subdivided into several groups based on the type of variable: structural housing characteristics (H_j), distances to urban areas (U_j), climate characteristics (C_j), education characteristics (E_j), and within-neighborhood land cover characteristics (T_j); within-neighborhood land cover characteristics include the dummy variables for whether a land cover type is adjacent to or within walking distance of house j or the percentage of house j 's census block covered by a particular vegetation type. Similarly, neighborhood-specific characteristics are subdivided into neighborhood-level non-land cover characteristics (N_k), and neighborhood-level land cover characteristics (V_k) where k is the neighborhood in which j is located. The neighborhood-level land cover characteristics consist of the percentage of neighborhood k that is covered by each land cover type or dummy variables for whether house j is within 1.0 km of each land cover type.

Econometric choices

¹⁰ If I assume that a census block is roughly a circle (square) then the mean and median radii (side) are 0.18 (.32) km and 0.09 (0.16) km, respectively.

¹¹ If I assume that a census block group is roughly a circle (square) then the mean and median radii (side) are 1.68 (2.97) km and 0.57 (1.00) km, respectively.

There are several econometric issues that must be addressed in this paper in order to obtain an unbiased estimate of the marginal willingness to pay for blue oak woodlands and other land cover types. First, the functional specification of the hedonic price function is unknown. As is standard in this literature, a Box-Cox transformation, Ramsey reset test, and the link test are utilized to select the preferred functional form for the hedonic price function.

Second, the proxy variables for the amenities of land cover types are endogenous and may suffer from omitted variable bias. A two-stage least squares estimator is utilized to instrument for endogenous land cover variables. Eight instruments are defined at each level of capitalization to control for potential endogeneity due to private ownership and omitted variable bias; at each level of capitalization, quadratic terms are included for three of these variables to capture the potential non-linearity of environmental relationships. To instrument adjacent land cover, eight instrumental variables are constructed (using average values across all soil types) at the parcel level: a dummy for a slope above 15%, a dummy for whether the property's dominant soil is characterized by poor drainage, a dummy for whether the property's dominant soil is characterized by good drainage, a dummy for whether there are prime agricultural soils, a dummy for whether there are agricultural soils of statewide importance, the average available water capacity of the parcel's soil, the average share of clay in the parcel's soil, and the average maximum depth of the parcel's soil; quadratic terms are included for the latter three characteristics.¹² The first five instruments parallel those used in Irwin and Bockstael (2001) and Kuminoff (2009). The latter three soil variables are utilized by Kueppers et al. (2005) to predict the future locations of blue and valley oaks because of their importance in regulating soil

¹² Available water capacity is the amount of water available to plants that is stored in the soil (USDA, 1998).

moisture.¹³ The use of these variables as instruments is justified because residential households do not have preferences over these specific soil variables per se, and these variables affect the type of vegetation on a property. To instrument for neighborhood-level land covers, the average of each variable is calculated at the census block group level using as weights the percentage of the census block group covered by each soil type. To instrument for census block-level and walking distance land cover variables, the average of each variable (measured using the dominant soil type) is calculated at the census block level using all types of properties, not just single family residential.¹⁴ The changes in the use of average versus dominant soil values and the use of weights when calculating instruments at the various spatial scales are in order to increase the overall amount of information captured by the set of instruments.

Third, omitted variable bias, not associated with spatial autocorrelation, can potentially result in biased coefficient estimates. In particular, Anderson and West (2006) argue that open space hedonic regressions omit many spatial variables that are correlated with open space variables. To avoid biased estimates, Anderson and West (2006) uses neighborhood fixed effects to absorb these omitted variables. However, the fixed effects only partially control for omitted variables because of either incorrect neighborhood definitions or within-neighborhood omitted variables.¹⁵ In addition, the inclusion of neighborhood fixed effects biases the overall value estimates of land cover types (Abbott and Klaiber 2011). Including neighborhood fixed effects achieves efficient

¹³ Kueppers et al (2005) also uses climate (temperature and precipitation) variables to predict the future locations of oaks. Specifically, Kueppers et al. (2005) uses mean temperatures of the coldest and warmest months, total annual precipitation, and April–August precipitation to predict the current and future locations of California blue and valley oak. However, climate variables are not valid instruments because households have strong preferences over climate.

¹⁴ The average of the soil dummy variables at the census block scale is equivalent to the percentage of parcels in the corresponding census block for which the dummy variable equals one.

¹⁵ Within-neighborhood omitted variables can exist even when the neighborhood is correctly defined. These omitted variables are defined at a spatial scale smaller than the neighborhood-level. For example, an important housing characteristic could be omitted, e.g. whether a house is adjacent to a particular land cover type, whose omission could potentially bias coefficient estimates and which will not be absorbed by a neighborhood fixed effect regardless of the neighborhood definition.

coefficient estimates for household-specific characteristics at the cost of omitting neighborhood-specific characteristics; this tradeoff exists at all spatial scales.¹⁶ Therefore, none of the econometric specifications in this paper include neighborhood fixed effects. Alternative solutions are to instrument open space variables (e.g. Irwin and Bockstael 2001; Irwin 2002) and to use a simple functional form that is robust to omitted variable bias (Cropper, Deck, and McConnell 1998). These latter two approaches are the methods that I employ.

Fourth, spatial autocorrelation due to omitted variables in a spatial error model can result in inefficient estimates. The open space literature uses two techniques to control for this type of spatial autocorrelation. Irwin and Bockstael (2001) use the Haining (1993) method of randomly drawing a subset of data from the pool of non-neighboring properties where the neighborhood is defined with varying radii. However, the coefficient estimates are not robust to the definition of nearest neighbor, indicating a potential problem with using this method to correct for spatial autocorrelation. In addition to using the Haining (1993) technique to correct for micro-level unobservables, Kuminoff (2009) also uses larger-scale spatial dummies (city and school district) to account for spatial autocorrelation at a macro-level. However, the Kuminoff (2009) estimates may not fully capture the value of amenities because the spatial fixed effects will absorb all amenities capitalized at the city and school district levels, and above (Abbott and Klaiber 2011). Due to the drawbacks of these two methods, I rely on the asymptotically unbiased properties of two-stage least squares because the dataset contains a large number of observations. In addition, I utilize cluster robust standard errors to control for intra-neighborhood spatial autocorrelation following Abbot and Klaiber (2011).

¹⁶ For example, fixed effects for smaller geographic units absorb more local unobserved variables and eliminate more global information about land cover amenities, while fixed effects for coarser geographic units preserve more information about land cover amenities and increase the possibility of inefficient estimates.

Last, standard error estimates are potentially biased downwards due to the imposition of neighborhood (census block group) level variables at the property level. Many of the variables in the model, including land cover variables, are at the neighborhood scale. As a consequence, the significance of neighborhood coefficients may be exaggerated due to standard errors that are biased downward (Moulton, 1990). This is because the heteroskedasticity robust standard errors often calculated when using ordinary least squares and two-stage least squares imply that the error terms are independently distributed when in fact they may be correlated; correlation is likely due to spatial autocorrelation within-neighborhoods. Because of the imposition of neighborhood variables at the micro-level, such that there is no variation for such variables at the neighborhood scale, the standard error estimates are potentially biased even for small levels of error correlation within the neighborhood. This is particularly true when the average number of observations per neighborhood, i.e. the number of housing sales per census block group, is large, such as in this paper (Moulton, 1990; Cameron and Trivedi, 2009). A solution is to estimate cluster robust standard errors at the census block group level (Cameron and Trivedi, 2009). Unlike the adjustment utilized by Moulton (1990), the standard error adjustments utilized in this paper vary by each variable's spatial correlation within neighborhoods (Cameron and Trivedi, 2009). Cluster robust standards errors are also robust to heteroskedasticity (Cameron and Trivedi, 2009).¹⁷

Data

The data for this model come from a variety of sources, including Kern County's Geographic Information System Development Services Agency, Tulare County's RMA GIS Mapping

¹⁷ The formula for the standard error estimate is given in Cameron and Trivedi (2009) as $\hat{V}(\hat{\beta}) = (X'X)^{-1} \left(\frac{K}{K-1} \frac{N-1}{N-L} \sum_k X_k u_k u_k' X_k' \right) (X'X)^{-1}$ where K is the number of neighborhoods, N is the number of observations, L is the number of regressors, X is the regressor matrix, X_k is the regressor matrix for the k^{th} neighborhood, and u_k is the vector of residuals for the k^{th} neighborhood.

Division, the Fresno County Public Works and Planning's Maps and GIS Information, the National Data Center (NDC), CoreLogic, the California Department of Forest and Fire Protection's Fire and Resource Assessment Program (FRAP), the California Department of Transportation (Caltrans), Cal-Atlas, the U.S. Census Bureau, the USDA's Natural Resource Conservation Service, WorldClim, and the California Department of Education. ArcGIS was used to integrate the data sets at the property level and to construct the spatial variables of interest at the property and neighborhood scales. Table 2.a defines all variables.¹⁸ Table 3.a and Table 4.a summarize the relevant variables at the property and census block group levels, respectively, and their predicted signs. Map 1 depicts the location of land cover types and the outline of census block groups, while Map 2 depicts the location of census blocks.

I group land covers into twelve land cover types by ecosystems. Following FRAP's ten major land cover classes, the initial ten land cover types are: agriculture, barren, conifers, desert, hardwood, herbaceous, shrubs, urban, water, and wetlands. To isolate the ecosystem of interest, blue oak woodlands, I further subdivide hardwood into hardwood forest and hardwood woodland following the FRAP's thirteen land cover subclasses. Finally, I subdivide hardwood woodland into other oak woodland and blue oak woodland. See Table 1 for a breakdown of acreage by land cover type (FRAP, 2002).¹⁹

¹⁸ Two data issues should be noted. First, the variables measuring urban land cover may include urban open space. The metadata for the land cover data indicates that urban land cover was constructed using census data at the census block level and using Department of Conservation (DOC) Farmland Mapping Program data. The data were updated to account for unhabitated publically owned land (FRAP, 2002). However, I identify few parks within the major urban areas of the study region using visual inspection. This may partly result from the spatial resolution of this land cover data being 0.1 square kilometers, which may result in the omission of smaller urban parks. Second, the zoning data were only available for future time periods. The zoning data are from 2010, except in the case of the Tulare County zoning data which are from 2007. I include these zoning variables to capture expected future zoning.

