Universidad de San Andrés

Departamento de Economía

Maestría en Economía

The Determinants of Land Conservation by Type of Easement: A Conditionally Parametric Multinomial Logit Approach

Autora: María Edisa Soppelsa

Legajo: 12632

Victoria, Buenos Aires, Febrero de 2017

The Determinants of Land Conservation by Type of Easement: A Conditionally Parametric Multinomial Logit Approach

Maria Edisa Soppelsa

February, 2017

Abstract

Conservation easements generally have a specific purpose: protect wildlife habitat, provide recreational use for the community, preserve a scenic view. This study focuses on these particular conservation purposes and how location and other variables play a role in that decision. A conditionally parametric multinomial logit model estimates how these variables affect the probability of each use. By allowing coefficients to vary throughout space I find effects not capture by standard or spatial logit models. Results show that effects not only differ by purpose but also spatially. These findings provide useful information for shaping regional policies that are able to address these differences and reach the desired social outcome.

1 Introduction

Land conservation is one of many posible uses for land and is a widely studied topic in land use change literature. Sometimes conservation arises as a mutually exclusive use, as it is the case of a highly developed area versus a protected one. Other scenarios allow conservation and other less invasive uses to coexist, such as conservation easements that allow some agriculture use. The level of interaction between posible uses will depend on public or private interests, other factors and land characteristics that can make certain areas more suitable for a specific use or combination of uses.

Understanding which factors drive land conservation and how they interact is crucial. For example, soil quality and high crop net returns can encourage intense agriculture use. Areas with high biodiversity and environmental value but close to the urban fringe may face a bigger threat for development. In some cases, specific factors like proximity to cities can have a favorable effect on areas used for recreation but a negative impact if the main goal is preservation of certain species. It is important then to also examine how these factors affecting land conservation have different effects when considering specific conservation purposes.

Protecting one particular area implies at the same time the decision of which particular purpose will be the priority concern. Conservation hot-spots and well-mapped species regions are clearly sought for their biodiversity richness and environmental benefits (Armsworth et al, 2004). Areas closer to cities deliver certain recreational value that can be observed in a declining willingness to pay when the reserve is further away from population centers (Ruliffson et al., 2003; Önal and Yanprechaset, 2007; Ando and Shah, 2010). Each different purpose for conservation may be affected by a variety of factors that differ when considering a particular purpose. Yet, not much have been done to identify these factors and their influence in reserve location with different objectives.

This paper aims to analyze the different effects that certain variables to be known to drive conservation decision have on specific conservation purposes. A simple theoretical model considers the problem of an agent that protects a particular area and chooses the conservation purpose which maximizes the utility function. Variables affecting the utility function are basically related to environmental value, demand-side factors and distance to points of interest. This theoretical framework translates into a discrete choice empirical model that estimates the probability of a specific purpose given a set of explanatory variables.

I use a conditionally parametric multinomial logit model to estimate the probability of different conservation purposes. This approach allows for spatial variation without imposing a functional form specification. A standard multinomial logit although useful fails to show how the effect of a particular variable is different throughout space. This distinction is useful as it can help developing policies specifically oriented to achieve the desired social outcome.

Results show how location of a protected area changes the effect of some of the factors affecting conservation purpose. Farm, ranch and forest uses attract bigger areas, further away from cities in the midwest, east and south, but closer to cities in the north west (Washington State). Easements used for environmental system are smaller and further away from highways in the east and south. Proximity to cities has a greater effect on Education and Recreational uses, with a positive effect on northern states (except New England) and decreasing the probability of this use for areas closer to cities in Virginia and Kentucky.

2 Conceptual Framework

In this section I present a simple theoretical framework to analyze what affects conservation purpose decision. I assume that agents making the decision maximize a utility function. The variables that affect that utility can be summarized in three groups: biodiversity, distance variables, and demand-side factors.

Land use change literature has studied different uses of land, how agents decide between them and the factors that affect that decision. From classic theory that explains how rent varies with location (Van Thünen, 1966; Muth, 1971) to spatial econometric models (Carrion-Flores and Irwin, 2004; Brady and Irwin, 2011; Wrenn and Sam, 2014; Chakir and Parent, 2009), this literature has concentrated on a variety of factors affecting land use. Costs and benefits for landowners, property taxes, location of parcels and distance to urbanized areas are just some of the most common variables driving land use change.

Land conservation arise as a particularly interesting land use because more often than not the decision for conservation is permanent. Restricting the possibility of land use change for protected parcels certainly emphasizes the importance of maximizing benefits or minimizing costs at the time of reserve-site selection.

Conservation planning literature has extensively studied how to maximize biodiversity protection in an adapting environment (Pressey, 2007). It has focused on the importance of identifying which areas to protect by also considering economical and social restrictions. It considers not only where to locate protected areas from an ecological point of view, but also the important role that budget constraints, land costs and opportunity costs play on conservation location (Adams et al. 2009).

Following that line, many authors have incorporated cost-benefit analysis to develop models that efficiently locate protected areas. Ando et al. (1998) show how accounting for land prices increases efficiency in either allocating a fixed budget or the coverage of biodiversity protected. Naidoo et al. (2006) focus on which costs to consider, how to quantify them and why they are important. Newburn et al. (2005) also includes the likelihood of land-use change as a variable that affects conservation planning. Development restrictions and threaten of development have also become hot topics in this literature

(Costello and Polasky, 2004; Lovell and Sunding, 2001; Wu and Irwin, 2008).

Besides biodiversity and cost-benefit analysis, other factors should also be considered when studying location of protected areas. From the demand-side point of view, willingness to pay and distance to urban areas influence the location of conservation. Ruliffson et al. (2003) develops a model that accounts for public access and species protection when allocating a limited budget on conservation. Önal and Yanprechaset (2007) incorporate population on cities near the reserve as an important factor to measure public access. Ando and Shah (2010) start the analysis on how willingness to pay for protected areas and location of people reflects on the location of single and multiple reserves. Distance to urban areas and population in those areas arise then as a factor to consider when deciding where to locate conservation.

Proximity to other protected areas also affects the location decision. Agglomeration bonuses pose some clear advantages in terms of biodiversity but they are not always the most efficient decision. Sometimes fragmentation appears as a valuable option as well. Theoretical and empirical models have contributed to this issue. Albers et al. (2008a) use a spatially-explicit game structure to explore conservation patterns accounting for contiguity and land trusts location. Albers et al. (2008b) use a linear spatial econometric model to study how public conservation can attract or repel private conservation. Considering spatial correlation between areas is definitely an important factor when studying landscape layouts.

Apart from affecting conservation use, these factors also influence the particular use given to a protected area. Conservation easements can have a variety of qualified purposes to count as a charitable donation and benefit from tax deductions. According to the Internal Revenue Code, those purposes are: "...public outdoor recreation or education; protection of a relatively natural habitat of fish, wildlife, plants, or similar ecosystem; the preservation of open space including farmland and forest land either for public scenic enjoyment or pursuant to governmental conservation policy; preservation of an historically important land area or certified historic structure" (Rissman et al. 2007; Treasury Regulation §1.170A-14(d)(1)-(5) 2001). Considering the main purpose for protected areas is useful to design better incentives and policies.

It is interesting to see how these factors have different effects if one looks at a specific easements' purpose. One may think that easements used primarily for recreation will have easy access to cities and highways. Biodiversity may need less human disruption, although some species can still be found near urban areas. This may result in landscape patterns that cluster certain conservation purposes in a particular area. Furthermore, policy makers can use this information to better target incentives destined to achieve specific conservation goals. This paper aims to explain how factors that affect the decision of conservation can also affect differently each of the conservation purposes.

Studies about specific conservation purposes are not that common. Chan et al. (2006) explore purpose to some extent by concentrating on trade-offs between conservation goals for biodiversity and the provision of ecosystem services. Rissman et al. (2007) focused more on specific purposes and the level of development allowed. They conducted a survey on easements managed by The Nature Conservancy and find out that 80% of the easements protect core habitat for species, 85% allow some residential or commercial use, and 46% is used as ranch, farm or forestry. However, none of these studies investigate which factors affect the decision of conservation purpose or how different the effect could be.

For this study, I use a simple model where an agent considers the afore mentioned variables and chooses a conservation purpose. An agent could be a Land Trust that works simultaneously with landowners when deciding which areas to protect and how to implement that protection. To simplify which variables to consider, I grouped them in three categories: biodiversity, demand-side factors affecting conservation and distance to points of interest. These categories are not mutually exclusive but can help to have a big picture of the three main factors that influence conservation purposes.

The agent chooses among k different conservation purposes, each one giving certain social utility value. I can assume the utility function depends on the three groups of variables, then the agent maximizes:

$$U = U_k(b, m, d) \quad \forall k = 1, \dots, K \tag{1}$$

where b are variables measuring biodiversity, m are demand-side variables, and d represents distance to points of interest, such as urban areas. The agent chooses k such that:

$$U_k \geq U_j, \quad \forall j \neq k$$

3 Empirical Model

Based on the conceptual framework, I define specific variables that can be used in the estimation. I use a conditionally parametric multinomial logit model to estimate the probability of each conservation purpose. A tri-cube kernel function determines the weights that depend on distance.

I propose a social utility for each easement as a function of the observable variables already mentioned and some other non-observable variables. The utility of choosing a particular conservation purpose is:

$$U_{ik} = x'_{b,ik}\beta_{b,k} + x'_{m,ik}\beta_{m,k} + x'_{d,ik}\beta_{d,k} + \epsilon_{ik} \quad \forall i = 1, \dots, N \quad \forall k = 1, \dots, K \tag{2}$$

where $x'_{b,ik}$ is a vector of biodiversity variables, $x'_{m,ik}$ a vector of demand-side variables, and $x'_{d,ik}$ a vector of distance variables, $\beta_{b,k}$, $\beta_{m,k}$, and $\beta_{d,k}$ are parameters to be estimated and ϵ_{ik} represents unobservable variables.

I use a multinomial logit to calculate the probability of an easement having a specific purpose k, given the observable and unobservable characteristics. The agent chooses the purpose that maximizes the social utility function by comparing utilities for each of the different purposes. If $y_i = 1, \ldots, K$ is the conservation purpose chosen, then:

$$y_i = j \quad \text{if} \quad U_{ij} \geq max\, U_{ik} \quad \forall j, k = 1, \ldots, K+1$$

I use purpose $K+1$ as a reference alternative such that the agent considers utilities of each purpose relative to the $K+1$ purpose:

$$y^*_{ik} = U_{ik} - U_{i,k+1}$$

Because proximity between easements may also play a role in the selection of a conservation purpose, including spatial interaction is useful. Spatial parametric models use a weight matrix with some measure of 'closeness' to account for space. This measure is generally based on contiguity, which may be difficult to justify when using parcel level data. They also require to invert a $n \times n$ weight matrix that could be computational challenging for large datasets. Some applications of spatial multinomial models to land use are Chakir and Parent (2009), Carrion-Flores et. al (2009), Li et. al (2013), and Wang and Kochelman (2009).

Locally weighted regression introduced by Cleveland and Devlin (1988) is a successful nonparametric procedure to account for space without imposing the structure of a functional form. Since variance tend to be high in nonparametric models, a conditionally parametric approach is useful. This assumes that a functional form is convenient to explain the data in small geographic areas by weighting observations within a certain specified distance. It also avoids the problem of defining and inverting a large $n \times n$ weight matrix while estimating at the same time a set of coefficients at particular target points. These coefficients then vary throughout space, providing a better insight of the spatial effect of the explanatory variables.

Based on Tibshirani and Hastie (1987) work, McMillen and McDonald (1999) extended the conditionally parametric method to discrete choice models. They estimate a conditionally parametric multinomial logit model of land use in Chicago after its zoning ordinance in 1920. Other applications using conditionally parametric methods that also focus on land use include McMillen and McDonald (2004), Wang et al. (2011), Wrenn and Sam (2014), and McMillen and Soppelsa (2015).

As is the case with all discrete choice models, I only observe y_i which show the particular purpose k that gives the maximum social utility, then:

$$y_i = \begin{cases} j & \text{if } y^*_{ij} = max\{y^*_{ik}, 0\} \\ 0 & \text{otherwise} \end{cases}$$

I estimate a conditionally parametric multinomial logit where the parameters to be estimated depend on geographic location. Formally:

$$y_i = x_i' \beta(lo_i, la_i) + \epsilon_i \tag{3}$$

where y_i is the dependent variable, x_i is a set of explanatory variables previously grouped in three groups: biodiversity, demand-side and distance, and ϵ_i is the error term. The coefficients β depend on the geographic coordinates, longitude and latitude, for observation i, and vary smoothly through space.

Kernel weights that depend on distance give more weight to nearby observations when estimating β. A tri-cube kernel is the most commonly used function[1]:

$$\omega_{ij} = \left[1 - \left(\frac{\delta_{ij}}{d_i}\right)^3\right]^3 I(\delta_{ij} < d_i)$$

where ω_{ij} is the kernel weight between observations i and j, δ_{ij} is the Euclidean distance between those two observations, d_i is the distance of the qth nearest observation (q is the window size), and I is an indicator function.

