# Do Monetary Subjective Well-Being Evaluations vary Across Space? Comparing Continuous and Discrete Spatial Heterogeneity

Mauricio Sarrias*
Universidad Católica del Norte
mauricio.sarrias@ucn.cl

July 11, 2018

**Abstract**

Using subjective well-being estimations, this study analyzes whether compensating variations vary across space using a cross-sectional dataset from Chile. To achieve this goal, I describe and compare two econometric ways of modeling unobserved spatial heterogeneity. Both approaches allow compensating variations to vary across spatial units by assuming some distribution a priori. One method assumes that the spatial heterogeneity can be represented by a discrete distribution (group of regions share the same coefficient) and the other that the preferences can be represented by a continuous distribution (each region have a different coefficient). The results show that focusing just on the average estimates of compensating variations, as the applied studies have done so far, masks useful local variation. More empirical studies are needed to assess the advantages and disadvantage of both econometric approaches and how their results compare across a wide range of conditions and samples.

*JEL classification:* C25, I31, R13.

*Key words:* life satisfaction; subjective well-being; compensating variation; continuous and discrete spatial heterogeneity.

---

*IDEAR, Department of Economics, Antofagasta, Chile; Avenida Angamos 0610; Email: mauricio.sarrias@ucn.cl. I am obliged to five anonymous referees; the editor for helpful comments; and Victor Iturra for enlightening conversations. I also express my thanks to FONDECYT, Grant 11160104 for full financial support.

# 1 Introduction

Subjective well-being (SWB) measures have been widely used in different fields, with particular interest in economics (Frey and Stutzer, 2002; Bruni and Porta, 2005; Bruni, 2007; Frey, 2008), sociology (Eid and Larsen, 2008; Deiner et al., 1999) and recently in regional and urban economics (DAcci, 2014; Lin et al., 2014; Glaeser et al., 2016; Goetzke and Islam, 2016). These measures allow researchers to obtain the impact of certain variables on a direct measure of welfare by estimating 'SWB or happiness equations' (Powdthavee, 2010b). The idea is to use self-rated happiness or life satisfaction measures, and to model individuals' SWB as a function of socio-economic characteristics. The resulting coefficients are the estimated average weights of each attribute on happiness.[1]

There is also a growing applied research line that uses self-reported well-being variables for computing compensating variation (CV). This method allows one to value in monetary terms a change in a certain variable (or condition) on individual's welfare, especially for attributes that do not have a market price (Clark and Oswald, 2002). This particular feature makes the CV method an attractive tool for policy makers that would like to know how individuals value intangible resources. For instance, this approach has been used for assessing the monetized value of health impairments, in particular cardiovascular disease (Groot and Maassen van den Brink, 2006) and chronic diseases (Ferrer-i Carbonell and van Praag, 2002). It has also been implemented to compute the value of social relationships (Powdthavee, 2008), crime (Moore and Shepherd, 2006), and for computing monetary values of externalities such as noise damage caused by aircraft noise nuisance (Van Praag and Baarsma, 2005) and other environmental conditions (Welsch and Kühling, 2009; Levinson, 2012).

In essence, CV method computes the trade-off between income and the variable of interest by taking the ratio of the estimated parameters. The resulting ratio is called the CV, which computes the increase or decrease in income necessary to compensate individuals for any given change in some attribute or situation. This simplicity makes this method very appealing compared with other methods such as the hedonic price method or the contingent valuation method, in which people are asked directly their willingness to pay for some attributes (Levinson, 2012).

However, one issue that has been neglected in previous studies is the spatial heterogeneity that might be present in how individuals value some attributes or situations (and hence in CVs).[2] Previous studies compute the CV using 'global

[1] For a more comprehensive review of how these measures have been used in the field in economic see Di Tella and MacCulloch (2006).

[2] It is worth mentioning that there are some studies that examine the spatial aspect of well-being in the economic literature (Brereton et al., 2008; Berry and Okulicz-Kozaryn, 2011; Moro et al., 2008; Rehdanz and Maddison, 2008). Other studies have investigated the spatial distribution of willingness to pay for amenities using choice experiments (Brouwer et al., 2010; Campbell et al.,

models' (as opposed to local models) assuming that the estimated coefficients (and hence the CVs) are unique for all spatial units in the sample. This approach might be suitable if the goal is to obtain and test average CVs for the whole sample or establish stylized facts, regardless of the spatial location of individuals.[3] However, one main drawback is that these methods may hide significant spatial variation and important local differences in the relationship between well-being and individuals' characteristics resulting in misleading and inadequate policy inferences for spatial units that follow a completely different process than the average pattern (Ali et al., 2007).

Spatial heterogeneity is particularly important for policy-making and cost-benefit analysis in problems that involve valuation of non-market attributes. For example, the government may be interested in how individuals value in monetary terms the negative externalities of noise pollution in their neighborhood, so as to create policies that compensate for the loss of welfare. Now imagine that a certain study indicates that individuals, regardless of their location, should be compensated say in average by US$50 in order to let them reach his previous welfare without the noise pollution. However, inhabitants in, let's say, region A may be more sensitive to noise to those in region B due to observed and/or unobserved factors. Suppose that given this spatial heterogeneity in the sensitivities, individuals living in region $A$ should be compensated by US$80, and those in region $B$ by US$50. As a result, a spatially blind policy based on a unique estimate of CV for all regions may ignore this fact creating a welfare-gap of US$30 not covered for inhabitants in region $A$, and thus producing undesired geographically uneven impacts.