¹⁹ In terms of predications about the signs of land cover types, I predict low density vegetation types (blue oak, other oak, herbaceous, and shrub) and water to have a positive effect on property prices at the within-neighborhood and neighborhood scales. At both spatial scales, I expect barren, desert, urban, and wetland land covers to have negative effects. At both spatial scales, I am uncertain with respect to agriculture because of its intensive nature in this area and with respect to conifers and hardwood forests due to the results of Irwin (2002) discussed earlier. Finally, I predict that urban land has a negative effect at the within-neighborhood scale. However, I am uncertain with respect

I examine residential houses sold between 1997 and 2003. Three factors drove the selection of the time period. First, the land cover data for the Forest and Range 2003 Assessment are most consistent for housing sales around 2003, so I exclude houses sold before 1997. Second, this period excludes houses sold after 2003 because of the housing bubble in the mid to late 2000s. Third, this choice of cut off dates places the 2000 U.S. Census at the center of the relevant time period.

Data cleaning is necessary to remove observations with missing and incorrect data and drop outliers that may potentially drive the results. As a means of addressing speculative transactions, I exclude any housing transaction for which the house was sold again within 365 days. In addition, I exclude homes that are sold before the current house is built in order to exclude any sales of empty lots, and I drop homes that are missing a building date. To ensure that I am looking at single family residences, I also drop housing sales for which CoreLogic and NDC do not agree in terms of their land use classification. After calculating the sales price of the house in terms of 1997 dollars, I apply two additional criteria in order to eliminate outliers. First, unlike Bishop and Timmins (2008) who drop houses in the top and bottom 1% of the housing price distribution, I drop housing sales that are in both the top and bottom 1% of the price per parcel acre and the price per square foot of housing distributions; this alternative method avoids dropping a disproportionate number of rural properties.²⁰ Second, I exclude houses whose area or whose number of floors, baths, or bedrooms are greater than five times the mean, or equal to zero in the case of the number of floors and building area. After applying these criteria, 168,271

to urban land at the neighborhood scale due to the potential household preference for living within neighborhoods with urban conveniences.

²⁰ The CoreLogic housing sales data exclude housing sales between family members. As a consequence, there are no housing sales with the price of zero that need to be dropped, so the bottom 1% contains only market sales with positive sale prices.

housing sales in 1,181 census block groups remain. Maps 3 and 4 depict the locations of housing sales.

Results

Overall, the results support the hypothesis that households do not differentiate between vegetation types, a subset of land cover types characterized by the presence of vegetation.²¹ Vegetation, which includes blue oak woodlands, appears to decrease property prices at the neighborhood scale relative to urban land. There is also evidence that vegetation decreases property prices relative to agriculture at the neighborhood scale, and increases prices relative to agriculture at the within neighborhood scale.

This section includes six subsections. The first four address five econometric problems (multicollinearity, heteroskedasticity, specification error, omitted variable bias, and endogeneity), and the last two present the final results with heteroskedasticity robust standard errors and cluster robust standard errors, respectively. In the final two subsections, I discuss the results under the a priori preferred specification and under a variety of sensitivity analyses: the number of endogenous land cover types, the number of instruments, the functional form, the proxy variables for land cover amenities, and the definition of land cover types.

Multicollinearity

I begin with an ordinary least squares regression including all potential variables; specification (1) in Table 5.²² Land cover amenities relative to urban land for eleven land cover types (agriculture, barren, conifers, desert, oak forest, other oak woodland, blue oak woodland, herbaceous, shrubs, water, and wetlands) are captured at the within-neighborhood and neighborhood scales using the percentages of a property's census block and census block group

²¹ Vegetation land cover includes conifers, desert, oak forest, other oak woodland, blue oak woodland, herbaceous land cover, shrubs, and wetlands.

²² See Table 29 for a mapping of specifications in Table 5 to specifications in all other following Chapter V tables.

covered by the corresponding land cover type.²³ Several unexpected signs (e.g. price decreases with the number of bedrooms, decreases with high school graduation, increases with poverty, and is non-decreasing with unemployment) indicate that multicollinearity is a potential problem. In response, I calculate two collinearity diagnostics: the variance inflation factor (VIF) for each variable, for which a value exceeding 10 indicates severe multicollinearity (Kennedy, 1998), and the common index for each corresponding coefficient, for which a value exceeding 30 indicates substantial multicollinearity (Hill and Adkins, 2003).²⁴ In addition to calculating the VIFs and common indexes, I assemble the corresponding variance-decomposition proportion matrix.²⁵ I find eleven and fourteen violations of the rules of thumb of 10 and 30, respectively, and a common number (the maximum common index) equal to 687.03. These violations indicate the presents of strong near dependencies (Hill and Adkins 2003). While multicollinearity is present, none of the variables of interest, i.e. the land cover variables, have aggregate variance-decomposition proportions over the threshold of 50%; this result holds even when we lower the critical conditional index value from 30 to 20, except for the percentage of the property's census block group covered by conifers (percveg30) which also has a VIF exceeding 10. This indicates that the multicollinearity does not harm the coefficients of interest (Hill and Adkins 2003; Belsley 1991), except for the coefficient corresponding to amenities from conifers at the neighborhood scale.

²³ Because all land cover types are mutually exclusive, I exclude urban land cover to avoid perfect multicollinearity.

²⁴ $VIF_j = \frac{1}{1-R_j^2}$ where R_j^2 is from regressing all other explanatory variables on variable j . The condition index corresponding to eigenvalue k of matrix X is the square root of the largest eigenvalue of matrix X divided by eigenvalue k of matrix X .

²⁵ Variance-decomposition proportions are depicted as a $K \times J$ matrix with the condition indexes as the first column (row titles) ranked from lowest to highest and variable names as the first row (column titles). The k - j variance-decomposition proportion is “the proportion of the variance of the j^{th} regression coefficient associated” with the k^{th} eigenvalue. In this matrix, the rows with indexes above 30 are interpreted as near dependent relationships and columns with aggregate variance-decomposition proportions (aggregated over the condition indexes exceeding 30) that exceed 0.50 are interpreted as variables involved in these near dependent relationships (Belsley, 1991).

Due to the high number of linearly dependent relationships, I redefine or drop variables that appear redundant and uncorrelated with the variables of interest to reduce the multicollinearity in the model. First, due to spatial multicollinearity resulting from the spatial configuration of Central Business Districts (Bakersfield, the City of Fresno, and Visalia) within the spatial range of the data (Tulare Lake Basin), I replace the distances to Bakersfield, City of Fresno, and Visalia with the distances to the nearest central business district and to the nearest urban area.²⁶

Second, I drop four variables regarding structural housing characteristics: whether a house has a garage because the variable is inconsistent between the NDC and CoreLogic datasets; I drop two housing quality variables because the variables are not defined clearly by the data provider, so it is difficult to ascertain what they capture; and I drop the number of bedrooms because it moves closely with the number of bathrooms (correlation coefficient of 0.57) and the square footage of the house (correlation coefficient of 0.59).²⁷

Third, I drop seven neighborhood demographic variables: the percentage of graduate/professionals, percentage of senior citizens, and percentage of children due to their unlikely connection to vegetation; I drop the percentage of vacancies and the percentage of unemployment due to the temporary nature of these 2000 Census variables, which are unlikely to hold over the study period; and I drop the percentage of high school graduates and the percentage of the population below the poverty line because both variables are highly correlated with median income, the percentage of college graduates, and the percentage of Hispanics.

²⁶ The correlation coefficient between the distances to Bakersfield and the City of Fresno is -0.86, and the correlation coefficient between the distances to the City of Fresno and Visalia is 0.70. See Heikkila (1998) for more on spatial multicollinearity.

²⁷ While the number of bathrooms is highly correlated with the square footage of the house (correlation coefficient of 0.72), it appears more frequently as an explanatory variable in the literature than the number of bedrooms. Plus, unlike the number of stories, there is no relationship between the number of bedrooms and vegetation that can be developed.

Fourth, I replace the seasonal temperature and precipitation variables with mean annual temperature and precipitation. In terms of affecting the values of land cover types, this change is potentially the most significant. However, the seasonal measures move so closely with the annual measures that little information is lost. The correlation coefficients between the mean temperatures of the warmest and coldest quarters and the annual mean temperature are 0.97 or above. Similarly, the correlation coefficient between precipitation of the wettest month and annual precipitation is 0.9975. Last, the correlation coefficient between precipitation of the driest month and elevation exceeds 0.93. Figures 1, 2, and 3 visually represents these correlations.²⁸

Last, I drop the measurement of land cover diversity (diversity10) at the neighborhood-level to avoid multicollinearity with land cover variables. In addition, this also conforms to the specifications in the related literature, e.g. Irwin and Bockstael (2001) and Irwin (2002). Last, dropping this variable simplifies the analysis and the marginal cost estimates of shifting blue oak woodland.

I test the sensitivity of the coefficient estimates and the multicollinearity diagnostics to the type of variables used to proxy for land cover amenities. Using my reduced set of variables, I estimate six additional specifications; specifications (2)-(7) in Table 5. Specifications (2)-(4) utilize twelve land cover types, while specifications (5)-(7) utilize six land cover types aggregated from the previous twelve. Specifications (2) and (5), like specification (1), capture land cover amenities at the within-neighborhood and neighborhood scales using the percentages of a property's census block and census block group covered by the corresponding land cover type. Specification (3) and (6) capture the within-neighborhood and neighborhood land cover amenities using a series of dummy variables for whether the corresponding land cover type is

²⁸ I also drop the measurement of land cover diversity (diversity10) at the neighborhood-level to avoid multicollinearity with land cover variables, and to simplify marginal cost estimates of shifting blue oak woodland.

within 0.1 km, 0.5 km, or 1 km of the property.²⁹ Finally, specifications (5) and (7) capture the within-neighborhood land cover amenities using the previous specifications' 0.1 and 0.5 km dummy variables and the neighborhood land cover amenities with the percentage of a property's census block group covered by the corresponding land cover type.