Following McMillen and McDonald (1999), locally weighted estimates for a multinomial logit model with $K+1$ alternatives are obtained by maximizing the pseudo log-likelihood function:

$$ln\, L_i = \sum_{j=1}^{n} \omega_{ij} \big(I_{i0}\, log(P_{0i}) + \cdots + I_{iK}\, log(P_{iK})\big) \tag{4}$$

where I_{ik} is a dummy variable indicating that purpose k was chosen for observation i, and P_{ik} is the probability of choosing purpose k for observation i. Normalizing on the base alternative such that $\beta_0 = 0$, the probabilities are:

$$P_{ik} = \frac{\exp(\beta_k' x_i)}{1 + \sum_{m=1}^{K} \exp(\beta_m' x_i)} \tag{5}$$

4 Data

I use a parcel-level conservation dataset, in combination with county-level census variables. I use census shapefiles to calculate distance variables. I also include a species occurrence dataset at the county level.

The basic dataset is the National Conservation Easement Dataset (NCED), an initiative of the U.S. Endowment for Forestry and Communities. This dataset compiles information on conservation easements throughout the United States. Last updated version (July 2016) contains $108,810$ easements managed by Land Trusts and public agencies, covering $24,557,278$ acres protected.

[1]Furthermore, I use a tri-cube kernel function for two main reasons. First, the tri-cube kernel function has a compact support that is necessary when using the nearest neighbor window size radial function on this estimation. Secondly, it was the weight function used when the Locally Weighted Regression Model was first introduced (Cleveland, 1988).

NCED is a parcel/area level dataset where the unit of observation is a conservation easement. Some of the attributes included in this dataset are characteristics about the holder or manager of the easement, intervention allowed, level of public access, duration and size of an easement. The main attribute for this study is the conservation purpose, with 8 categories following the Internal Revenue Code: environmental system, recreation and education, historic preservation, open space for farms, open space for ranch, open space for forest, open space for other uses and an unknown purpose category.

Conservation purpose of a particular easement is the dependent variable in the model. From the previous purpose categories, I structure 5 different possible uses: (1) Environmental System, (2) Education and Recreation, (3) Open space for farm or ranch, (4) Open space for forest, and (5) Open space for other uses. I drop observations which use is unknown. I also drop those who have historic preservation purposes as the decision for this use is probably determined by factors other than the ones studied here.

I combine NCED dataset with county-level census variables (2010 population census and 2012 agriculture census), and species occurrence data from the Biodiversity Information Serving Our Nation dataset (BISON, 2015). I overlay NCED dataset with county shapefiles (US Census TIGER, 2015) to assign a county *geoid* to each easement. I find 118 observations without a matched county. The most common reason for this is some county boundary issues (many observations were located on or near water). To overcome this issue, specific counties should be assigned by hand to each of the 118 easements. However, the decision is to drop these easements because the lack of a matching county only occurs for a small number of observations with a total area of approximately 1,200 acres (a small percentage in the total of acres protected). I restrict the analysis to continental US.

The final dataset consists of $85,561$ easements grouped in 5 different conservation purposes. The base alternative for the multinomial logit model is "Open Space for Other Uses". Explanatory variables from census and biodiversity datasets can be grouped in three categories: biodiversity, demand-side factors, and distance variables.

Biodiversity is difficult to measure, especially when the area considered is the whole continental United States. As a proxy for species richness I use the number of species observed in a particular county, adjusted by the size of the county. This ranges between approximately 0 to almost 20 different species per acre.

Demand-side factors are generally measured by WTP, which is commonly find by surveying people in the area. Once again, this is difficult to estimate when the study covers almost the whole country. One way to have a proxy for this missing variable is to consider income and some distance variable that can estimate how accesible the site is. This has of course many limitations, for example, it assumes that people with higher income will be willing to pay more for a conservation site, which could not always be the case. Nevertheless, it is still true that people on the lower end of the income distribution will probably have a lower willingness to pay. I use the log of median household income, distance to the nearest city with at least 50,000 people, and population density.

Distance variables are also important to determine conservation purpose. On the one hand, one may expect that places near a city are mostly used for recreation purposes. On the other hand, species preservation may be more common in more remote areas, away from cities and highways. I calculate distance to nearest city, nearest highway, rail lines, and to the coast to incorporate distance factors. Distance to the coast includes not only maritime coast but also the Great Lakes.

I present summary statistics in Table **??**. A little bit more than half of the easements are used for environmental system. Easements used for farm and ranch are approximately 18,000 (21%), and forest is just 1.3%. The base alternative of Open Space - Other (use different than farm, ranch or forest) include almost 7,400 easements. Distance to the city ranges from less than a mile to approximately 79 miles, and distance to the highway can reach up to 185 miles.

Table 1: Summary Statistics

Variables				
Dependent Variable: Conservation Purpose		*N*		
Open Space - Other		7,376		
Environmental System		56,983		
Rec. and Education		1,732		
Open Space - Farm and Ranch		18,315		
Open Space - Forest		1,155		
Right-hand Side Variables	*Mean*	*St. Dev.*	*Min*	*Max*
Distance to City	11.144	7.547	0.046	79.115
Distance to Highway	16.888	19.456	0.004	187.465
Distance to Rail Line	4.461	5.055	0.002	90.311
Distance to Coast	126.916	154.190	0.003	846.055
Species per Acre	0.577	0.780	0.0003	19.523
Income (log)	10.909	0.316	9.845	11.638
Pop. Density	436.847	726.708	0.300	69,467.500
Easement Size (log acres)	3.092	2.217	−9.546	12.788
Ag. Value per Acre	1,285.484	807.283	1	2,474

Note: N = 85,561

5 Results

Multinomial logit estimation shows how the effects of the explanatory variables affect different conservation purpose (Table **??**). The probability of all conservation uses increases with the number of species per acre and with higher land values. Coefficients for distance to nearest city and rail line show that the closer the easement is, the higher the probability of that easement being used for all purposes other than open space. These results are somehow intuitive, although one may expect some variation for these coefficients. For example, distance to city having a positive effect for recreation and education purpose.

What is more interesting is to focus on those variables where the effect is different depending on the particular purpose. Higher population density increases the probability of environmental system and recreational uses, but less dense areas have a higher probability for conservation where farms or ranches are present. This is reasonable if we think about people living closer to recreational areas and farms located in rural land. Greater distance to the highway translates in higher probability for all purposes except recreational and educational use. I measure the distance to the nearest highway point, not necessarily a highway exit, which could probably have a positive effect for this particular use.

Although different effects from a particular variable are interesting, they do not tell the whole story on how these variables affect conservation purposes. A standard multinomial logit model fails to incorporate spatial effects that can also influence the estimated probabilities. A spatial multinomial logit will include these spatial interactions. However, it will still assume that one coefficient will be enough to explain the effect of a particular variable. Studying spatial interaction in large geographic areas could show that the effect of a particular variable presents different patterns throughout space.

5.1 A Distribution of Coefficients

A second step is to estimate a conditionally parametric multinomial logit model (CPAR ML). This approach estimates a matrix of $n \times k$ coefficients for each purpose. Spatial interaction is then incorporated by allowing $\beta(lo_i, la_i)$ to vary throughout space. This highlights how each variable affect the probability of a particular conservation purpose differently depending on where the easement is located. This method proves to be useful for large datasets and even if the functional form is not that well specified. For this particular estimation, I use a window size of 25% which indicates that only the closest 25% observations are given positive weights. These weights are calculated using a tri-cube kernel function.

Coefficients means and standard deviations are presented as a summary to do a quick comparison with the standard multinomial logit (Table **??**). Signs of coefficients differ slightly compared to the standard multinomial logit. For the CPAR ML model, the probability for environmental purpose increases for easements located closer to highways or in less populated counties. Distance to highway also increases the probability of farm, ranch and forest. Distance to rail lines increases the probability of an easement being used for education and recreation, while dis-

Table 2: Standard Multinomial Logit Model

	Dependent variable: Conservation Purpose			
	Env. System	Educ. and Rec.	Farm and Ranch	Forest
Constant	16.733***	3.200*	2.900***	4.093**
	(0.712)	(1.636)	(0.827)	(1.707)
Distance to City	−0.019***	−0.023***	−0.043***	−0.018***
	(0.002)	(0.005)	(0.003)	(0.005)
Distance to Highway	0.009***	−0.005**	0.012***	0.004**
	(0.001)	(0.002)	(0.001)	(0.002)
Distance to Rail Line	−0.027***	−0.032***	−0.061***	−0.029***
	(0.003)	(0.007)	(0.004)	(0.007)
Distance to Coast	−0.0002	−0.001***	0.006***	−0.002***
	(0.0002)	(0.0004)	(0.0003)	(0.0004)
Species per Acre	1.281***	0.845***	1.394***	0.022
	(0.045)	(0.083)	(0.050)	(0.152)
Income (log)	−1.125***	−0.227	−0.225***	−0.485***
	(0.065)	(0.150)	(0.075)	(0.158)
Pop. Density	0.0001*	0.0003***	−0.001***	−0.0001
	(0.00005)	(0.0001)	(0.0001)	(0.0001)
Easement Size (log Acres)	−0.301***	−0.380***	0.009	0.144***
	(0.008)	(0.014)	(0.009)	(0.020)
Ag. Value per Acre	0.0001***	0.0003***	0.0005***	0.0005***
	(0.00002)	(0.00004)	(0.00002)	(0.00005)
Observations	85,561			
R^2	0.316			
Log Likelihood	−55,578.740			
LR Test	51,243.910***			
	(df = 72)			

Notes:
*p<0.1; **p<0.05; ***p<0.01
Base Alternative: Conservation for open space with a use different than farm, ranch or forest (*Open Space - Other*).
Fixed Effects by Census Regions

tance to the coast decreases the probability of environmental system and recreation.

Table 3: CPAR Multinomial Logit Model

	Dependent variable: Conservation Purpose			
	Env. System	Educ. and Rec.	Farm and Ranch	Forest
Constant	25.031	7.728	12.286	13.948
	(4.558)	(7.936)	(4.890)	(8.084)
Distance to City	-0.026	-0.023	-0.073	-0.034
	(0.010)	(0.019)	(0.012)	(0.020)
Distance to Highway	-0.00001	-0.048	-0.013	-0.019
	(0.005)	(0.016)	(0.006)	(0.011)
Distance to Rail Line	-0.049	0.011	-0.069	-0.021
	(0.019)	(0.041)	(0.022)	(0.033)
Distance to Coast	-0.007	-0.004	0.005	0.004
	(0.002)	(0.003)	(0.002)	(0.003)
Species per Acre	0.596	0.256	1.063	0.299
	(0.288)	(0.563)	(0.344)	(0.571)
Income (log)	-1.986	-0.696	-1.191	-1.517
	(0.414)	(0.722)	(0.444)	(0.748)
Pop. Density	-0.0001	0.0003	-0.0002	-0.001
	(0.0002)	(0.0003)	(0.0002)	(0.001)
Easement Size (log acres)	-0.304	-0.274	0.360	0.351
	(0.033)	(0.064)	(0.039)	(0.082)
Ag. Value per Acre	0.0002	0.00001	0.0003	0.0003
	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Observations	85,561			
Log Likelihood	−41,776.78			

Note:
Base Alternative: Conservation for open space with a use different than farm, ranch or forest (*Open Space - Other*).

However, one of the advantages of CPAR models is showing a range of effects for the right-hand side variables. The densities of coefficients for each particular use show all coefficients taking values that range from negative to positive (Figure **??**). On the one hand, some variables present higher frequencies for all uses in either positive or negative values. This indicates that the effect of that particular variable is mainly negative or positive for all uses. On the other hand, higher frequencies could be positive or negative depending on the particular use. This is a more interesting scenario because policies oriented to increase conservation should account for this variation when considering the conservation purpose they are trying to achieve.

In general, distance variables show negative effects for all uses. Distance to the coast in particular shows highest frequencies around 0, for most of the uses, although the distribution for environmental system has higher variance with a group of coefficients affecting negatively and another group with a positive effect. Distance to highway is particularly interesting because it affects negatively the probability of recreation and education, but the effect is positive for farm, ranch and forest.