Focusing solely on the sign and significance of coefficients, the results are relatively robust across the specifications. None of the significant coefficients in specification (1) change signs and remain significant in specifications (2)-(7). In general, the reduction in explanatory variables results in an increase in the significance of the remaining coefficients. While some of the proxy variables for land cover amenities change significance depending on the specification, the signs of statistically significant variables are relatively stable for specifications (2)-(7); the only changes in signs for statistically significant variables occurs for the dummy variables for whether a property is within 0.5 km of conifers (p5kmdistw_30) and other oaks (p5kmdistOak) between specifications (3) and (4) and the dummy variable for whether a property is within 0.5 km of barren land cover (p5kmdistw13_20) between specifications (6) and (7). Excluding the percentage of a property's census block group that is publicly owned (public) from these regressions has no effect on the consistency of these estimates in terms of the sign of the coefficients (see Table 6).

Reducing the number of explanatory variables reduces the severity of multicollinearity. The common number is approximately cut in half, the mean VIF is greatly reduced, and the number of violations of the rules of thumb of 10 and 30 greatly decrease. In all but two specifications, multicollinearity does not harm the coefficients of interest as measured by aggregate variance-decomposition proportions over the threshold of 50% where a condition index of 30 is utilized as

²⁹ In this specification, none of the land cover types should be excluded from the model because the dummy variables are not mutually exclusive. In other words, a land cover type being within a particular distance of a property does not prevent another land cover type from also being within that specified distance.

the critical value for a near dependency. In specifications (3) and (6), which utilize only dummy variables to capture land cover amenities, violations occur for dummy variables capturing distance to urban and/or man-made land cover types at the 0.5 km and 1.0 km scales. If the critical value for a conditional index is reduced to 20, I find similar violations at the 0.5 km scale in specifications (4) and (7) and at the 0.1 km scale in specifications (6) and (7) for urban and man-made land covers. In all but one specification, multicollinearity does not significantly affect the coefficients corresponding to the proxy variables for land cover amenities as measured by the variance inflation factors. In specifications (4), which combine the use of dummies and percentages of neighborhood land cover, violations occur for the percentage of neighborhood covered by conifers. From these results, specification (2) is chosen as the a priori preferred specification (or base model) because the multicollinearity does not harm the coefficients corresponding to the proxy variables for land cover amenities according to the standard rules of thumb discussed earlier.^{30, 31}

These multicollinearity tests indicate that the dummy variables corresponding to urban land cover and the variables corresponding to conifer land cover should be interpreted carefully. For urban land cover variables, particular care should be taken when land cover types are aggregated. In addition, the near dependent relationship between the 0.5 km and 1.0 km urban variables indicate that both proxy variables likely measure overlapping urban amenities. Similarly, conifer

³⁰ If the critical value for a VIF is lowered to 9, I find violations in specification (2) similar to that of specification (4).

³¹ In general there are six near dependent relationships as identified by a critical value of approximately 30 for the condition index, which I will rank from strongest to least strong: (1) the intercept, the mean annual temperature, and elevation (2) mean elementary academic performance index, and to a lesser extent distance from the CBD (3) annual precipitation and the county fixed effects, and to a lesser extent distance from the CBD (4) the mean neighborhood tax and to a lesser extent elementary academic performance index (5) median neighborhood income, and (6) house square footage and the number of baths (full and half). Because strong collinear relationships may hide weaker ones, the variables in the stronger relationships may be included in the weaker relationships. Note that in specifications (3) and (6) whether a property is within 0.5 km and 1.0 km of urban and man-made (urban and agricultural) land cover results in a seventh near dependency, and to a lesser extent in specification (7).

coefficients should be interpreted with caution, particularly at the neighborhood scale when land cover types are more disaggregate. Table 6 re-estimates the specifications in Table 5 without public land included. While the aggregate measures for multicollinearity are approximately the same, conifer land cover types no longer violate the VIF rule of thumb of 10. Urban/man-made (urban and agricultural) land cover at the 0.1 km scale and urban/man-made land cover at the 0.5 km and 1 km scales still exceed an aggregate variance-decomposition proportion of 0.5 using the condition index cut offs of 20 and 30, respectively. In other words, the inclusion of the share of the neighborhood that is publically owned exacerbates multicollinearity for conifers, but not urban land.

The Breusch-Pagan/Cook-Weisberg test indicates that there is heteroskedasticity in all seven specifications estimated using OLS. As a consequence, I calculate Huber–White standard errors for all following specifications.

Model specification tests

Using a linear Box-Cox transformation, I test alternative model specifications for each definition of adjacency; see Table 7. The left side variable, the real price of housing, and all strictly positive right-hand side variables are transformed such that the transformation coefficient (λ) that all transformed right hand side variables share differs from the transformation coefficient (θ) for the left hand side variable. Because of the restriction that a transformed variable cannot equal zero, all of the land cover variables are untransformed. In specification (2), the left hand transformation coefficient equals 0.327 and the right-hand transformation coefficient is 0.944; the transformation coefficients are robust to the model specification. This is closest to a log-linear model of the simple specifications. Assuming specification (2), I also re-estimate the linear

Box-Cox transformation three more times assuming that $\lambda = 0$, $\theta = 0$, and then $\lambda = \theta$; see Table 8.

I utilize the Ramsey reset test and the link test to test a variety of simple functional forms (square root linear, linear-linear, log-linear, linear-log, and log-log) and the four transformations estimated above under specification (2). Both tests reject all nine functional forms (Table 9).³² The failure of these tests is likely due to omitted variable bias. Based on these results, I adopt the log-log specification because it is the simple functional form that performs best on both tests, i.e. has the lowest F-test and t-test values, and this matches the functional form chosen by Irwin and Bockstael (2001) and Irwin (2002). I will check the sensitivity of my final results to functional form, particularly with respect to the log-linear form and the unrestricted linear Box-Cox model.

Omitted variable bias

Using the log-log specification, I rerun ordinary least squares with spatial fixed effects at the neighborhood-level using model specifications (2), (4), and (5). I then test for whether the resulting coefficients are significantly different than the corresponding log-log specification without fixed effects. I also jointly test whether the within-neighborhood land cover variables and whether the within-neighborhood non-land cover variables differ with and without neighborhood fixed effects; see Table 10. In model specification (2), I reject the null hypothesis that coefficients corresponding to within-neighborhood land cover variables are jointly unaffected by the inclusion of neighborhood fixed effects. Within-neighborhood non-land cover variables are also jointly affected by the inclusion of neighborhood fixed effects. While these results are consistent with the presence of omitted variable bias and/or spatial autocorrelation, not all of the regression coefficients may be affected by omitted variables. Many of the coefficients corresponding to the land cover variables of interest, such as those corresponding to

³² Table 2.c defines all transformed variables.

blue oak woodland and herbaceous land covers, do not differ statistically on an individual basis between the two specifications.³³ While six coefficients gain or lose statistical significance with the inclusion of neighborhood fixed effects, only the log of distance to the nearest urban area changes signs and remains statistically significant; this change and the loss of significance of annual precipitation and elevation may be the result of insufficient within-census block group variation to achieve identification.

Similar results hold for model specifications (4) and (5). For model specification (5), the only significant differences from the results for model specification (2) is that the coefficients corresponding to herbaceous and shrub land covers become statistically insignificant when fixed effects are included. For (4), while the proxy variables for amenities of adjacent land cover types (0.1 km) have similar signs as the census block variables in (2) for the corresponding land cover types, the adjacent land cover variables are more statistically stable across the specifications with and without fixed effects than the census block variables. However, the statistical significance of proxy variables for amenities of land cover types within walking distance (0.5 km) is more unstable than their census block counterparts in model specification (2), and the coefficient corresponding to blue oak woodland at the 0.5 km scale significantly changes when fixed effects are included.

First stage of two-stage least squares

The previous results indicate that omitted variables are likely a problem and an earlier discussion indicates that privately owned land cover types are likely endogenous. As a

³³ Before accounting for the endogeneity of majority privately owned land cover types, blue oak woodlands appear to have a negative effect on household welfare compared to urban land at the within-neighborhood scale. In (2), the replacement of neighborhood-specific variables with neighborhood fixed effects slightly increases the magnitude of the negative and statistically significant coefficient corresponding to the percentage of the census block covered by blue oaks. However, this change is not statistically significant. The inclusion of neighborhood fixed effects switches the effect of agriculture at the census block scale from positive and insignificant to negative and significant. Lastly, the inclusion of neighborhood fixed effect has no statistically significant effect on the coefficient corresponding to herbaceous land cover, which is significant and negative.

consequence, I instrument for land cover variables using the opportunity cost and soil variables discussed earlier. Table 2.b defines the instrumental variables, and Table 3b provides summary statistics at the property level. Table 4.b summarizes the neighborhood instrumental variables at the census block group level.

In order to determine whether the instruments are strongly correlated with land cover variables, I regress each proxy variable for land cover amenities corresponding to majority privately owned land cover types (agricultural, blue oak, herbaceous, other oaks, and urban) and near majority privately owned land cover types (desert and shrubs) on the instruments at the corresponding spatial scale; see Tables 11 to 17. I regress the indicator variable for whether a land cover type is within 0.1 km of a property, which is the proxy variable for adjacent land cover amenities, on soil variables at the property scale. I regress the indicator variable for whether a land cover type is within 0.5 km of a property, which is the proxy variable for walking distance land cover amenities, on soil variables at the census block scale. I regress the percentage of a census block covered by a land cover type, which is the proxy variable for adjacent and walking distance land cover amenities, on instruments at the census block scale. Finally, I regress the indicator variable for whether a land cover type is within 1.0 km of a parcel and the percentage of a census block group covered by a land cover type, which proxy for neighborhood amenities, on soil variables aggregated at the census block group scale.

Table 13 displays these regressions for blue oak. While R^2 is relatively high for the percentage of blue oaks within a census block group (0.329), it is considerably lower for whether blue oaks are within 0.1 km of a property (0.066); the R^2 for whether blue oaks are within walking distance and the R^2 for the percentage of blue oaks covering the census block are 0.128 and 0.190, respectively. Almost all of the instruments are highly significant, including the

quadratic terms, in all regressions. In all five specifications, I reject the null hypothesis of the instruments jointly equaling zero at the 1%, 5%, and 10% significance levels using both the F-test and likelihood ratio test. Overall, the instruments appear to be relatively good measurements of blue oak variation based on their joint significance.

Table 11 displays the comparable regressions for agriculture. Overall, the variables appear to better explain the land cover variables for agriculture than the land cover variables for blue oak. The R^2 is relatively high when the dependent variable is at the neighborhood scale (0.242 to 0.318) and walking distance scale (0.250), while it is lowest at the adjacent scale (0.075) and at the census block scale (0.104). Almost all of the instruments are highly significant with the exception of the dummy variable for poor drainage at the property level (PoorDrain) when the dependent variable is at the 0.1 km spatial scale and average maximum soil depth at the neighborhood scale (cbg_avg_wgt_maxdepth) when the dependent variable is at the 1.0 km spatial scale. In all five specifications, I reject the null hypothesis of the instruments jointly equaling zero at the 1%, 5%, and 10% significance levels using both the F-test and likelihood ratio test. As was the case for blue oaks, the instruments seem to be relatively good measurements of agricultural variation based on their joint significance in each regression.

I conduct the same analysis for the other oak woodland, herbaceous, and urban layers as I did for blue oak woodlands and agriculture (Tables 12, 14 and 15). The R^2 values for the estimated specifications are between 0.035 and 0.239 for other oak woodland, 0.03 and 0.164 for herbaceous vegetation, and 0.127 and 0.279 for urban land cover. All of the potential instruments are individually and jointly significant, except available water capacity in some of the other oak wood land specifications and, in the case of the herbaceous layer, maximum soil depth when the dependent variable is at the 0.1 km spatial scale. In all five specifications for each land cover

type, I reject the null hypothesis of the instruments jointly equaling zero at the 1%, 5%, and 10% significance levels using both the F-test and likelihood ratio test. Based on the criterion of joint significance, the instruments seem to be relatively good measurements of other oak woodland, herbaceous, and urban variation.

I conduct the same analysis for the desert and shrub layers (Tables 16 and 17). The R^2 values for the estimated specifications are between 0.031 and 0.071 for desert land cover, and between 0.086 and 0.595 for shrub vegetation. While the R^2 values for desert land cover appear relatively low in all specifications, the exogenous variables are individually and jointly significant in all specifications for both for land covers. In all five specifications for desert and shrub land covers, I reject the null hypothesis of the instruments jointly equaling zero at the 1%, 5%, and 10% significance levels using both the F-test and likelihood ratio test. Based on the criterion of joint significance, the instruments seem to be relatively good measurements of both land covers.

The previous tables, Tables 11 to 17, do not take into account collinearity between instruments at multiple spatial scales, which may result in weak instruments. To explore whether multicollinearity is a potential problem, i.e. whether instruments at the property, census block, and census block group levels are collinear, I implement two sets of tests. First, I check for multicollinearity using the rules of thumb discussed earlier for sets of instruments: property and census block group instruments; census block and census block group instruments; and, finally, property, census block, and census block group instruments. While I find that together the property and census block group instruments have four common indexes over 30, indicating multicollinearity, there are no VIFs that exceed 10. Similar, results hold when I test the census block and census block group instruments together for multicollinearity. However, I find that all three sets of instruments (property, census block, and census block group) collectively violate the

common index rule of 30 nine times and the VIF rule of 10 thirteen times. This indicates that the use of two sets of instruments to identify land cover variables at two spatial scales is justifiable if care is taken to check for weak instruments. However, the use of three sets of instruments to identify land cover variables at three spatial scales is unadvisable due to the likelihood of weak instruments.

Second, I regress the instrumental variables at the property and census block group scales and all exogenous explanatory variables on the endogenous proxy variables for land cover amenities at the census block and census block group scales; see Table 18.a. Similarly, I conduct an identical analysis with the instrumental variables at the census block and census block group scales; see Table 18.b. These regressions assume that majority privately owned land cover types are endogenous. All of the instruments remain significant in a majority of the land cover specifications. Because multicollinearity affects the individual significance of coefficients, these regressions further support the argument that the multicollinearity of two sets of instruments at different spatial scales is unlikely to be a significant problem.

Two-stage least squares with heteroskedasticity robust standard errors

The results of the two-stage least squares regressions with heteroskedasticity robust standard errors are consistent with the current literature that the type of surrounding land cover affects housing price. In addition, this section contributes to the literature by demonstrating that land cover types capitalize at multiple spatial scales: within-neighborhood and neighborhood. Like the previous literature, the results imply that vegetation and natural (vegetation, barren, and water) land covers have generally positive effects on nearby housing prices, i.e. at the within-neighborhood scale. Unlike the previous literature, these results also demonstrate that vegetation and natural land covers have generally negative effects on housing prices at the neighborhood

scale; households desire to live within neighborhoods with urban conveniences.³⁴ However, the results also suggest a need to calculate cluster robust standard errors.

Preferred specification

The a priori preferred specification, specification (1) in Table 19.a, is chosen based on the results from the tests in the previous section and a priori expectations. First, I choose the percentage of a property's census block and census block group covered by a land cover type as the proxy variables for that land cover's amenities primarily because of their strong performance under the multicollinearity tests. In addition, distances, which traditionally measure the cost of access, are less ideal proxies for amenities from privately owned land cover types. Second, I choose the log-log specification because I reject all model specifications using the Ramsey reset and link tests and find that the log-log specification has the lowest F-test and t-test statistics of the simple functional forms (square root linear, linear-linear, log-linear, linear-log, and log-log). In addition, this matches the functional form chosen by Irwin and Bockstael (2001) and Irwin (2002). Third, I assume that land cover types that are majority privately owned (agriculture, other oak woodland, blue oak woodland, herbaceous, and urban) are endogenous based on a priori expectations. In addition, I do not have enough instruments (or strong enough instruments) for all land cover types, and this definition implies that blue oak woodland and the land cover types mostly likely to replace blue oaks are endogenous. This cut off for endogeneity also enables me to test the validity of a larger set of instruments using Wooldridge's robust score test of overidentification (Cameron and Trivedi 2009). Last, though I reject below the null hypothesis that the full set of instruments developed in this paper are valid, I maintain the validity of these

³⁴ The following results for land covers grouped by ecosystems demonstrate that the effects of individual land cover types, including vegetation types, differ across spatial scales. These more general results for vegetation and natural land covers are calculated by replacing the corresponding land cover variables at the within-neighborhood and neighborhood levels with aggregate measures of vegetation and natural land covers. These general results are not reported below.

instruments based on prior arguments and because a rejection of this null hypothesis can result from misspecification, as well as invalid instruments (Cameron and Trivedi 2009).

The key results regard the effects of blue oaks, agriculture, herbaceous, and urban land use on housing prices; all coefficients of land cover at the census block or census block group levels should be interpreted as the change in housing price for a substitution of one percent of the corresponding land cover type for urban land at that particular spatial scale. Blue oaks have a negative and statistically significant effect on housing prices at the census block level as compared to urban land. At the neighborhood (census block group) level, blue oaks have a positive and statistically significant effect. Agriculture has a negative and statistically significant effect on property prices at the census block level and census block group levels. Herbaceous land cover has a statistically insignificant negative effect on property prices at the census block scale, and a statistically significant negative effect at the census block group scale. In addition, other oak woodland has a significant positive effect at the within-neighborhood (census block) scale, and a significant negative effect at the neighborhood (census block group) scale. In terms of overall land cover types, I reject the null hypotheses that all vegetation types (conifers, desert, oak forest, other oak woodland, blue oak woodland, herbaceous, shrubs, and wetland), vegetation and urban land cover types, non-urban land cover types (agriculture, barren, vegetation, and water), and all land cover types (agricultural, barren, vegetation, urban, and water) have equal effects on property prices.

In comparison to the OLS log-log results reported in Table 5, the coefficient estimates corresponding to agricultural, blue oak, other oak, and herbaceous variables increase in magnitude when significant, as do the coefficients for many of the other land cover types. The

land cover coefficients become significant or increase in significance, except for herbaceous vegetation at the census block level, and some change signs.

In general the signs of variables match their predicted signs in Table 3.a. There are four exceptions for land cover variables at the within-neighborhood scale: the sign of the coefficients corresponding to blue oak woodland, shrubs, and water are unexpectedly negative, and the effect of desert is unexpectedly positive. There are two exceptions for land cover variables at the neighborhood scale: other oak woodland land cover has an unexpectedly negative effect and wetland land cover has an unexpectedly positive effect on property prices.³⁵ In terms of the non-land cover variables, the percentage of the neighborhood that is Hispanic and the percentage of the neighborhood that is publically owned have unexpectedly positive and negative signs, respectively. The signs of these two variables are likely the result of collinearity: the percentage of the neighborhood that is Hispanic is highly correlated with the percentage of the neighborhood that has a college degree (-68%) and the neighborhood's median income (-66%); the percentage of a neighborhood that is publicly owned is highly correlated with the proxy variable for conifers at the neighborhood scale (75%). Finally, the quadratic effect of zoning whereby low and high density zoning have negative effects on property prices is expected, but the exact points where the zoning switches effects differ.³⁶

I test for endogeneity (Table 19a) and weak instruments (Table 19b). Using the Durbin-Wu-Hausman (DWH) test, I reject the null hypothesis that the majority privately owned land cover types are exogenous. Examining the first stage regression results to detect weak instruments, the

³⁵ Given that urban land is the omitted land use from my regression and my predicted signs for urban land cover, technically the positive coefficient corresponding to desert at the within-neighborhood scale and the negative coefficient corresponding to other oak woodland at the neighborhood scale are possible. However, I would have deemed them unlikely a priori.

³⁶ I expected the effect of zoning to be quadratic because high density zoning implies future high density, which has a negative effect on housing price, and low density zoning prevents future subdivisions that may be valuable due to diminishing marginal utility to plot area. Both types of zoning affect the price of housing through their effects on future rents.