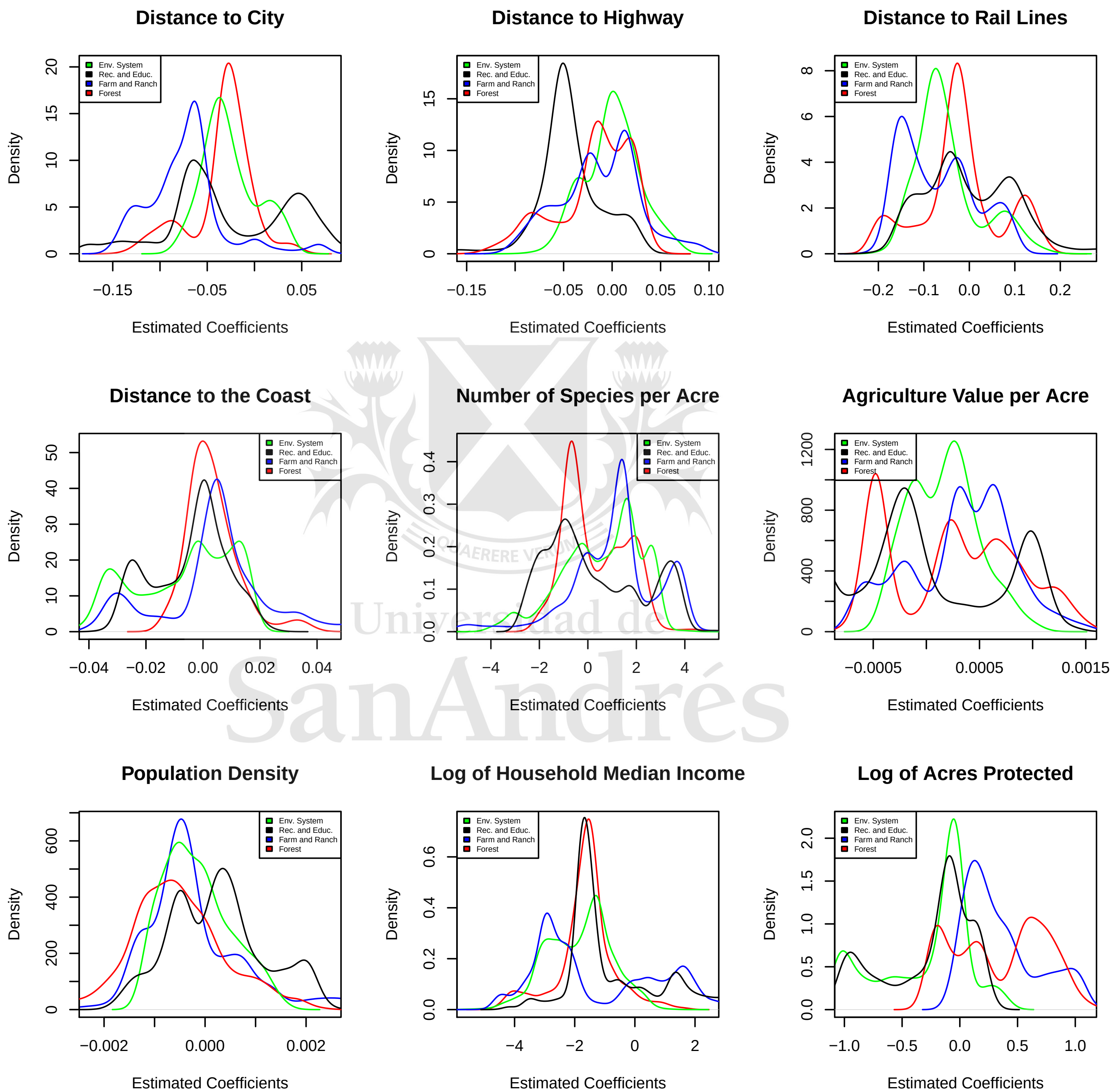
Other variables also have either positive or negative effects depending on the use. The number of species have mainly positive effects on environmental system and farm and ranch uses, but negative effects on the probability of forest use. Agriculture value per acre has mostly a negative effect on recreation and education, and forest. However, when considering the probability of environmental system use, the effect can be either positive or negative.

Maps provide a better picture to analyze how these effects differ widely among observations and throughout space. One of the key advantages of using this methodology is in fact observing smooth changes and how location affect coefficients' variability. I present three sets of maps, each one showing spatial variation on the coefficients for a particular variable, for each conservation purpose (Figures **??** to **??**). I concentrate on the three variables with most evident varying effects: distance to city, distance to highway and easement size.

Distance to the nearest city affect each conservation purpose particularly differently. Increasing distance to city slightly increase the probability for environmental system use in the West of the country (except in California), but the effect is negative for the North East and some parts of the Midwest. Effects are stronger for recreation and education use, where there is a clear distinction between northern and southern areas. Getting farther away from the city seems to increase probability of recreation and education sites in the north, except around New England area, and the opposite happens in southern states, with a higher effect around Virginia and Kentucky. The effect on farm and ranch is somehow similar to environmental system, with higher negative coefficients on the Mid Atlantic area. Finally, distance to city does not seem to have a big effect on forest.

Distance to highway presents a different story. Effects are almost zero for the West and North West, with the exception of recreation and education use that

Figure 1: CPAR Logit Estimated Coefficients



shows a slight negative effect of distance to highway on the northern area. This is probably because highways are highly concentrated on the eastern region of the country, which makes their effect stronger. Increasing distance to highway also increase the probability of environmental system and farm and ranch use for the southern eastern region. It is particularly interesting to notice that this effect turns highly negative for recreation and education around Virginia and Kentucky, but is highly positive for farm and ranch in the same area. The probability of forest does not seem to be greatly affected by distance to highway, although effects are a bit noticeable around New York and Massachusetts states.

The spatially effect that the size of the easement has on the probability of a particular purpose looks interesting. Contrary to what one would expect, bigger areas have a lower probability of environmental system use in the eastern region, especially in the Mid Atlantic. The effect is similar for recreation and education, although it turns positive for southern states. Bigger areas have a higher probability of being used for farm, ranch or forest, particularly in the North and West, but the probability of forest is lower for bigger areas around Kentucky and Virginia.

5.2 Spatial Effects

Allowing the estimation of a distribution of coefficients also proves interesting to detect some patterns across the country. Effects vary throughout space, showing smooth changes between positive and negative effects for all variables and uses. This is interesting because it shows a range of effects that are not captured when estimating standard or spatial logit models.

Two ways of presenting these effects are useful. I first concentrate on regional effects and show a big picture of how each variable has a different effect depending on the region. Then I change focus to particular purposes, explaining how the same purpose is influenced differently depending on where the easement is located. Extra maps are available in the Appendix.

5.2.1 Spatial Effects by Regions

In the West, variables such as distance to city, population density, easement size and agriculture value have a positive effect increasing the probability for almost all purposes. However, recreation and education are negatively affected by easement size and agriculture value in that area. Other variables have a negative effect, although median income increases the probability of recreation and education. For distance to highway and rail lines the effect is still negative but close to zero.

The area around the Great Lakes and northern Midwest shows some differences. Effects are not easy to generalize since area limits are not that well defined and effects vary depending on the purpose. Distance to city and population density have a negative effect for almost all purposes (except recreation and environmental system, respectively). Income's effect is mainly negative, but for farm, ranch and recreation it changes from negative to positive when moving west. Distance to the coast and highway have a negligible effect and the rest of the variables have a positive effect.

The South of the country presents some clear regional effects for some of the variables while other effects appear to be similar to the west region. Distance to highway and population density have a distinct effect, increasing probabilities for all purposes. Agriculture value also have a positive effect for farm, ranch and forest. Size of the easement and distance to the rail line effects are similar to the west, positive and negative, respectively. Other variables such as distance to the coast and number of species seems to have no significant effect in this area.

One area that arises as particularly different is the northern region of the south east (Virginia, Kentucky, West Virginia, North Carolina and Tennessee). For almost all variables, effects are more intense of even sometimes opposite in this area compared to neighboring regions. Distance to the city and easements' size have a negative effect, with some variation depending on the purpose, reaching high values for recreation and education. Distance to highway also have a high negative effect for this purpose, but a high positive effect for farm and ranch purposes. Once more, this makes clear the importance of analyzing different purposes, as not all variables have the same effect on conservation as a whole.

5.2.2 Spatial Effects on Different Conservation Purposes

Recreation and Education appears as a distinct purpose, with many variables affecting it differently than they affect other purposes in the same region. Recreation areas are further from cities and closer to highways in the north and great lakes. In this last area, probability of recreation increases further from rail lines and in counties with higher species occurrence. Effects are reversed in the south, although in the northern part of that region recreation areas are still closer to highways. As mentioned before, this particular region is interesting since variables seem to have different or greater effects than they have in surrounding areas. Easements in a county with higher agriculture value and higher median income have a higher probability of recreational use than being use just as open space. Smaller easements also have more probability of being used for recreation in that area. This last effect is also true for the northwest, but not in the Great Lakes region.

Figure 2: Spatial Variation in the Estimated Coefficients for Distance to the Nearest City

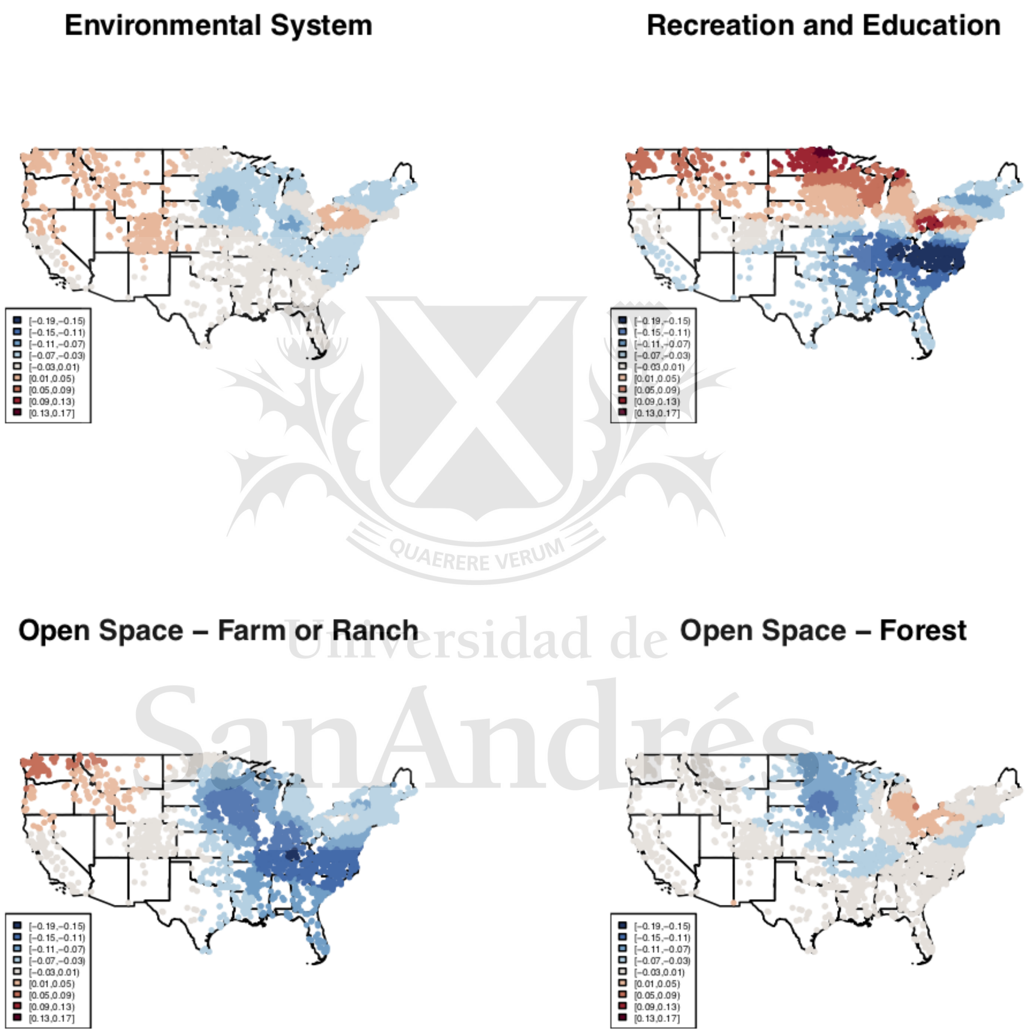


Figure 3: Spatial Variation in the Estimated Coefficients for Distance to the Highway