R^2 and adjusted- R^2 estimates are fairly high. Using the joint F-statistic, I strongly reject the null hypothesis that all of the coefficients in the first stage analyses are equal to zero. While the partial R^2 estimates and the minimum eigenvalue also appear high enough that weak instruments is not a critical problem, Shea's Adjusted Partial R^2 estimates are low enough to engender some caution.

In my a priori preferred specification, specification (1) in Table 19, I implement two-stage least squares with all of the proposed instruments measuring land cover amenities at the census block and census block group levels. Wooldridge's robust score test of over-identified restrictions rejects the null hypothesis that the full set of instruments is valid. However, the rejection of the null hypothesis can result from model misspecification instead of invalid instruments (Cameron and Trivedi 2009).³⁷ Because all functional forms were rejected using the link and Ramsey reset tests, including the log-log form, and omitted variables are present, as demonstrated in a previous sub-section, Wooldridge's robust score test statistic may reflect model misspecification in this case. As a consequence, the full set of instruments cannot be rejected. Instead, sensitivity analysis is conducted with respect to the number of instruments. Results of the sensitivity analysis are reported in a following sub-section.

Sensitivity Analysis

Using heteroskedasticity robust standard errors, I estimate five additional specifications: (2)-(6) in Tables 19 to 22. First, in specification (2) in Table 19.a, I relax the assumption in the a priori preferred specification that majority privately owned land cover types are endogenous to all land cover types that are greater than one-third privately owned.³⁸ While I again reject the

³⁷ This is because Wooldridge's robust score test of over-identified restrictions has power in multiple directions (Cameron and Trivedi 2009).

³⁸ The endogenous land cover types increase from agriculture, other oak woodland, blue oak woodland, herbaceous, and urban land covers to include desert and shrub land covers.

null hypothesis that all land cover types are exogenous, the expansion of the number of endogenous land cover types moves the analysis towards weaker instruments as measured by decreases in the minimum eigenvalue in Table 19.a and the first stage regression results (R^2 , adjusted- R^2 , Shea's adjusted partial R^2 , and the joint F-statistics) in Table 19.b. Second, I reduce the set of instruments due to the potential invalidity of my initial choice of instruments. I drop the available water capacity and the poor drainage variables because it is possible that households may be willing to pay less for a water-logged property.³⁹ I also drop the variables for a slope above a 15% grade because a property with a high slope may have a beautiful view. I then rerun the previous two specifications with the set of remaining instruments; see specifications (3) and (4) in Table 21.a. Using Wooldridge's robust score test of over-identified restrictions, I fail to reject the null hypothesis that the remaining instruments are valid (Table 21a). However, weak instruments are a potential problem as measured by the low minimum eigenvalues in Table 21.a and the first stage regression results in Table 21.b. Therefore, the interpretation of the key variables is most likely to be accurate for the a priori preferred specification, i.e. specification (1) in Table 19, as compared to specifications (2)-(4). Third, specifications (5) and (6) result from re-estimating the a priori preferred specification using the left-hand side linear box transformation and the log-linear functional forms, respectively.

Across all six specifications, the coefficients corresponding to land cover variables are not robust in terms of statistical significance. This is particularly true for specifications (5) and (6) where the statistical significance of land cover variables varies greatly from specification (1) at both the neighborhood and within-neighborhood spatial scales.

Summary

³⁹ I also drop the available water capacity variables because, of the instruments, they are both the most collinear with other instruments and highly correlated with the real price of housing.

The results for the preferred specification, i.e. specification (1) in Table 19, predict that a decrease of blue oaks within the Tulare Lake Basin due to climate change and a corresponding increase in herbaceous land cover will increase the property prices of immediately surrounding properties, and decrease them within the same neighborhood. Similar results hold if urban or agricultural development replaces blue oak woodlands.⁴⁰ However, these results do not always hold under a variety of robustness checks, particularly when I change the functional form. This is because the sign and statistical significance of the coefficients corresponding to land cover variables are highly variable across specifications. Potential explanations for this lack of robustness are multicollinearity or standard errors that are biased downwards due the imposition of census block group level data at the property level. A solution for this latter problem is to calculate cluster robust standard errors.

Cluster robust standard errors at the neighborhood-level

The calculation of cluster robust standard errors adjusts all standard errors upward as compared to the two-stage least squares regressions with heteroskedasticity robust standard errors. As a consequence, all t-tests adjust downwards, and previously significant variables may become insignificant. None of the coefficient estimates change from the previous two-stage least square estimates. Therefore, any discussions of signs in the previous subsection still hold.

When I control for the clustering of standard errors, standard errors substantially increase. Standard errors for land cover variables increase on average by a factor of between four and five. As a consequence, many of the coefficients become statistically insignificant. This may indicate

⁴⁰ In terms of the theoretical model in Chapter III, these results violate the assumption that blue oak woodlands are preferred to herbaceous land cover at all spatial scales. These results also violate the monotonicity assumptions for the net land use externality function because I assume in Chapter III that there is a negative location-dependent land use externality from urban land and a positive location-dependent land use externality from private open-space. Therefore, the Chapter III results do not apply if these empirical estimates are valid. However, many of these empirical estimates are no longer significant when I calculate cluster robust standard errors.

that households do not care about proximity to particular vegetation or land cover types. Joint hypothesis tests at both the neighborhood and within-neighborhood spatial scales find some support for this claim. I find that households do differentiate between vegetation and urban land covers and households differentiate between vegetation and other non-urban land covers, but do not differentiate between vegetation types. This latter result holds regardless of whether land cover variables are constructed to reflect ecosystem types, as in previous sections of this paper, or tree cover density. The overall finding is that there is no cost to nearby residents of blue oaks being replaced by herbaceous land cover as an effect of climate change, and there may be a positive benefit at the neighborhood scale to developing woodlands for urban and agriculture use.

Preferred specification with Cluster Robust Standard Errors

I re-estimate the a priori preferred specification using cluster robust standard errors; see specification (1) in Table 23. Controlling for biased standard errors using clustering of standard errors around neighborhoods, as defined by census block groups, I find that results for the a priori preferred specification demonstrate a general decline in the statistical significance of land cover variables at both the neighborhood (census block group) and within-neighborhood (census block) levels compared to when the specification was estimated using heteroskedasticity robust standard errors. Examining the a priori preferred specification with cluster robust standard errors (specification 1, Table 23), I find that the only neighborhood land cover variable that remains statistically significant at the 1% significance level is desert (as compared to specification 1, Table 19); desert has a negative effect on property prices as expected. The statistical significance of the negative effect of agriculture at the neighborhood-level on property values decreases from the 1% significance to the 5% significance level, while the negative effects of other oaks and

herbaceous vegetation at the neighborhood scale decrease from the 1% significance level to the 10% significant level. None of the remaining neighborhood land cover variables that were statistically significant in the a priori preferred specification with heteroskedasticity robust standard errors (blue oaks, shrubs, and water) are statistically significant with the calculation of cluster robust standard errors. Additionally, none of the land cover variables measured at the census block scale are statistically significant in the preferred specification with cluster robust standard errors. In comparison, eight census block land cover variables were significant at the 1% significance level in the two-stage least squares regression of the preferred specification with heteroskedasticity robust standard errors (specification 1, Table 19).

While some of the land cover variables are still individually significant at the neighborhood-level in the a priori preferred specification with cluster robust standard errors, I find a general loss of joint significance of vegetation variables and land cover variables at both the within-neighborhood and neighborhood spatial scale. I fail to reject the following null hypotheses: all coefficients corresponding to vegetation variables are equal at the neighborhood scale; all non-urban land cover variables at the neighborhood scale are equal to zero and hence are equal to urban land; all vegetation variables at the within neighborhood scale are equal to zero and hence are equal to urban land; and all non-urban land cover variables at the within neighborhood scale are equal to zero and hence are equal to urban land. While I am able to reject the null hypothesis that vegetation variables at the neighborhood scale are jointly equal to zero, i.e. equal in value to urban land cover, at the 5% significance level, I am unable to reject the null hypothesis that all coefficients corresponding to neighborhood vegetation variables are jointly equal to zero when I exclude desert land cover (with a p-value of 0.566) and when I exclude both desert and wetland land covers (with a p-value of 0.4711). I exclude desert and wetland based on the a priori

assumption that these vegetation types negatively affect property values at the neighborhood scale.⁴¹ I fail to reject the null hypotheses that the effect of blue oak woodlands on property prices at both the within-neighborhood and neighborhood scales differs from either herbaceous, agriculture, and urban land covers at even the 10% significance level.⁴²

Focusing on the estimation results for the a priori preferred specification, the take-away message is that households do not differentiate between vegetation types at both spatial scales. While households do differentiate between urban and vegetation land covers at the neighborhood scale, it is only in the sense of preferring not to live within a neighborhood with two undesirable vegetation land cover types, here defined as desert and wetlands. These results indicate that, in general, households do not care about the type of land cover types within the census block and census block group in which they live once one uses clustering to control for spurious correlation. Finally, marginal replacements of blue oak woodlands with herbaceous land cover types due to climate change and marginal replacements of blue oak woodlands with urban and agricultural land covers due to development will have no effect on household welfare.⁴³

The apparent statistical significance of many of the non-land cover variables with more naïve statistical methods no longer holds with the use of cluster robust standard errors. Many of the non-land cover neighborhood variables become insignificant, including the percentage of the neighborhood that is Hispanic, the percentage of the neighborhood that is publicly owned, and the average neighborhood tax rate.⁴⁴ As expected, variables which vary little within

⁴¹ See footnote 47.

⁴² Testing whether the effect of blue oak woodland on property prices is equal to urban land is equivalent to testing whether the individual blue oak woodland coefficient is equal to zero. This is because urban land is the land cover type dropped at both the census block and census block group scales.

⁴³ While marginal replacements of blue oak woodlands by man-made land covers have no effect on household welfare, the accumulation of marginal changes by many households over space and time will result in non-marginal changes. This analysis does not apply in this case, and changes in household welfare may occur.

⁴⁴ 57% of neighborhood non-land cover variables and 50% of neighborhood land cover variables remain statistically significant after the use of cluster robust standard errors.

neighborhoods experience large increases in their standard errors. As a consequence, distance to the nearest central business district, several of the dummy variables for zoning, and the dummy variables for Fresno County lose statistical significance. However, the school quality variable, the annual amount of precipitation, the dummy variable for Tulare County, dummy variables for high density housing and the dummy variable for mobile home zoning remain significant. Also as expected, all of the structural housing variables remain significant at the 1% significance level due to their variability at the property level. Lastly, the intercept is no longer significant.

Sensitivity Analysis

Using cluster robust standard errors, I estimate twelve additional specifications: (2)-(13) in Tables 23 to 27. For each specification, I conduct the same joint hypothesis tests as I did for the a priori preferred specification with cluster robust standard errors. See Table 28 for a summary of all joint hypothesis tests across all thirteen specifications.

Specification (2) in Table 23, specifications (3) and (4) in Table 24, and specifications (5) and (6) in Table 25 are the results from re-estimating specifications (2)-(6) in the previous subsection using cluster robust-standard errors. As in the a priori preferred specification, many of the land cover variables lose individual statistical significance when using cluster robust standard errors instead of heteroskedasticity robust standard errors.

In terms of the joint hypothesis tests, I find general support that households do not differentiate between vegetation types and do differentiate between vegetation and urban land covers. In specifications (2)-(4), I fail to reject the null hypotheses that households do not differentiate between: vegetation land covers, vegetative and urban land covers, non-urban land covers, and all land covers; these results differ from the a priori specification only in the failure of

households to differentiate between vegetation and urban land covers.⁴⁵ In specifications (5)-(6), I reject all four of these null hypotheses. However, I fail to reject the null hypotheses that households do not differentiate between vegetation land covers and between vegetation and urban land covers when I exclude undesirable vegetation: desert and wetlands. Like the a priori specification, the latter two specifications support the argument that households do not differentiate between vegetation types and between vegetation and urban land covers, except to avoid neighborhoods with undesirable land covers. However, the ability to reject the third and fourth null hypotheses, even after excluding desert and wetlands, supports the possibility that households value vegetation and non-vegetation land cover types (agriculture, barren, urban, and water) differently. In terms of blue oak land cover, like specification (1), the effect of blue oak woodlands on property price is equal to the effects of herbaceous, agricultural, and urban land covers in all six specifications.

Many papers in the open space literature that find that the type of open space matters in terms of its effect on property price, use proxies for land cover amenities calculated uniquely for each property and do not utilize cluster robust standard errors.⁴⁶ In order to examine whether these alternative definitions of land cover remove the need to utilize cluster robust standard errors, I re-estimate the a priori preferred specification three times with different land cover variables (Table 26). Specifications (7) and (8) replace each of the within-neighborhood land cover variables in the a priori preferred specification with a dummy variable for whether a property is within 0.1 km or 0.5 km, respectively, of the corresponding land cover type. Specification (9) replaces each

⁴⁵ The results in specifications (3) and (4) should be interpreted with caution given the evidence of weak instruments provided in the previous sub-section about the corresponding specifications with heteroskedasticity robust standard errors. Though weak instruments are not as significant of a concern for specification (2), the change in the results may result from weaker instruments.

⁴⁶ Some papers control for spatial autocorrelation using spatial fixed effects or the Haining method (1993). Spatial fixed effects bias the estimate of overall capitalization. The Haining method may correct for spatial autocorrelation if neighborhoods are defined appropriately.

of the within-neighborhood and neighborhood land cover variables in the preferred specification with dummy variables for whether a property is within 0.5 km and 1.0 km, respectively, of the corresponding land cover type. While the standard errors of the proxy variables for land cover amenities as measured by distance to a land cover type increase with clustering by a smaller factor on average than the census block and census block group variables utilized earlier, the increases are still substantial enough to result in the statistical insignificance of most land cover variables.

I conduct similar joint hypothesis tests as before for the three new specifications, and find again that the results only differ slightly from the a priori preferred specification. I fail to reject the null hypotheses that the effect of vegetation variables on property price are jointly equal and jointly equal to the effect of urban land at both the within-neighborhood and neighborhood scales. In a majority of specifications, specifications (8) and (9), I also find that households differentiate between non-urban and all land cover types. While these results again indicate that households do not differentiate between vegetation types, the results differ from the a priori preferred specification in that households do not differentiate between vegetation and urban land covers and households may differentiate between vegetation land covers and other non-urban land covers (agriculture, barren, and water).⁴⁷ In terms of blue oak land cover, with one exception, I again fail to reject the null hypotheses that blue oak woodland has the same effect on property prices as herbaceous, agricultural, and urban land covers. In specification (9), I reject the equality of blue oak and urban land covers at the within neighborhood scale; however, neither corresponding coefficient is individually significant. These results, along with the results

⁴⁷ Unlike previous results where I found that households avoid neighborhoods within undesirable vegetation types, I find evidence that households may prefer not to live in close proximity to agricultural land and prefer to live in the same neighborhood as herbaceous and barren land.

in the previous subsections, indicate that there is no cost in terms of decreased property prices of replacing blue oak woodlands with herbaceous, agricultural, or urban land covers.

The aggregation of land cover types is the imposition of restrictions that results in biased estimates if false. As a consequence, analysis is conditional on assuming that the restrictions are true. While the evidence is fairly clear that households do not care about particular vegetation types as defined by ecosystems, they may care about vegetation in terms of its density (non-tree greenery, woodland, and forest). To analyze the sensitivity of the previous results to various categories of aggregation, I re-estimate the a priori specification using four unique groupings of vegetation; see specifications (10)-(13) in Table 27. In Specification (10), I replace the vegetation variables (conifers, desert, oak forest, other oak woodland, blue oak woodland, herbaceous, shrubs, and wetland) in the a priori specification with an aggregate measure of vegetation at the census block (*CbwVeg*) and census block group (*PerVeg*) spatial scales.⁴⁸ In specification (11), I group vegetation variables into tree-vegetation (woodland and forest) and non-tree vegetation (herbaceous, shrubs, and wetlands). In specification (12), I split tree vegetation in specification (11) into woodland and forest vegetation. Finally, in specification (13), I separate woodland land cover from specification (12) into hardwood woodland and non-hardwood woodland land covers and non-tree land cover from specification (12) into herbaceous and shrub land covers.⁴⁹ This last specification is the only disaggregation of the four based partially on ecosystem type.⁵⁰

⁴⁸ The continued separation of agriculture and urban land is statistically supported by the consistent significance of agriculture in many previous specifications. I maintain the separation of barren and water land covers because they are non-vegetation land cover types that are clearly distinct in nature; in addition, barren land is strongly significant in several previous specifications.

⁴⁹ I include wetlands in herbaceous land cover following the CWHR system's classification of wetland and herbaceous land covers as common life forms. I also include desert shrubs in shrub land cover.

⁵⁰ In specification (10), I group land cover types in the following way: agriculture, barren, vegetation, urban, and water. In specification (11), I group land cover types in the following way: agriculture, barren, tree vegetation, non-tree vegetation, urban, and water. In specification (12), I group land cover types in the following way: agriculture,

I conduct the same joint hypothesis tests as before for the four new specifications, which clarify previous results. Though untestable in specification (10), I fail to reject the null hypothesis that households do not differentiate between vegetation types in specifications (11)-(13). In specifications (10)-(12), I reject the null hypothesis that households do not differentiate between urban and vegetation land covers.⁵¹ Again in specifications (10)-(12), I reject the null hypothesis that households do not differentiate between non-urban land cover types; this difference is driven by a difference between vegetation land covers and agriculture in specifications (10) and (11).⁵² Finally, in specifications (11)-(13), I fail to reject the null hypothesis that all land cover types have equal effects on property prices. These results provide strong evidence supporting the results under the a prior specification that households do not differentiate between land cover types, but do differentiate between vegetation and urban land covers. Contrary, to the preferred specification, the results also provide strong evidence that households differentiate between vegetation land cover and other non-urban land covers, particularly agricultural land cover.

In these four specifications, I measure the cost of the marginal loss of blue oak woodlands using the aggregate land cover type that contains blue oak woodland. This corresponds to vegetation, tree, woodland, and hardwood woodland land covers in specifications (10), (11), (12), and (13), respectively. Though immeasurable in specification (10), I again find that there is no statistically significant difference between the effect of blue oak woodland and herbaceous

barren, woodland, forest, non-tree vegetation, urban, and water. In specification (13), I group land cover types in the following way: agriculture, barren, hardwood woodland, non-hardwood woodland, forest, herbaceous, shrub, urban, and water.

⁵¹ The failure to reject in specification (13) may result from the incorrect disaggregation of land cover types by ecosystem type, instead of just by density.

⁵² In specification (10), I reject the null hypothesis that the marginal prices of vegetation and agriculture are equal at the census block and census block group scales with p-values of 0.0017 and 0.0573, respectively. In specification (11), I reject the null hypothesis that the marginal prices of tree and agricultural land covers are equal at the census block group scales with p-values of 0.0299.

land cover in specifications (11)-(13).⁵³ Unlike previous results, in specifications (10) and (11), I find that there is a significant difference between the marginal implicit prices of blue oak woodland and agriculture. When statistically significant, there is an economic benefit (cost) from replacing blue oak woodland with agriculture at the neighborhood (within-neighborhood) scale. In specifications (10)-(12), I also find a statistically significant difference between the marginal implicit prices of blue oak woodland and urban land cover at the neighborhood scale. Unlike previous results regarding blue oak woodland, I find that there is a positive externality at the neighborhood scale from developing blue oak woodland for urban or agricultural use.⁵⁴

Though households do not differentiate between vegetation land covers, the last four specifications provide evidence that households consider vegetation density, and not ecosystem type, when making housing decisions. First, as initially expected, I find that households do differentiate between vegetation and non-vegetation (agriculture, barren, water, and urban) land covers when I group vegetation by density. This differentiation only breaks down in the last four specifications when I again group vegetation land cover by ecosystem type in specification (13). Second, collectively, the rejection of the equivalence of forest and non-tree vegetation and the familiar failure to reject the equivalences of woodland and non-tree vegetation and of woodland and forest vegetation in specification (12) support the hypothesis that households account for tree density. Therefore, the results under specifications (10)-(12) should be given greater weight than the other alternative specifications to the a priori preferred specification.