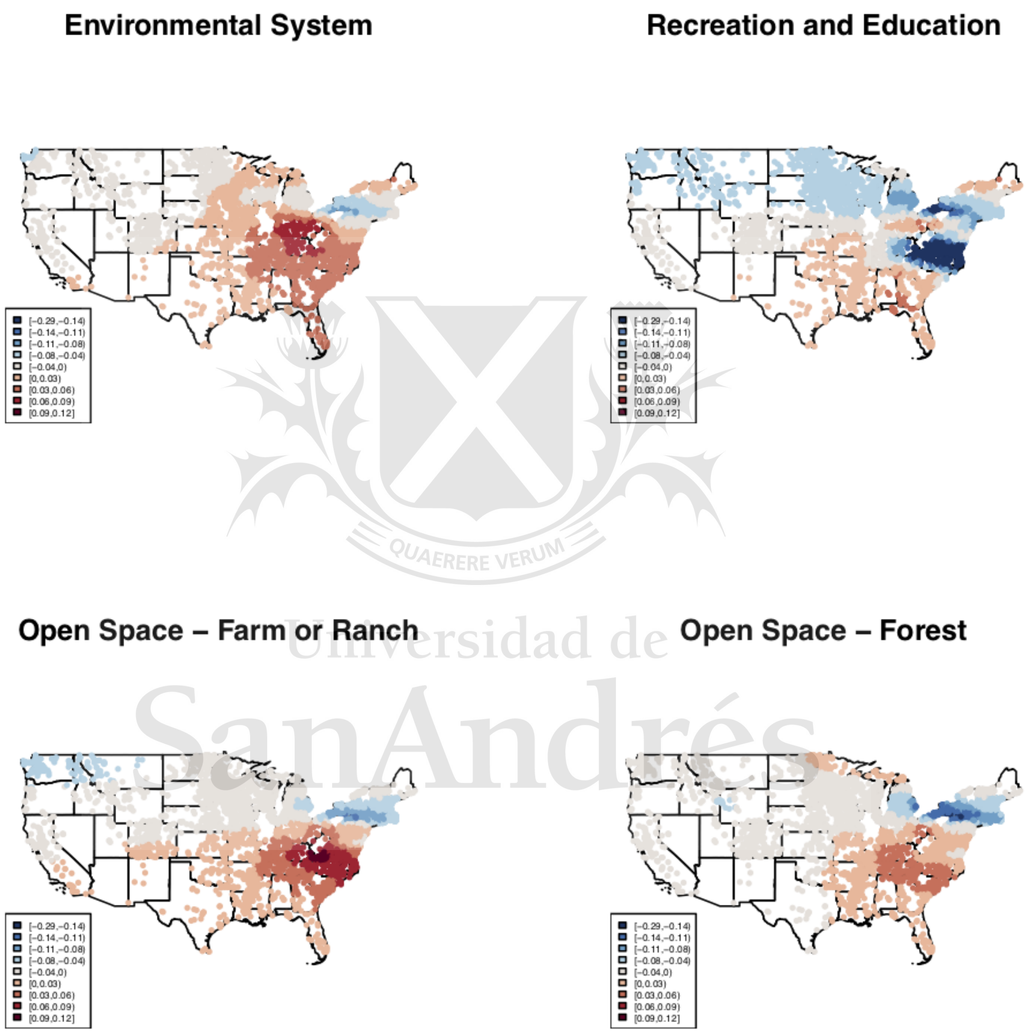
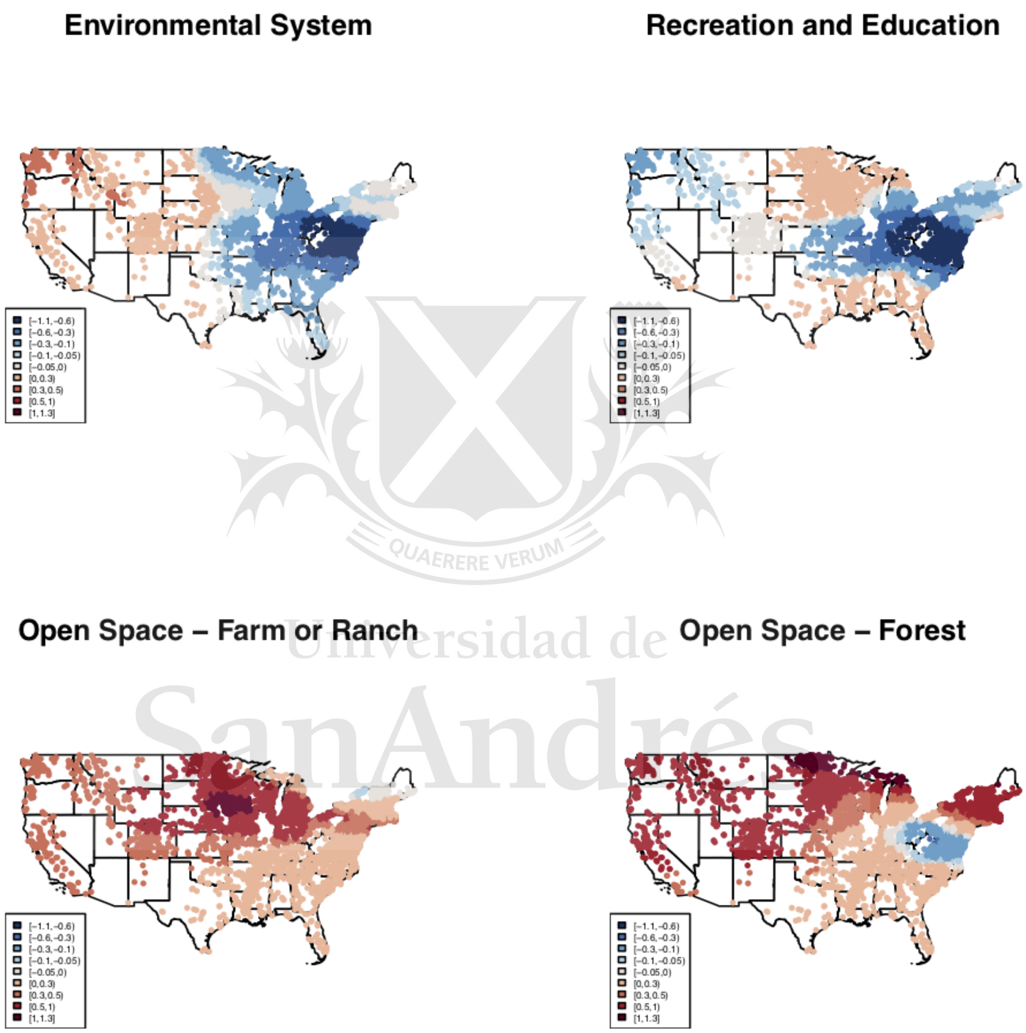


Figure 4: Spatial Variation in the Estimated Coefficients for Easement Size (Log Acres)



The effects for Environmental System somehow follow the standard regions. Easements are closer to cities in the Great Lakes region and North East, but further away in the west. They are further away from highways in the Midwest and South. The probability of this purpose increases for smaller easements in the area to the East of the Mississippi River, and this is true for bigger easements in the West. Income and agriculture value have negative and positive effects respectively for almost the whole country, with the exception of the northern part of the south area. This purpose also seems more likely for easements in southern counties with higher population density.

Many variables influence open space for farm, ranch or forest in the same way. Probability for these uses increases for easements distant to highways in the South. It also increases for smaller easements and easements in counties with higher median income. In the North East, the probability is higher for easements closer to the coast and further away from rail lines. Agriculture value has a positive effect, but median income decreases the probability for these uses.

The effect differ between farm and ranch, and forest for a few variables. Easements closer to the city have a higher probability for farm and ranch in all the East region, but the probability of forest only increases for easements closer to the city in the Great Lakes area. Easements in counties with higher population density have a higher probability for farm and ranch in the south and lower probability in the north central area. Even though the effect for forest is similar, the intensity is much lower. The effect of the number of species is also more intense for farm and ranch than for forest, negative in the South and positive in the North and North East.

5.3 Shift in the Distribution of Estimated Probabilities

An interesting exercise is to calculate the distribution of estimated probabilities for discrete changes in one of the explanatory variables. I show these distributions for changes in two variables: distance to city and easement size (Figures **??** and **??**). Distribution of probabilities are also different for each of the four purposes. Values chosen for this exercise depend on the densities of the variable allowed to change.

Changing distance to city shows opposite effects in the distribution of probabilities of environmental system versus the rest of the uses. Moving away from the city shifts the distribution of estimated probabilities for environmental system to the left and increases its variance. The distribution is multi-modal, but it gets flatter when distances move to 20 miles from the city, showing this variable has a lower effect after a certain point. The opposite happens for other uses, where all

Figure 5: Shift in the Distribution of Conservation Use - Distance to City

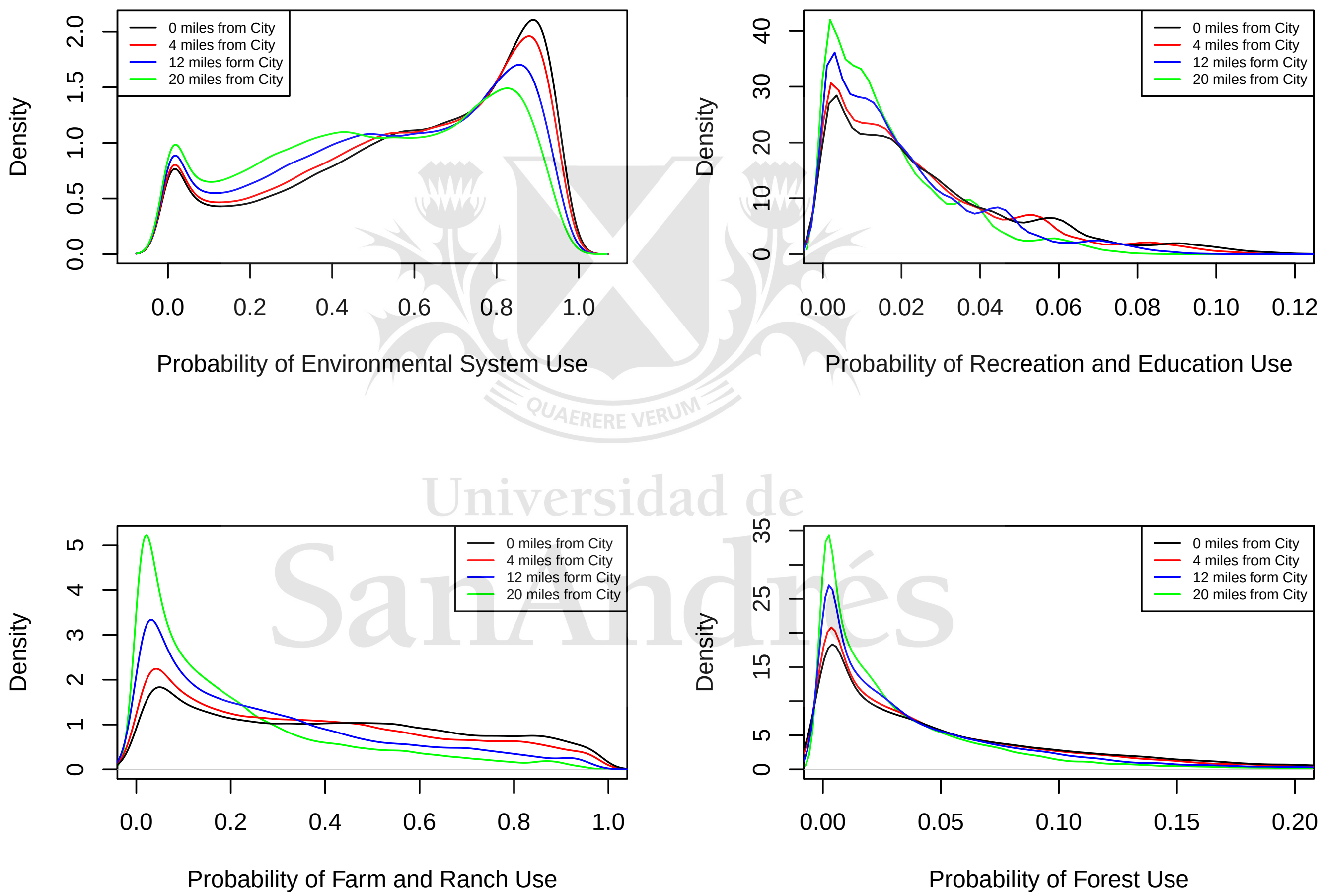
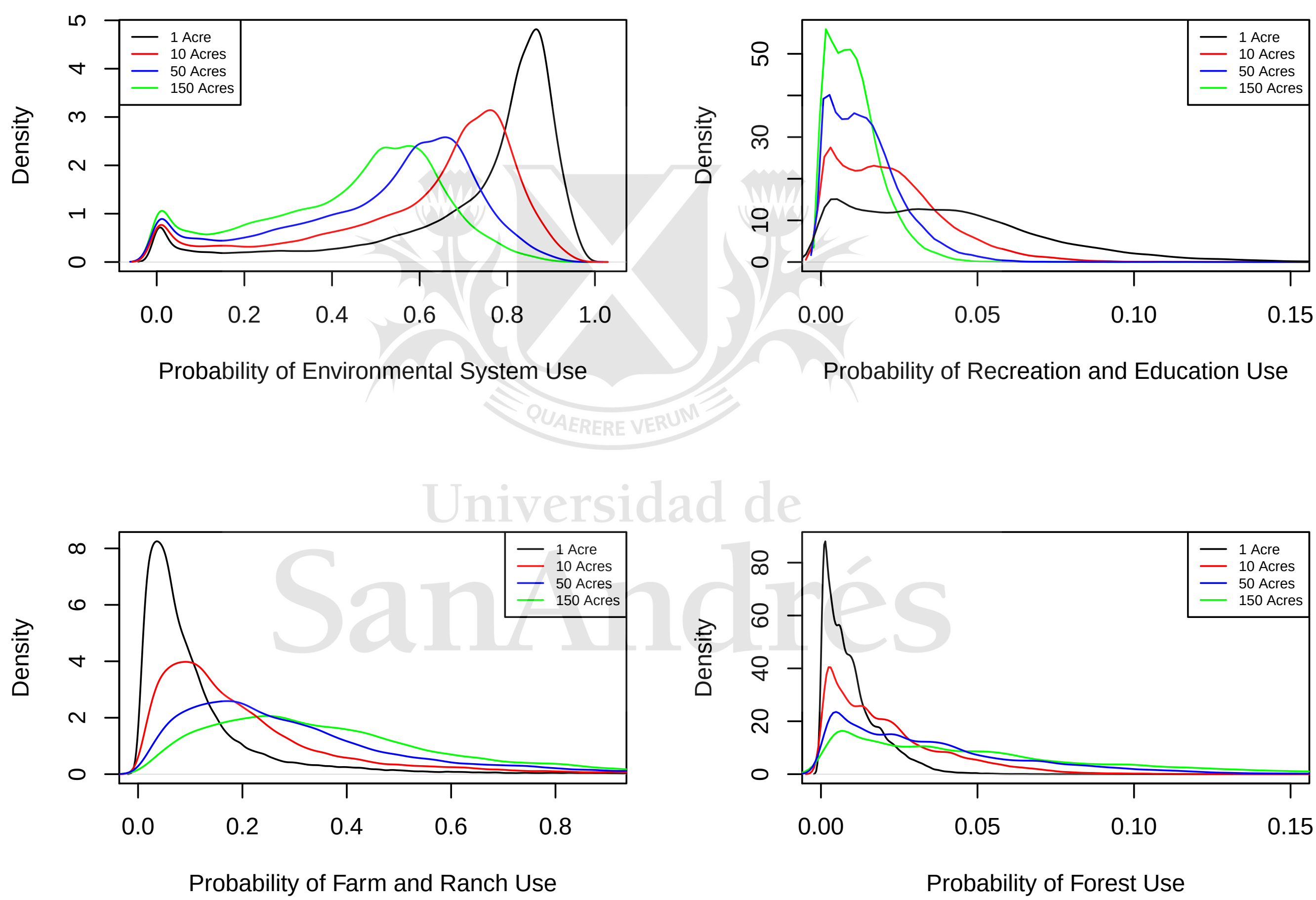


Figure 6: Shift in the Distribution of Conservation Use - Easement Size



distributions are uni-modal. Probabilities are in general lower for theses uses and moving away from the city reduces probabilities and variance.

Easement size has some mixed results. The bigger the easement, the less probable it is to be used for environmental system or recreation. Increasing the easement size shifts the distribution to the left increasing the variance for environmental use, but decreasing the variance for recreation. Distributions of estimated probabilities for farm, ranch and forest shift to the right and have a higher variance when one increases the size of the easements. At 150 acres, both distributions are almost flat.

6 Conclusions

Different purposes of conservation provide different services which affect the decision of where to protect land and for what. Environmental Systems protect terrestrial or aquatic ecosystems, including preservation of endangered species. Open spaces provide scenic view and restrict development, and include subcategories such as farms, ranches, forests or other uses. Recreation and Education easements provide direct benefits to the public. All these different benefits and other factors such as location and biodiversity influence the decision of choosing a particular conservation purpose.

A CPAR ML model presents two advantages with respect to standard or spatial multinomial models. First, it incorporates space without using a weight $n \times n$ matrix that requires a satisfactory definition of neighboring observations and needs to be inverted for the estimation. Second, by allowing the estimation of a distribution of coefficients, one can observe a range of effects that vary smoothly throughout space.

Effects are different by region and conservation purpose. This variability in the results make it difficult to have a general rule that applies to all purposes or all regions. Promoting conservation of recreation easements in the West may require different incentives than in the East Coast. Understanding how different purposes and regions interact is useful to design incentive-based policies and redirect conservation to the desired goals. These results also highlight the importance of designing regional incentives that can account for the observed spatial effects.

An easement in the West has a higher probability of being used for Recreation and Education if it is smaller compared to the ones used for open space. They tend to be away from the city in states like Washington and Oregon, and closer to the city in California. This probability is also higher for easements in counties with low agriculture value or higher median income. Bigger easements and further away from cities have a higher probability of Environmental System use.

The biggest easements are more likely used for farm, ranch and forest in this region.

Easements in the Midwest have a higher probability of Recreation use further away from cities but closer to highways, especially in the north. Easements closer to the city have a higher probability of environmental system, open space for farm and ranch. The bigger the easement, the higher the probability of farm, ranch and forest. Smaller easements are more likely to be used for recreation and environmental system, particularly around the Great Lakes.

In the South, smaller easements also have a higher probability of environmental system use. Getting closer to the city increases the probability of recreation and open space for farm and ranch. Easements located in counties with lower median income have a higher probability of environmental system use, while areas in richer counties are more likely to be used as open space for farm and ranch. Although effects are similar for the whole southern region, the northern part has some distinct characteristics. Recreation and Education is more likely for smaller easements, closer to cities and highways, and in counties with higher agriculture value.

The North East region can be separated in two areas: New England and Mid Atlantic. Easements in New England have a higher probability for environmental system and recreation further away from highways and in more populated counties. Farm, ranch and forest are more likely near highways. In the Mid Atlantic, the probability for all uses is higher away from highways. The probability for Environmental System and Recreation is higher away from cities and for smaller easements.

The fact that coefficients vary throughout space and by purpose highlights again the advantages of using a CPAR Multinomial Logit approach to estimate probabilities of different conservation purposes. Further studies may concentrate on particular reasons driving these differences. Policy makers may find this findings useful to adapt incentive-based strategies to better target specific conservation uses.

References

Adams, V. M., Pressey, R. L., and Naidoo, R. (2010). Opportunity costs: Who really pays for conservation? *Biological Conservation*, 143(2):439–448.

Albers, H. J., Ando, A. W., and Batz, M. (2008a). Patterns of multi-agent land conservation: Crowding in/out, agglomeration, and policy. *Resource and Energy Economics*, 30(4):492–508.

Albers, H. J., Ando, A. W., and Chen, X. (2008b). Spatial-econometric analysis of attraction and repulsion of private conservation by public reserves. *Journal of Environmental Economics and Management*, 56:33–49.

Ando, A. W., Camm, J., Polasky, S., and Solow, A. (1998). Species Distributions, Land Values, and Efficient Conservation. *Science*, 279(5359):2126–2128.

Ando, A. W. and Shah, P. (2010). Demand-side factors in optimal land conservation choice. *Resource and Energy Economics*, 32(2):203–221.

Armsworth, P. R., Kendall, B. E., and Davis, F. W. (2004). An introduction to biodiversity concepts for environmental economists. *Resource and Energy Economics*, 26(2):115–136.

Brady, M. and Irwin, E. G. (2011). Accounting for Spatial Effects in Economic Models of Land Use: Recent Developments and Challenges Ahead. *Environmental and Resource Economics*, 48(3):487–509.

Carrion-Flores, C. and Irwin, E. G. (2004). Determinants of Residential Land-Use Conversion and Sprawl at the Rural-Urban Fringe. *American Journal of Agricultural Economics*, 86(4):889–904.

Carrión-Flores, C. E., Flores-Lagunes, A., and Guci, L. (2009). Land Use Change : A Spatial Multinomial Choice Analysis. *Agricultural and Applied Economics Association 2009 Annual Meeting*.

Chakir, R. and Parent, O. (2009). Determinants of land use changes: A spatial multinomial probit approach. *Papers in Regional Science*, 88(2):327–344.

Chan, K. M. A., Shaw, M. R., Cameron, D. R., Underwood, E. C., and Daily, G. C. (2006). Conservation planning for ecosystem services. *PLoS Biology*, 4(11):2138–2152.

Cleveland, W. S. and Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, 83(403):596–610.

Costello, C. and Polasky, S. (2004). Dynamic reserve site selection. *Resource and Energy Economics*, 26(2):157–174.

Li, M., Wu, J., and Deng, X. (2013). Identifying Drivers of Land Use Change in China: A Spatial Multinomial Logit Model Analysis. *Land Economics*, 89(4):632–654.

Lovell, S. J. and Sunding, D. L. (2001). Voluntary Development Restrictions and the Cost of Habitat Conservation. *Real Estate Economics*, 29(2):191–206.

McMillen, D. P. and McDonald, J. F. (1999). Land use before zoning: The case of 1920' s Chicago. *Regional Science and Urban Economics*, 29:473–489.

McMillen, D. P. and McDonald, J. F. (2004). *Locally Weighted Maximum Likelihood Estimation: Monte Carlo Evidence and an Application.* Advances in Spatial Science. Springer Berlin Heidelberg, Berlin, Heidelberg.

McMillen, D. P. and Soppelsa, M. E. (2015). A Conditionally Parametric Probit Model of Microdata Land Use in Chicago. *Journal of Regional Science*, 55(3):391–415.

Muth, R. F. (1969). *CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE.* University of Chicago Press, Chicago.

Muth, R. F. (1971). The Derived Demand for Urban Residential Land. *Urban Studies*, 8(3):243–254.

Naidoo, R., Balmford, A., Ferraro, P. J., Polasky, S., Ricketts, T. H., and Rouget, M. (2006). Integrating economic costs into conservation planning. *Trends in Ecology and Evolution*, 21(12):681–687.

Newburn, D., Reed, S., Berck, P., and Merenlender, A. M. (2005). Economics and Land-Use Change in Prioritizing Private Land Conservation. *Conservation Biology*, 19(5):1411–1420.

Önal, H. and Yanprechaset, P. (2007). Site accessibility and prioritization of nature reserves. *Ecological Economics*, 60(1996):763–773.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M., and Wilson, K. A. (2007). Conservation planning in a changing world. *Trends in Ecology and Evolution*, 22(11):583–592.

Rissman, A. R., Lozier, L., Comendant, T., Kareiva, P., Kiesecker, J. M., Shaw, M. R., and Merenlender, A. M. (2007). Conservation easements: Biodiversity protection and private use. *Conservation Biology*, 21(3):709–718.

Ruliffson, J. a., Haight, R. G., Gobster, P. H., and Homans, F. R. (2003). Metropolitan natural area protection to maximize public access and species representation. *Environmental Science and Policy*, 6:291–299.

Tibshirani, R. and Hastie, T. (1987). Local Likelihood Estimation. *Journal of the American Statistical Association*, 82(398):559–567.

Von Thunen, J. H. (1966). *Der Isolierte Staat in Beziehung der Landwirtschaft und National ökonomie.* Pergamon Press, Oxford, hall p (ed edition.

Wang, X. and Kockelman, K. M. (2009). Baysian Inference for Ordered Response Data with a Dynamic Spatial-Ordered Probit Model. *Journal of Regional Science*, 49(5):877–913.

Wang, Y., Kockelman, K. M., and Wang, X. (2011). Anticipating Land Use Change using Geographically Weighted Regression Models for Discrete Response. *Transportation Research Record: Journal of the Transportation Research Board*, 2245:111–123.

Wrenn, D. H. and Sam, A. G. (2014). Geographically and temporally weighted likelihood regression: Exploring the spatiotemporal determinants of land use change. *Regional Science and Urban Economics*, 44:60–74.

Wu, J. and Irwin, E. G. (2008). Optimal Land Development with Endogenous Environmental Amenities. *American Journal of Agricultural Economics*, 90(1):232–248.

7 Appendix

The first section of the Appendix include some extra maps that help to have a better picture of how all variables affect different purposes. Some variables have little or no effect in certain regions, but are still useful to understand spatial patterns. The second section shows the shifts in the distribution of Conservation Use for the rest of the variables in the model.

7.1 Spatial Variation in the Estimated Coefficients

Figure 7: Spatial Variation in the Estimated Coefficients for Distance to the Coast

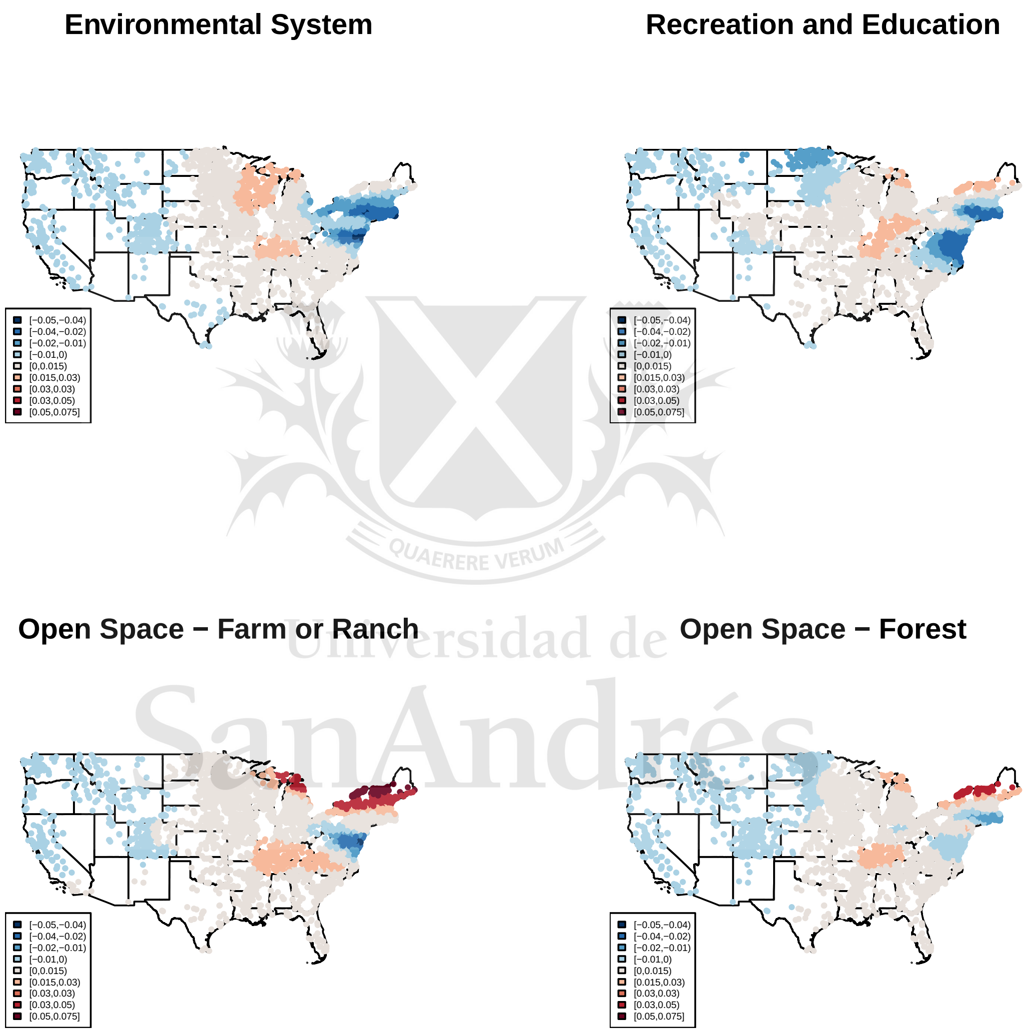


Figure 8: Spatial Variation in the Estimated Coefficients for Distance to Rail Line