Summary

⁵³ Like blue oak woodland, I measure the cost of replacing blue oak land cover with herbaceous land cover using the aggregate land cover type that contains herbaceous land cover. This corresponds to non-tree land cover in specifications (11) and (12) and herbaceous land cover in specification (13).

⁵⁴ This matches the result in Irwin (2002) for private forests.

When I use cluster robust standard errors, many of the results that held previously no longer hold. First, households do not differentiate between vegetation types, regardless of whether I aggregate by ecosystem or tree density. Second, I find strong evidence, particularly when I disaggregate vegetation by tree density, of households differentiating between vegetation and urban land covers at the neighborhood scale, and not at the within-neighborhood scale. Using the estimates in specification (10) in Table 27, I find that a 1% increase of urban land cover at the expense of a 1% decrease of vegetation land cover at the census block group level increases the housing price of the average (mean) priced house by \$773.57.⁵⁵ Third, in some specifications, particularly those that disaggregate vegetation by ecosystem type, I find evidence that this difference between the marginal implicit prices of vegetation and urban land at the neighborhood scale is driven by a preference to live in neighborhoods without undesirable land cover types, such as desert and wetlands. Lastly, there is some evidence that households differentiate between vegetation land cover and other non-urban land cover at the both spatial scales. According to specification (10) in Table 27, this result is driven by a difference between the marginal implicit prices of vegetation and agriculture at both spatial scales; the difference in the marginal implicit prices of agricultural and vegetation land covers is jointly statistically significant at both spatial scales. I find that a 1% increase of agricultural land cover at the expense of a 1% decrease of vegetation land cover at the census block group level increases the housing price of the average priced house by \$593.67. Additionally, I find that a 1% increase of agricultural land cover at the

⁵⁵ Due to the non-linear functional form of the hedonic price equation, the marginal willingness to pay estimate for blue oaks at the neighborhood scale is the product of the estimated coefficient corresponding to blue oaks at the neighborhood scale and the estimated price of housing. Following the common literature practice, I calculate marginal willingness to pay at the mean price of housing by substituting the mean price of housing for the estimated price of housing. An alternative is to evaluate the hedonic price equation at the mean values of housing characteristics. In both cases, the statistical significance of the marginal willingness to pay estimate for blue oaks is not equivalent to the statistical significance of the coefficient corresponding to blue oaks at the neighborhood scale. Rather, it depends on the statistical significance of all coefficients in the hedonic price equation. Future work will address this issue.

expense of a 1% decrease of vegetation land cover at the census block level decreases the housing price of the average priced house by \$321.06.

The estimates of the cost of a marginal change in blue oak woodlands vary based on the type of vegetation that will replace it. Because I find that households do not differentiate between land cover types, the cost of a marginal shift in blue oak woodlands for herbaceous land cover due to climate change is \$0. Across some specifications, particularly those with vegetation types disaggregated by tree density, I find evidence that households differentiate between tree vegetation and man-made land cover types (agriculture and urban) at the neighborhood-level. Using specification (11) in Table 27, I find that a 1% increase of urban land cover at the expense of a 1% decrease of tree land cover at the census block group level increases the housing price of the average priced house by \$1,538.74. A similar loss of tree land cover for agricultural land cover increases the housing price of the average priced house by \$1,247.30. Using specification (12) in Table 27, I find that a 1% increase of urban land cover at the expense of a 1% decrease of woodland at the census block group level increases the housing price of the average priced house by \$1,272.49. However, according to the specifications with vegetation disaggregated by ecosystem type and specification (13), there is no statistical difference between the marginal implicit prices of blue oak woodland and man-made land cover types (agriculture and urban).

The results of this paper differ from the results in the literature that open spaces, including woodlands and forests, have positive effects on surrounding property prices. There are many factors that may contribute to this difference. First, none of the surveyed papers utilize cluster robust standard errors. As a consequence, the statistical significance of some of the coefficients estimated in these papers may be overstated. If I do not correct for cluster robust standard errors, my results are consistent with the previous literature in that vegetation has a positive effect on

property prices at the within neighborhood scale. Second, this paper focuses on the Tulare Lake Basin, one of the most important agricultural areas in the country. Because household preferences determine inter-regional sorting, in addition to intra-regional housing choice as analyzed in this paper, the preference structure of individuals in this area may differ from the other study areas. Third, while the percentage of a neighborhood that is publically owned is accounted for, this paper does not differentiate explicitly between privately and publically owned land cover types or between urban land cover types by density of housing. Nor does it account for preserved lands.⁵⁶ Fourth, recreational areas within urban areas are likely designated as urban land in the land cover data and variables. As a consequence, the excluded open spaces within urban areas, which increase nearby property prices, bias downward the value of non-urban land cover types. Fifth, the proxy variables for land cover amenities utilized in this paper, the percentage of the census block or census block group covered by land cover types, may fail to capture the full set of amenities. While these proxy variables more accurately reflect land cover amenities from private land covers than distance measurements, they suffer from the possible shortcomings of not varying by property and non-uniformity of size. Future work should potentially use the percentage of area covered by a land cover type within 0.5 km and 1.0 km radii of a property as proxies, following Irwin and Bockstael (2001) and Irwin (2002).

Conclusion

To capture the full value of capitalized land cover services, this paper used two-stage least squares to calculate asymptotically unbiased estimates of the marginal implicit prices of land

⁵⁶ Irwin (2002), which does include such differentiations, finds that conserved land, publically owned land, and privately owned pasturelands produce positive externalities in excess of surrounding development. However, Irwin (2002) also finds the cost of developing privately owned forests to surrounding landowners may be negligible in the case of low density development, and actually a benefit in the case of high density development. This latter result is similar to this chapter, and supports further analysis accounting for ownership and preservation status to find whether the results remain robust.

cover variables. In addition to the opportunity cost variables used as instruments for endogenous open space in previous papers, this paper developed several soil variables to use as instruments for endogenous land cover types. To be able to include multiple proxy variables for each endogenous land cover type, this paper calculated these instrumental variables at various scales of capitalization using different weighting methods to reduce collinearity. While the resulting estimates are inefficient in the presence of spatial autocorrelation, they are asymptotically unbiased and capture the full capitalized values of land cover types. To address spatial autocorrelation within neighborhoods and to adjust standard errors for the imposition of neighborhood-level data at the property level, cluster robust standard errors were calculated. Many of the common econometric results in the urban forestry and open space literatures no longer held after this adjustment.

Blue oak woodlands face two primary threats: development and climate change. While the current literature recognizes the direct welfare effects of climate change through its effect on agricultural productivity and local climates, it has failed to recognize the indirect welfare effects of climate change through its effect on vegetation. By estimating the marginal values of several land cover types using the hedonic model, this paper has demonstrated that the indirect welfare effects of climate change, in terms of the effect of climate change on property prices through its effect on surrounding vegetation, are insignificant. So this omission has no implications in the specific empirical context that I consider. In addition, property owners may actually benefit from the conversion of vegetation, particularly woodlands, to urban and agricultural uses. Therefore, this paper has demonstrated that Tulare Lake Basin households are unlikely to be negatively affected by, and may actually benefit from, marginal losses of blue oak woodlands. These

benefits range from \$0 to \$1,538.74 for the urban development of 1% of a neighborhood's blue oak woodlands, and from \$0 to \$1,247.30 for agricultural development.⁵⁷

One of the key findings in this paper is that property owners do not differentiate between surrounding vegetation land covers or natural (vegetation, barren, water) land covers. This implies that there is no location-dependent cost of climate change in the short run from shifting vegetation or natural land cover types. This does not imply that there is not a cost of shifting land cover types due to climate change. While the marginal implicit prices of vegetation types, location-independent use values (such as the values of ecosystem services), and non-use values (such as bequest, altruist, and existence values) of vegetation types are constant for marginal shifts in land cover types, they may change for non-marginal shifts. This result has several implications. First, research should focus on estimating the non-use values and the values of ecosystem services of the vegetation types most likely to be negatively affected by climate change. This will require the use of stated preference methods, instead of revealed preference methods as used in this paper. Second, like most of the costs of climate change, the bulk of the costs of climate change in terms of its effect on vegetation will occur in the long-run. As for many of the issues surrounding climate change, this raises the problem of how to encourage policymakers to adjust current behavior to avoid or reduce costs in the long-run, which may be substantial. A first step is to use existing estimates of non-use values and the values of ecosystem services to demonstrate the potential magnitude of the indirect costs of climate change, e.g. Chiabai et al (2009).

Third, in terms of previous theoretical papers, this result implies that location-dependent land use externalities from privately owned open space are unaffected by vegetation type. As a

⁵⁷ I assume that statistically insignificant differences between the marginal implicit prices of blue oak and man-made (urban and agricultural) land covers imply no cost or benefit from the loss of blue oak woodland.

consequence, the only uncertainty that local policymakers face is over the uncertain future value of location-independent land use externalities. These location-independent externalities are the non-use values and the location-independent use values of vegetation. Because of the difference in population size, the non-use and location-independent use values that non-residents attribute to localized vegetation types, such as blue oak woodlands, and the species they support are likely to vastly outweigh the non-use and location-independent use values that municipality residents attribute to them. This implies that the optimal adjustments of local policies, as discussed in the previous paper, to account for future learning about climate change are likely to be small. However, if local policymakers account for only their constituents' welfare, they then fail to account for the non-use and location-independent use values that non-residents place on their municipality's land cover, and the corresponding value of information about the future effects of climate change. As a consequence, local policymakers are under preserving private open space as compared to what is socially optimal from the state, national, and international points of view. This implies that land use policies should be set and coordinated at a higher spatial scale than at the municipality level in order to achieve the socially optimal land use allocation and to appropriately account for the value of future learning about the effects of climate change.