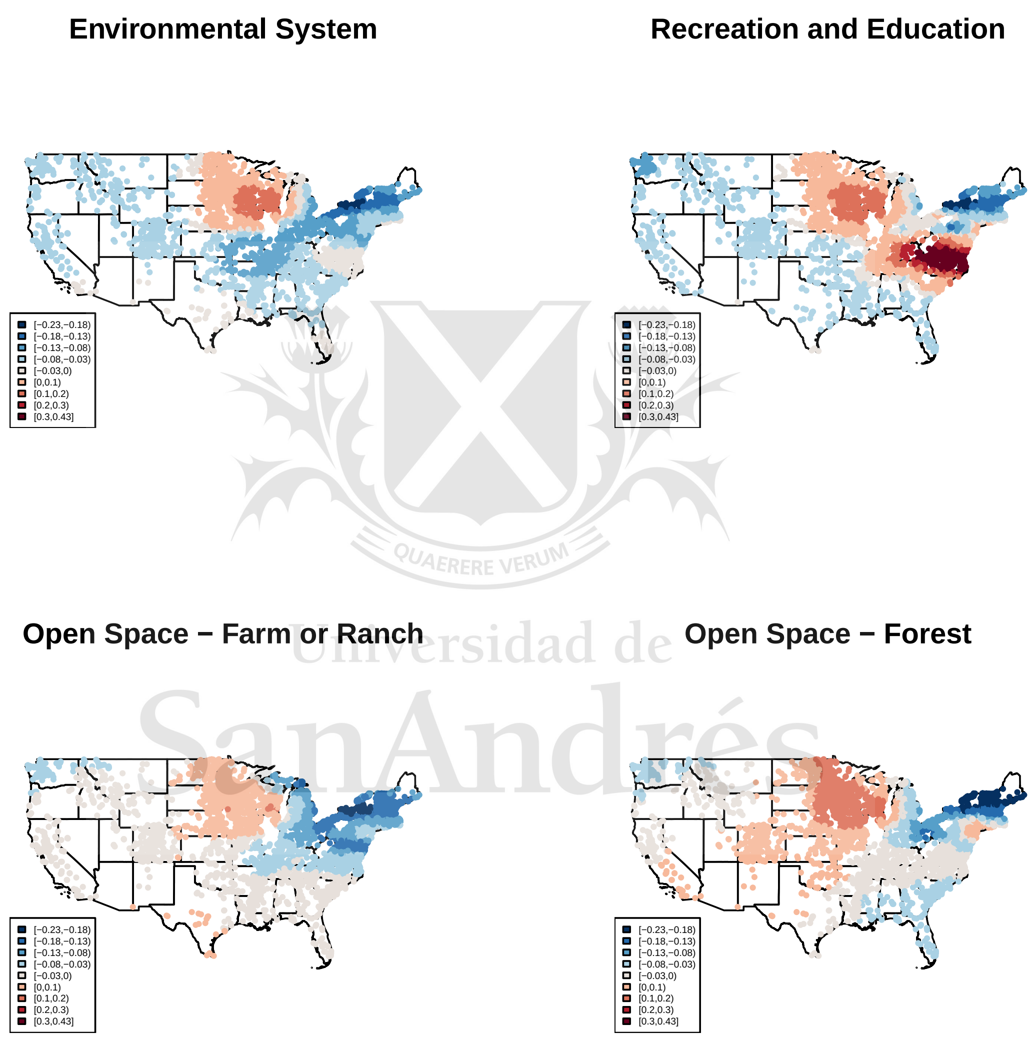


Figure 9: Spatial Variation in the Estimated Coefficients for Median Income

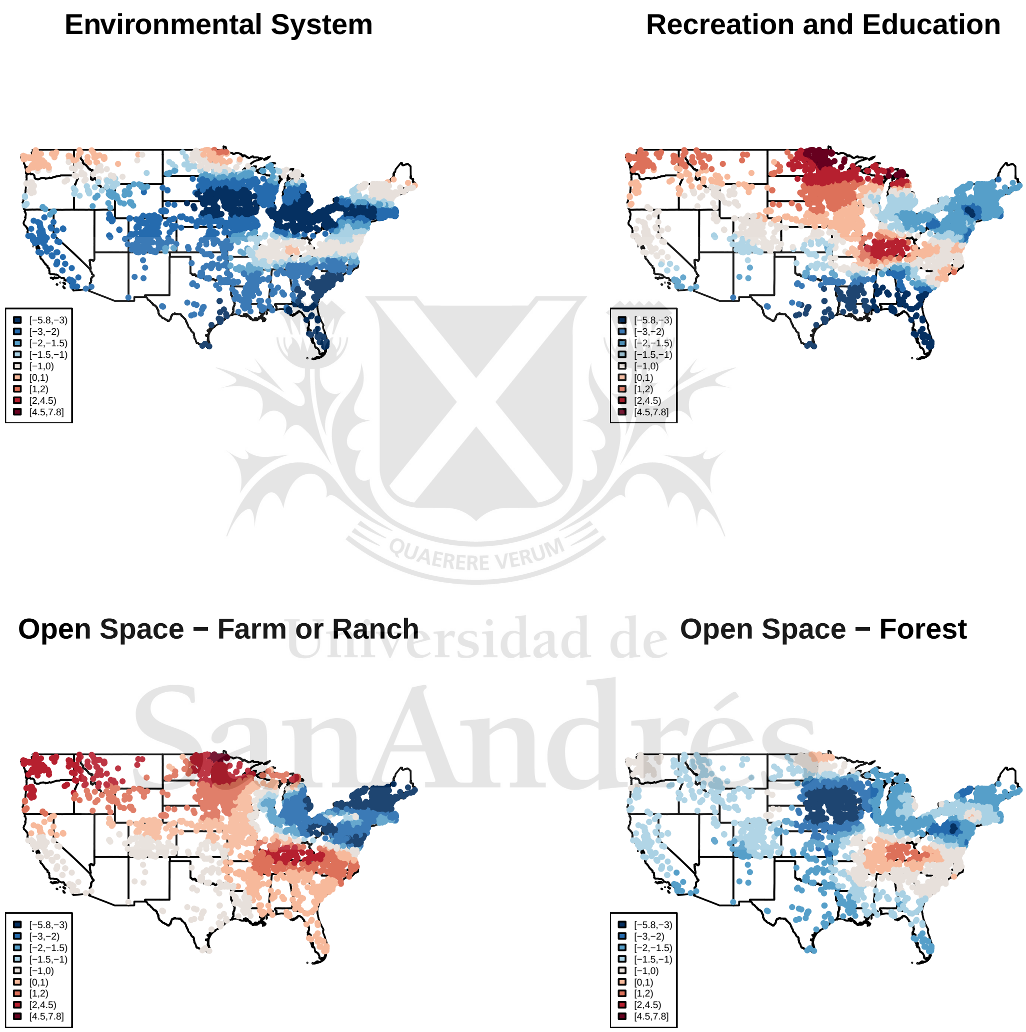


Figure 10: Spatial Variation in the Estimated Coefficients for Population Density

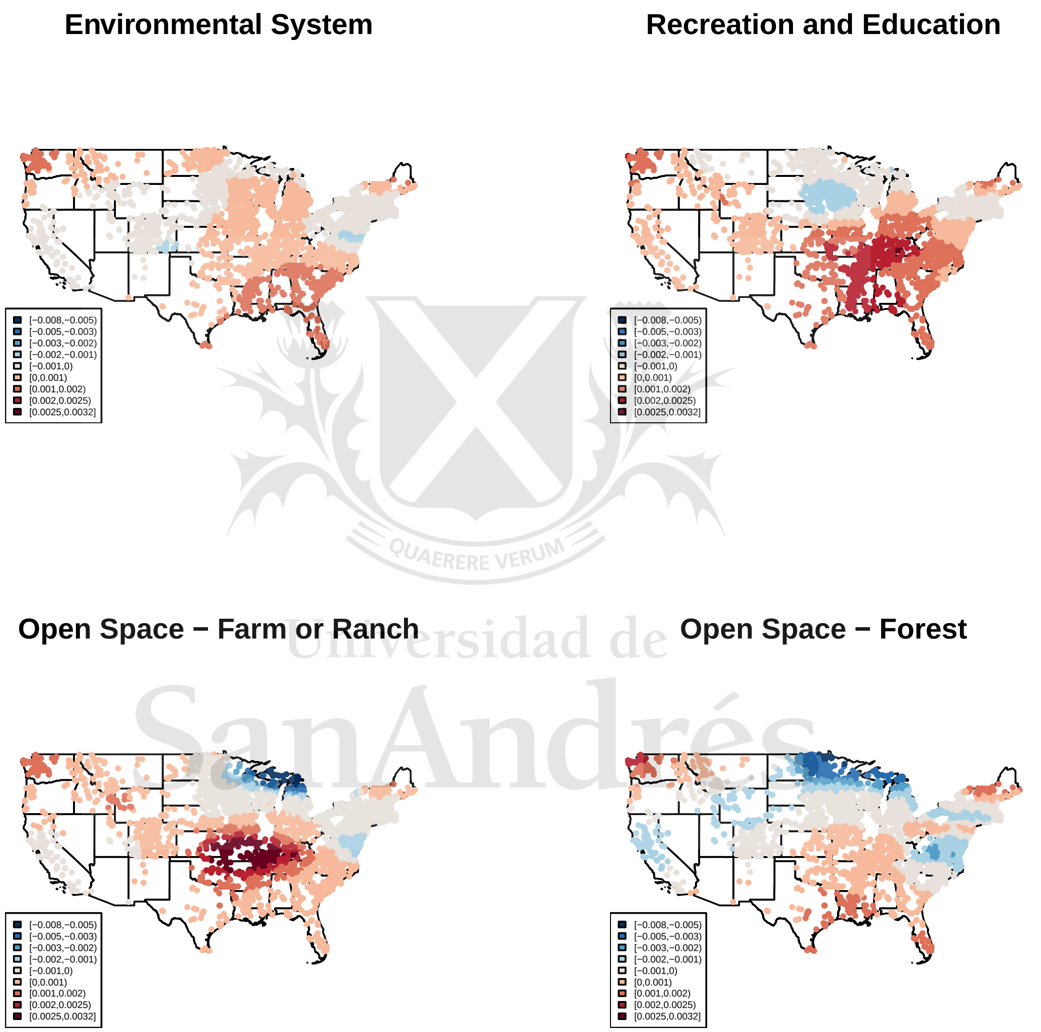


Figure 11: Spatial Variation in the Estimated Coefficients for Number of Species

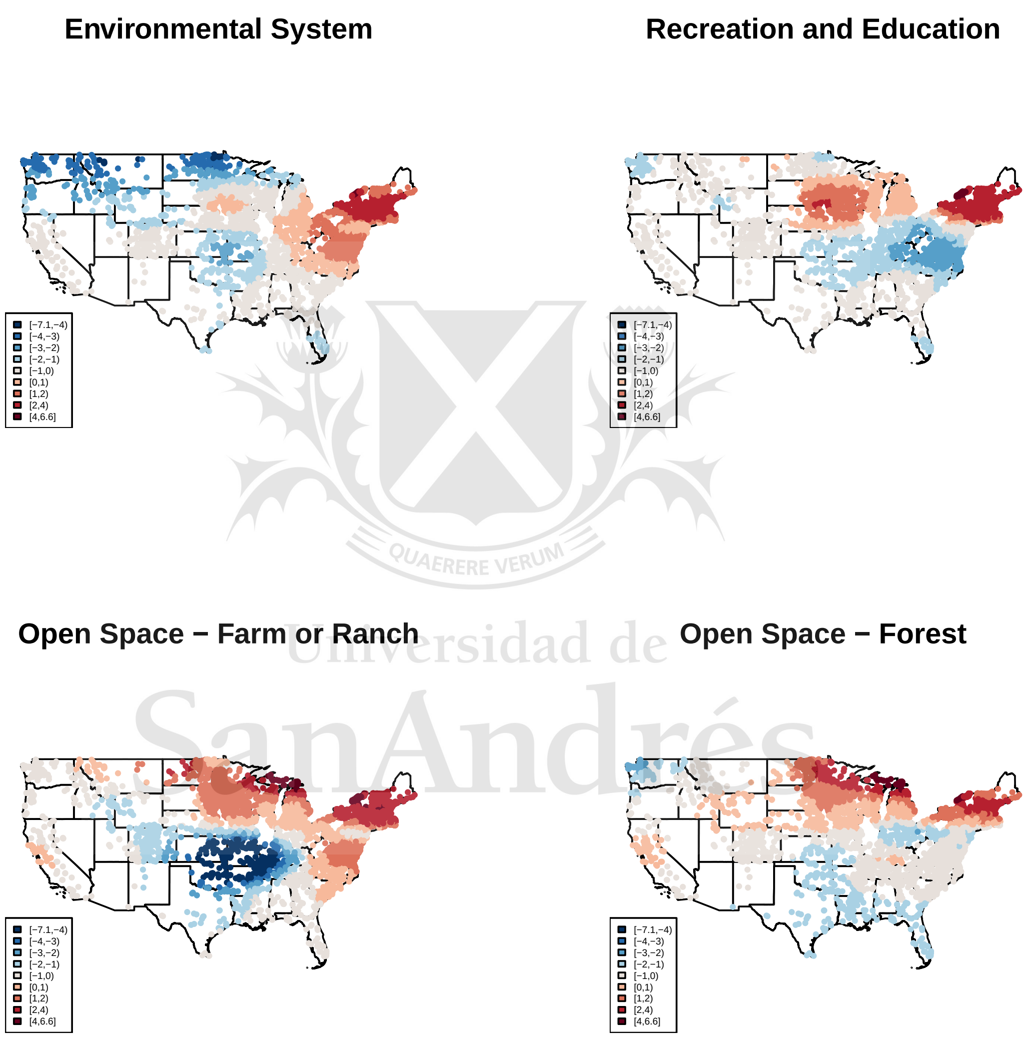
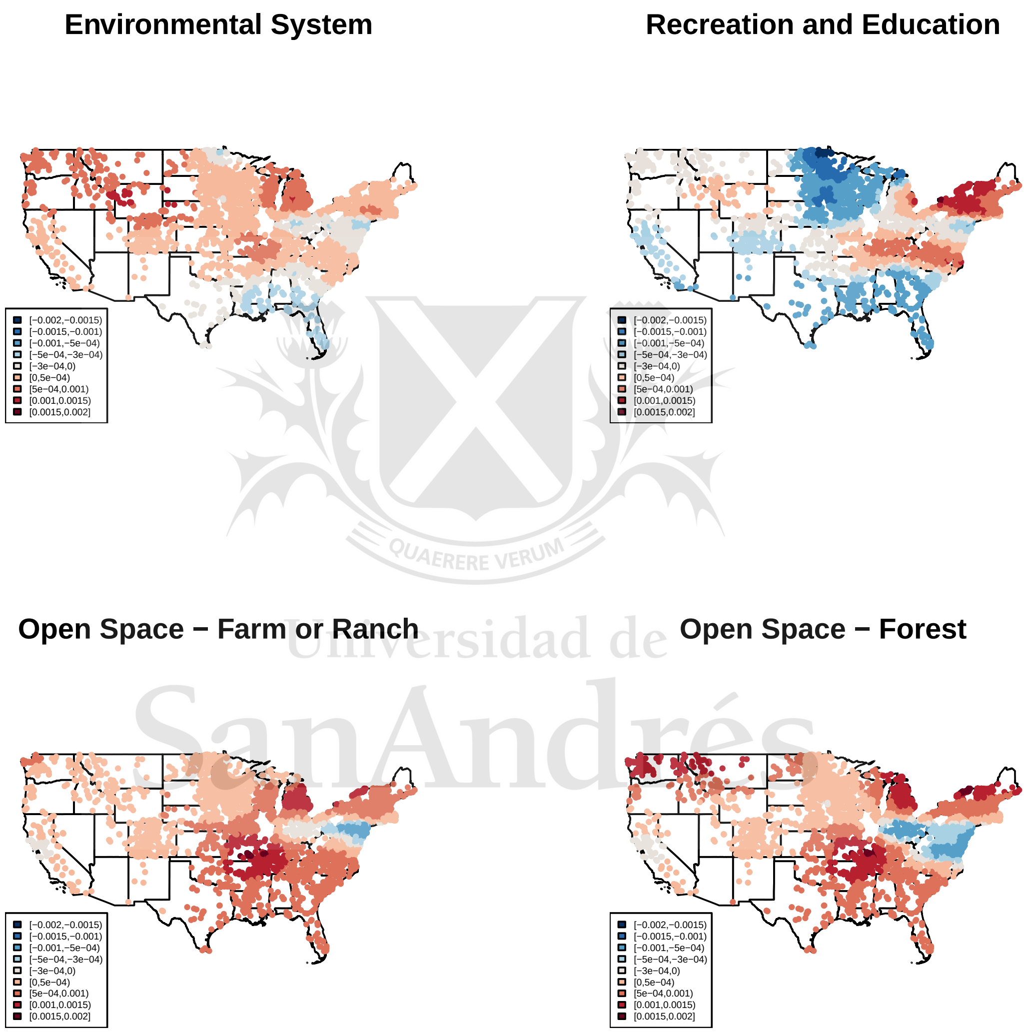