Last, the results of this analysis imply that there is little benefit from distorting land use conservation policies to benefit nearby landowners. In other words, the equivalence of location-dependent externalities across vegetation types implies that the socially optimal conservation choice is equivalent to maximizing location-independent land use externalities. Because location-independent land use externalities are made up of non-use values, the optimal choice is the one that maximizes the probability of future existence. One striking implication of the results is that property owners are less likely to pressure policymakers to distort conservation policies in

any way that will be detrimental to the future survival of any one species because they do not differentiate between land cover types.

References

- Abbott, J.K., and H.A. Klaiber. 2010. "Is all space created equal? Uncovering the relationship between competing land uses in subdivisions." *Ecological Economics* 70:296-307.
- . 2011. "An Embarrassment of Riches: Confronting Omitted Variable Bias and Multi-Scale Capitalization in Hedonic Price Models." *Review of Economics and Statistics* 93(4):1331-1342.
- Anderson, S.T., and S.E. West. 2006. "Open space, Residential Property Values, and Spatial Context." *Regional Science and Urban Economics* 36:773-789.
- Bajari, P., and C.L. Benkard. 2005. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy* 113(6):1239-1276.
- Bajari, P., and M.E. Kahn. 2005. "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities." *Journal of Business & Economic Statistics* 23(1): 20-33.
- Belsley, D.A. 1991. "A Guide to Using the Collinearity Diagnostics." *Computer Science in Economics and Management* 4:33-50.
- Bishop, K., and C. Timmins. 2008. "Simple, Consistent Estimation of the Marginal Willingness to Pay Function: Recovering Rosen's Second Stage without Instrumental Variables." Working Paper, Department of Economics, Duke University.
- Brown, J.N., and H.S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." *Econometrica* 50(3):765-768.
- [FRAP, 2005] California Department of Forest and Fire Protection. 2005. "Metadata: Census 2000 Urbanized Areas (CEN00UA02_2)." Available at: <http://frap.cdf.ca.gov/data/frapgisdata/data%20dictionaries%5Ccen00ua.xml> (accessed April, 2011).
- [FRAP, 2002] California Department of Forest and Fire Protection. 2005. "Metadata Record: Multi-source Land Cover Data (v02_2)." Available at: http://frap.cdf.ca.gov/data/frapgisdata/output/fveg02_2.txt (accessed May, 2012).
- Cameron, A.C., and P.K. Trivedi. 2009. *Microeconometrics Using Stata*, 1st ed. College Station TX: Stata Press.
- Chattopadhyay, S. 1999. "Estimating the Demand for Air Quality: New Evidence Based on the Chicago Housing Market." *Land Economics* 75(1):22-38.
- Chiabai, A., C. Travisi, H. Ding, A. Markandya, and P. Nunes. 2009. "Economic valuation of forest ecosystem services: methodology and monetary estimates." Working Paper No. 12.2009, Fondazione Eni Enrico Mattei Nota di Lavoro, Milano.
- Cragg, M., and M. Khan. 1997. "New Estimates of Climate Demand: Evidence from Location Choice." *Journal of Urban Economics* 42:261-284.
- Cropper, M.L., L.B. Deck, and K.E. McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." *Review of Economics and Statistics* 70(4):668-675.
- Ekeland, I., J.J. Heckman, and L. Nesheim. 2004. "Identification and Estimation of Hedonic Models." *The Journal of Political Economy* 112(1):S60-S109.

- Freeman, M.A. 1996. "Hedonic Pricing Methods." In D.W. Bromley, ed. *Handbook of Environmental Economics*. Oxford UK: Blackwell, pp.672-686.
- Garrod, G., and K. Willis. 1992. "The Environmental Economic Impact of Woodland: A Two-Stage Hedonic Price Model of the Amenity Value of Forestry in Britain." *Applied Economics* 24:715-728.
- Geoghegan, J., L.A. Wainger, and N.E. Bockstael. 1997, "Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS." *Ecological Economics* 23(3):251-264.
- Haining, R. 1993. *Spatial Data Analysis in the Social and Environmental Sciences*. New York: Cambridge University Press.
- Hannah, L., G. Midgley, I. Davies, F. Davis, L. Ries, W. Thuiller, J. Thorne, C. Seo, D. Storms, and N. Snider. 2008. "BIOMOVE – Creation of a Complex and Dynamic Model for Assessing the Impacts of Climate Change on California Vegetation." California Energy Commission, PIER Energy-Related Environmental Research Program. CEC-500-2008-060, December.
- Heckman, J.L., R.L. Matzkin, and L. Nesheim. 2003. "Simulation and Estimation of Hedonic Models." IZA Discussion Paper No. 843, Institute for the Study of Labor.
- Heikkila, E. 1988. "Multicollinearity in Regression Models with Multiple Distance Measures." *Journal of Regional Science* 28(3):345-362.
- Hill, C.H., and L.C. Adkins. 2003. "Collinearity." In B. Baltagi, ed. *A companion to theoretical economics*. Oxford: Blackwell, pp.256-278.
- Howard, P.H. 2011. "Optimal Preservation of Environmental Land within Optimal Preservation of Agricultural and Environmental Land within a Municipality under Irreversibility and Uncertainty." Working Paper, Department of Agricultural and Resource Economics, University of California, Davis.
- Howitt, R., J. Medellín-Azuara, and D. MacEwan. 2009. "Estimating the Economics Impacts of Agricultural Yield Related Changes for California." California Climate Change Center. CEC-500-2009-042-F, August.
- Irwin, E.G. 2002. "The Effects of Open Space on Residential Property Values." *Land Economics* 78(4):465-480.
- Irwin, E.G., and N.E. Bockstael. 2001. "The Problem of Identifying Land Use Spillovers: Measuring the Effects of Open Space on Residential Property Values." *American Journal of Agricultural Economics* 83(3):698-704.
- Kennedy, P. 1998. *A Guide to Econometrics*, 4th ed. Cambridge MA: MIT Press.
- Klaiber, A.H., and D.J. Phaneuf. 2009. "Do Sorting and Heterogeneity Matter for Open Space Policy Analysis? An Empirical Comparison of Hedonic and Sorting Models." *American Journal of Agricultural Economics* 91(5):1312-1318.
- Klaiber, A.H., and V.K. Smith. 2009. "Preference Heterogeneity and Non-market Benefits: The Roles of Structural Hedonic and Sorting Models." In J. Bennett, ed. *The International Handbook on Non-market Environmental Valuation*. Cheltenham UK: Edward Elgar, pp. 222-253.

- Kueppers, L.M., M.A. Snyder, L.C. Sloan, E.S. Zavaleta, and B. Fulfroost. 2005. "Modeled regional climate change and California endemic oak ranges." *Proceedings of the National Academy of Sciences of the United States of America* 102(45):16281-16286.
- Kuminoff, N.V. 2009. "Using a Bundled Amenity Model to Estimate the Value of Cropland Open Space and Determine an Optimal Buffer Zone." *Journal of Agricultural and Resource Economics* 34(1):68-90.
- Lenihan, J.M., R.J. Drapek, D. Bachelet, and R.P. Neilson. 2003. "Climate change effects on vegetation distribution, carbon, and fire in California." *Ecological Applications* 13(6):1667-1681.
- Maddison, D., and A. Bigano. 2003. "The amenity value of the Italian climate." *Journal of Environmental Economics and Management* 45:319-332.
- Moulton, B.R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics* 72(2):334-338.
- Palmquist, R.B. 1999. "Hedonic models." In J.C.J.M. van den Bergh, ed. *Handbook of Environmental and Resource Economics*. Cheltenham U.K.: Edward Elgar, pp. 765-776.
- Pascual U., R. Muradian, L. Brander, E. Gómez-Baggethun, B. Martín-López, M. Verma, P. Armsworth, M. Christie, H. Cornelissen, F. Eppink, J. Farley, J. Loomis, L. Pearson, C. Perrings, and S. Polasky. 2010. "The economics of valuing ecosystem services and biodiversity." In P. Kumar, ed. *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations*. London: Earthscan Publications Ltd.
- Powe, N.A., G.D. Garrod, C.F. Brunson, and K.G. Willis. 1997. "Using a Geographic Information System to Estimate an Hedonic Price Model of the Benefits of Woodland Access." *Forestry* 70(2):139-149.
- Price, C. 1995. "The Pros and Cons of Alternative Valuation Methods." In K.G. Willis and J.T. Corkindale, ed. *Environmental Valuation New Perspectives*. Willingford: CAB International, pp.160-177.
- . 2003. "Quantifying the Aesthetic Benefits of Urban Forestry." *Urban Forest and Urban Green* 1:123-133.
- Rehdanz, K. 2006. "Hedonic Pricing of Climate Change Impacts to Households in Great Britain." *Climate Change* 74:413-434.
- Rehdanz, K., and D. Maddison. 2008. "The amenity value of climate to households in Germany." *Oxford Economic Papers*:1-18.
- Ritter, L.V. 1988. "Blue Oak Woodland Vegetation." In K.E. Mayer and W.F. Laudenslayer, Jr., ed. *A Guide to Wildlife Habitats of California*. Sacramento: California Department of Forestry and Fire Prevention.
- Rosen, S. 1974. "Hedonic markets and implicit prices: product differentiation in pure competition." *Journal of Political Economy* 82:34-55.
- Standiford, R.B., and T. Scott. 2001. "Value of Oak Woodlands and Open Space on Private Property Values in Southern California." *Sistemas y recursos forestales* 10(1): 137-152.

Timmins, C. 2003. "If You Can't Take the Heat, Get Out of the Cerrado...Recovering the Equilibrium Amenity Cost of Non-Marginal Climate Change in Brazil." Working Paper, Department of Economics, Yale University.

Tyrväinen, L., and A. Miettinen. 2000. "Property Prices and Urban Forest Amenities." *Journal of Environmental Economics and Management* 39(2):205-223.

[USDA, 1998] United States Department of Agriculture. 1998. "Soil Quality Resource Concerns: Available Water Capacity." Available at: <http://soils.usda.gov/sqi/publications/files/avwater.pdf> (accessed April, 2011).

Walsh, R. 2007. "Endogenous open space amenities in a locational equilibrium." *Journal of Urban Economics* 61(2):319-344.