Figure 12: Spatial Variation in the Estimated Coefficients for Agriculture Value



7.2 Shifts in the Distribution of Conservation Use

Figure 13: Shift in the Distribution of Conservation Use - Distance to Highway

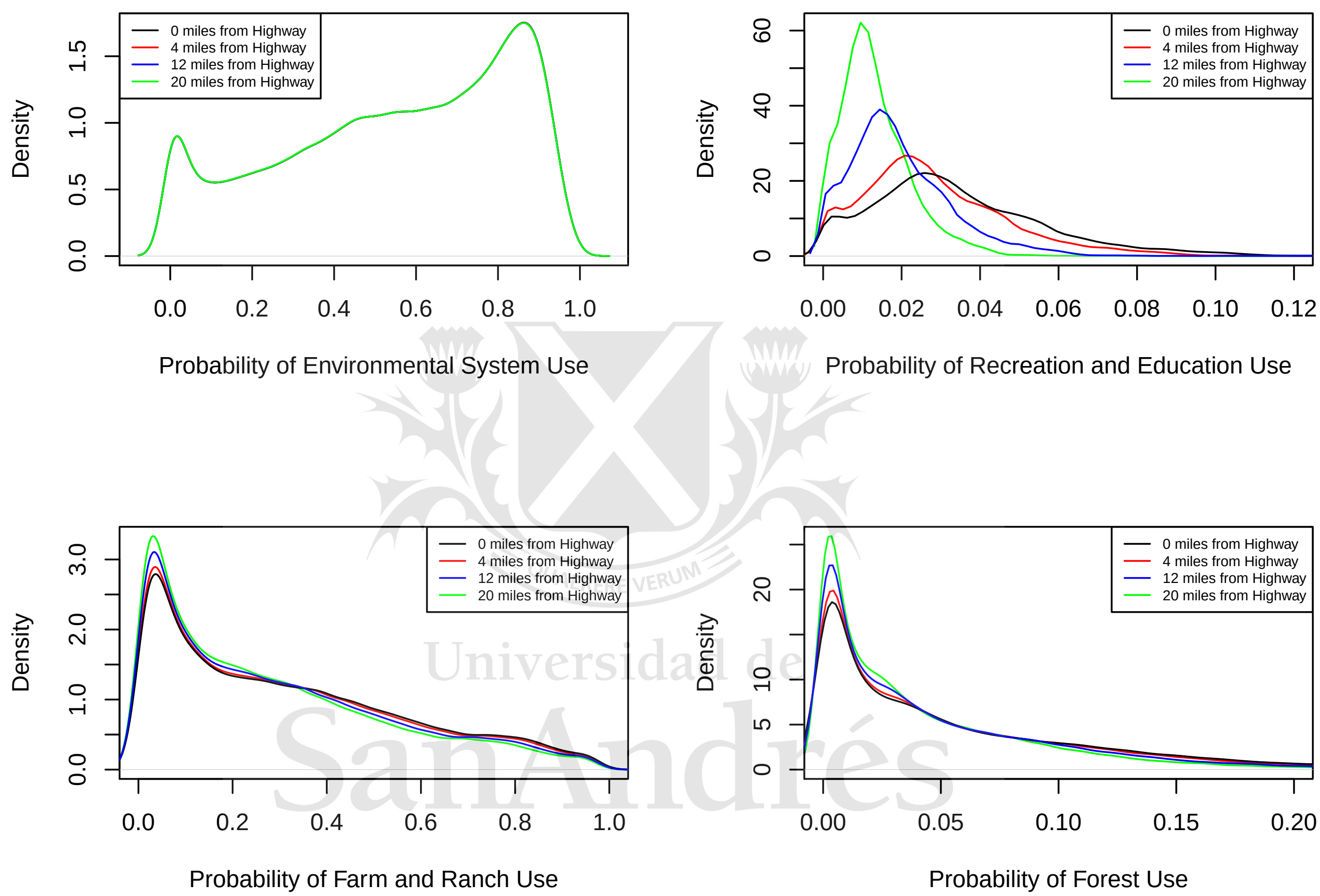


Figure 14: Shift in the Distribution of Conservation Use - Distance to Rail Line

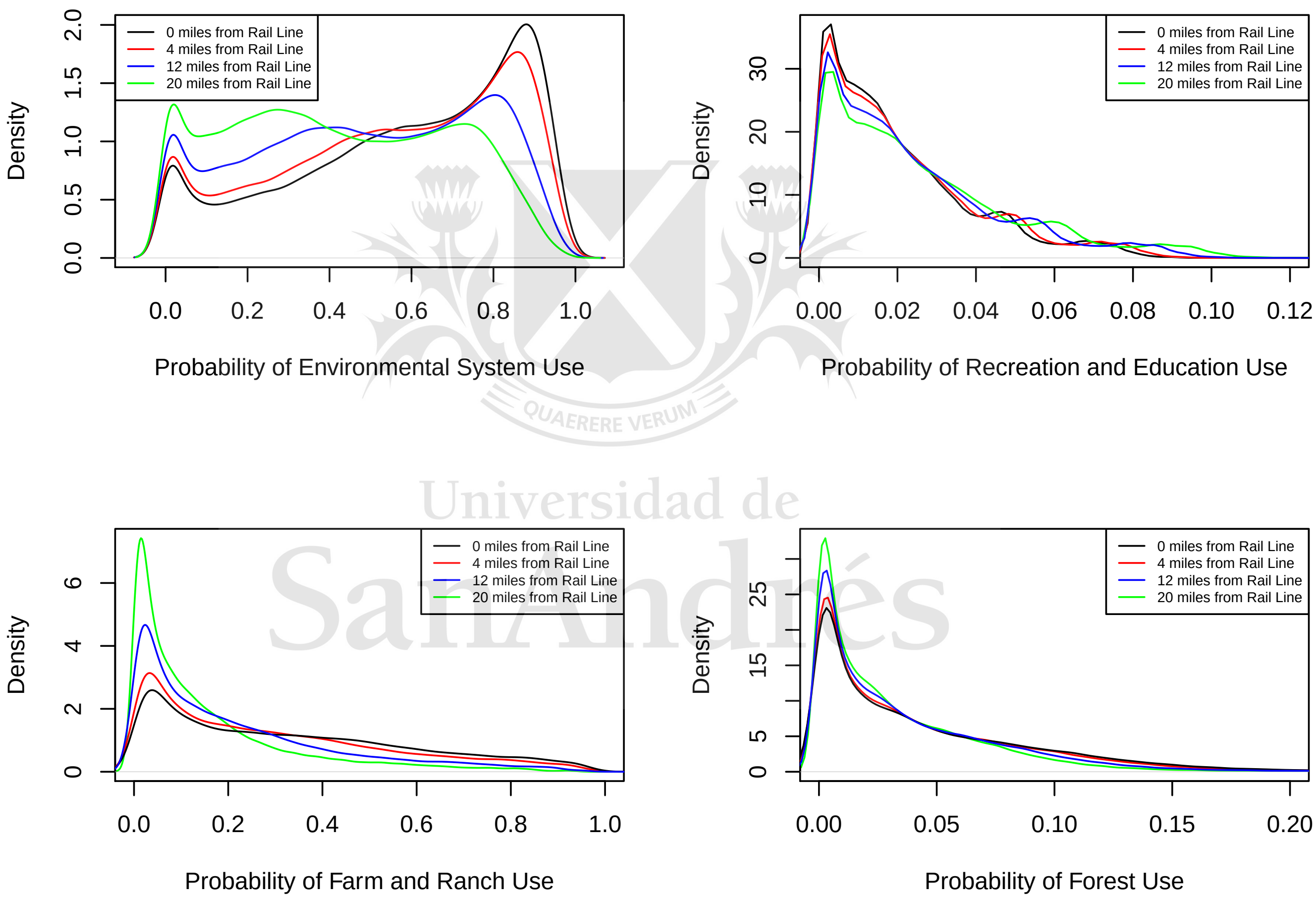


Figure 15: Shift in the Distribution of Conservation Use - Distance to the Coast

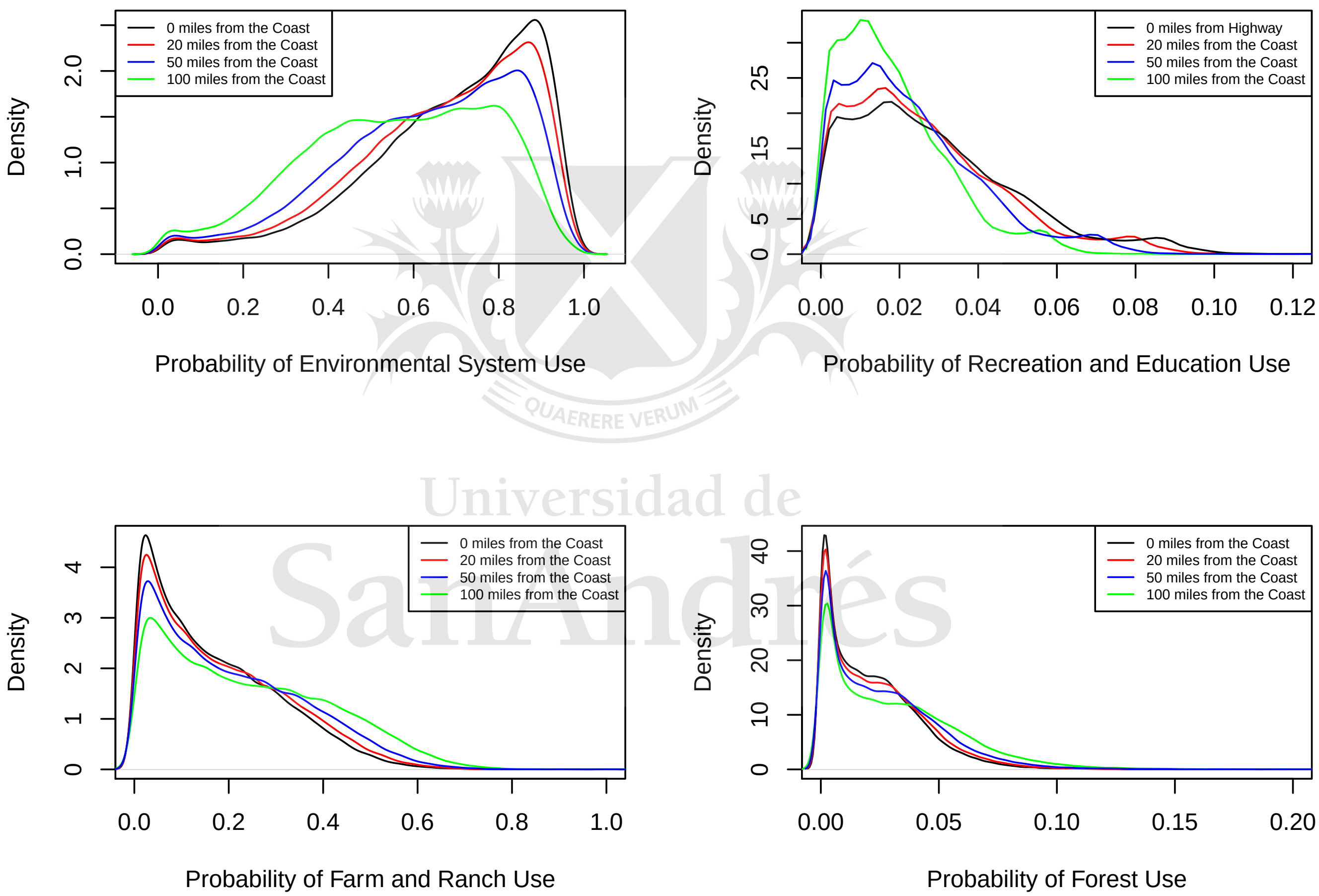


Figure 16: Shift in the Distribution of Conservation Use - Species per Acre

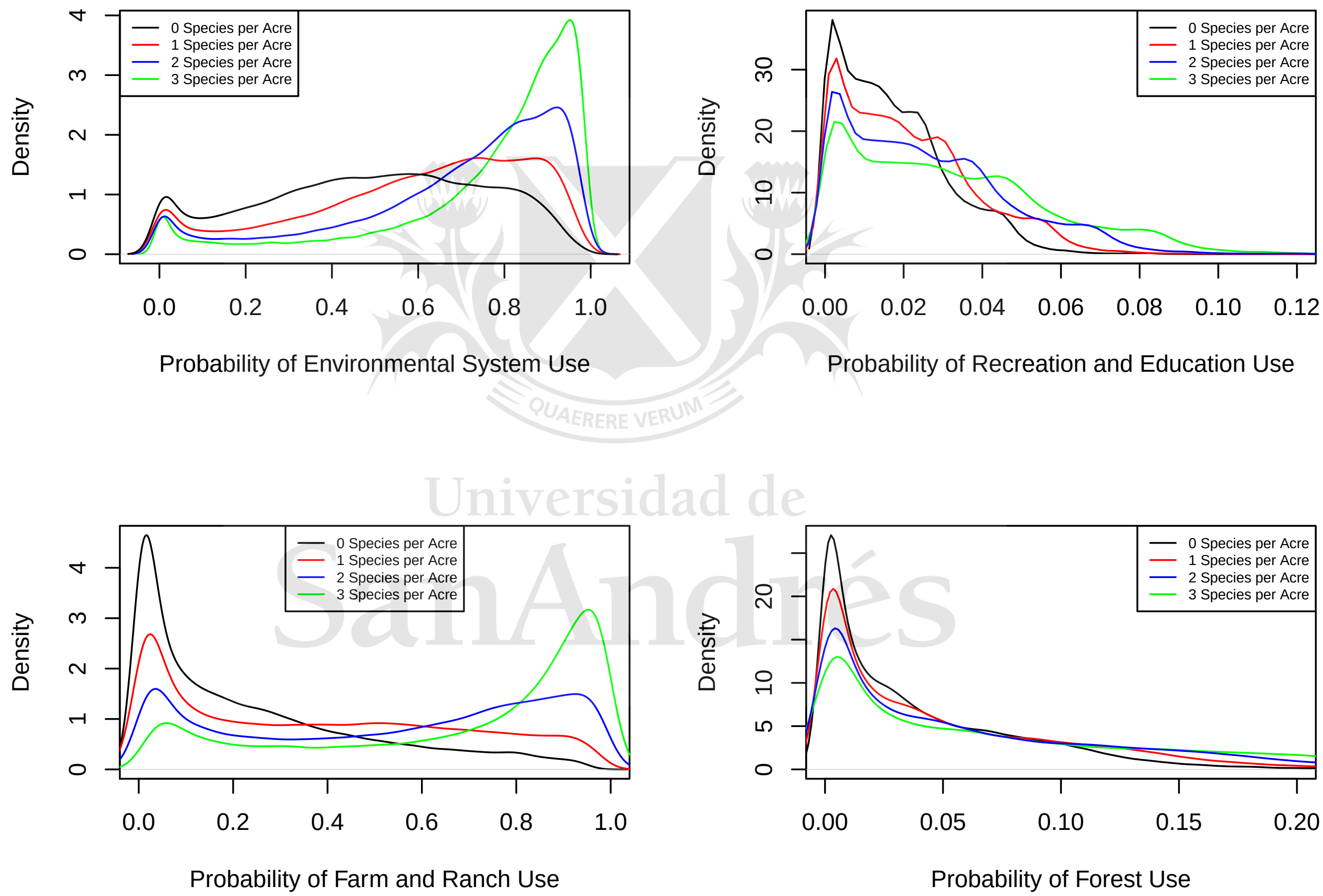


Figure 17: Shift in the Distribution of Conservation Use - HH Median Income

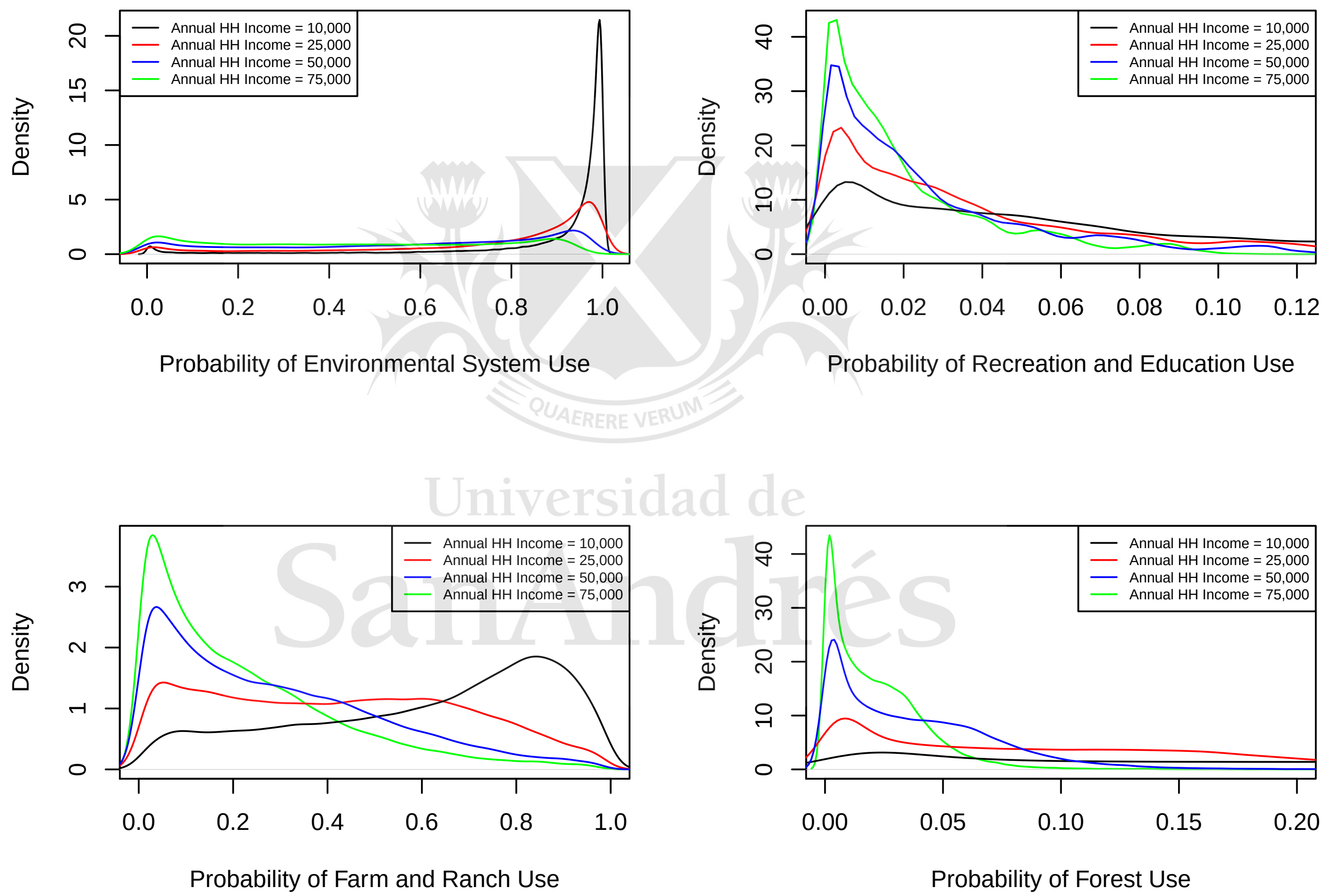


Figure 18: Shift in the Distribution of Conservation Use - Population Density

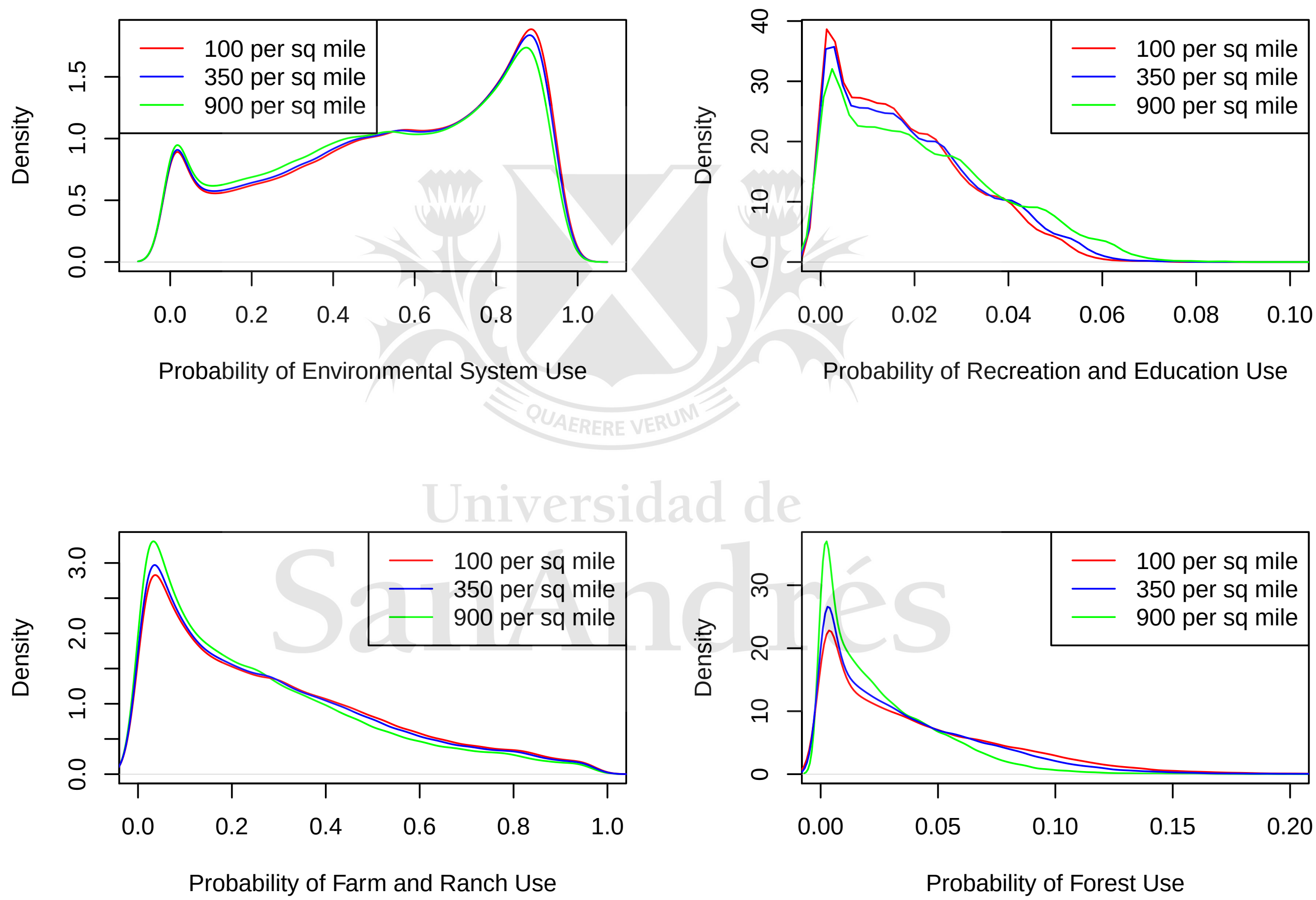


Figure 19: Shift in the Distribution of Conservation Use - Ag. Value per Acre