Under this context, this paper seeks to address the limitation of previous studies by taking into consideration the spatial heterogeneity that might be present in individuals' evaluations, where spatial heterogeneity is understood as different realizations of the parameters for each spatial unit. To accomplish this, I describe and compare two econometric ways of modeling unobserved spatial heterogeneity. The idea in both approaches is to allow the coefficients to vary across spatial units by assuming some distribution specified a priori. By allowing spatially varying coefficients I will be able to obtain location-specific CVs. The main difference between both approaches is whether the chosen distribution is continuous or discrete. The hypothesis of spatial heterogeneity in CVs is tested using a Chilean database with 'communes' as spatial units.[4]

I present two main results. First, both approaches reveal that there exist significant amounts of spatial variation on individuals' valuations for some attributes.

---

2009).

[3]Throughout this study I will use location unit, region, or geographical area interchangeably.

[4]Communes are the smallest geographical areas in Chile for which individual data is available.

This is a very important result for the literature on happiness and economics: neglecting spatial heterogeneity an ignoring local differences on CVs might lead to inadequate policy design for some group of regions. In effect, if different regions have different valuations of a certain observable attribute, then one should account for this heterogeneity when analyzing the potential outcomes of a policy. The second result is that the discrete case shows a slight superiority over the continuous case, at least for this sample. This might be explained by the complexity of the estimation procedure when assuming continuous unobserved spatial heterogeneity. However, the results are not conclusive.[5]

The remainder of this study is organized as follows. Section 2 discusses a simple model used to understand spatial heterogeneity in CVs, along with its underlying assumptions. Section 3 explains the continuous and discrete strategies for modeling spatial heterogeneity. Section 4 presents the data and results. Finally, Section 5 discusses the main findings and concludes.

# 2 Model, assumptions and spatial heterogeneity

## 2.1 Model and assumptions

Similarly to Ferrer-i Carbonell and van Praag (2002), I assume that the true latent well-being of individuals is determined by income $Y$, specific health conditions $H$, socio-economic $X$ characteristics, and neighborhood attributes $Z$:

$$W_i^* = W(Y_i, H_i, X_i, Z_i). \tag{1}$$

This latent function can be thought of as an operationalization of an indirect utility function, which gives the individuals' maximal attainable utility when faced with those characteristics (see for example Van Praag and Baarsma, 2005; Bockstael and Freeman, 2005). Adopting this notation, one could ask for example: *what is the change in income needed to compensate individuals for the marginal change of neighborhood characteristics?* Mathematically, this is achieved by totally differentiating Equation (1) in equilibrium, setting $dW^* = 0$, and holding all the other variables constant such that:

$$CV = \left.\frac{dY}{dZ}\right|_{dW^*=0} = -\frac{\frac{\partial W}{\partial Z}}{\frac{\partial W}{\partial Y}}. \tag{2}$$

This ratio measures the marginal change in $Y$ needed to bring the individual to

[5] Although the main focus of this article is to highlight the spatial heterogeneity of the CVs, it is also important to underline that the results of this article are not free of potential endogeneity problems. Section 4.2 discusses the potential biases that can be expected in terms of CVs. Thus, this study makes strong assumptions about the exogeneity of the variables, but offers valuable insights when analyzing the spatial heterogeneity of CVs.

his original level of well-being given a marginal change in $Z$, *ceteris paribus*. Due to the negative sign, CV will be negative if $Z$ measures good quality of the neighborhood and positive if it measures a bad characteristic, holding other things equal.

In order to operationalize Equation (1), I proxied $W_i^*$ by self-reported subjective well-being using the answer to the following question "*How satisfied are you with life?*" The answer ranges from 1 ("completely unsatisfied") to 10 ("completely satisfied"), and then I use a dichotomy version of this categorical answer. According to Ferrer-i Carbonell and Frijters (2004), this approach has three main theoretical assumptions.

- *Monotonicity*: It is assumed that SWB questions are a positive monotonic transformation of the underlying concept called welfare/utility.
- *Ordinal comparability*: Using the welfare approach ordinal comparability is assumed, that is, if $SWB_i > SWB_j$, then $W_i > W_j$ for individual $i$ and $j$. That is, individuals use discrete SWB rating in the same way and share a common opinion of what happiness is.
- *Interpersonal cardinality*: This assumption amounts to assuming that the difference between a satisfaction answer between 9 and 10 is the same as the difference between 2 and 3, and therefore we can make interpersonal comparisons.

Monotonicity assumption implies that SWB actually measures welfare/utility. The problem is that welfare cannot be easily and objectively measured, and thus it is difficult to test some sort of correlation between both. Daniel Benjamin and colleagues have made contributions to this debate using experimental studies. For example Benjamin et al. (2012) analyze whether SWB questions are a good proxy for person's preferences, and hence a good proxy for utility. They conduct an experiment where respondents are faced with hypothetical scenarios and derive both choice and anticipated SWB ranking of two alternatives.[6] According to the results, the authors argue that individuals do not seek to maximize SWB exclusively, but that SWB is a uniquely important argument of the utility function and thus it can be used as a proxy for utility. They also found that hypothetical choices maximize predicted life satisfaction (compare with other SWB questions) for most of the respondents in most of the scenarios. However, in settings where one alternative involves higher income or more money, they found that respondents are systematically more likely to choose the money alternative than they are likely to predict it will yield higher SWB. In other words, individuals are more sensitive to money in hypothetical choices than in

[6]In order to obtain anticipated SWB they ask whether the respondent anticipates more life satisfaction for a particular hypothetical situation.

anticipated SWB measures. In view of Equation (2), the implication of this result is that CVs obtained by any regression may be upward biased.

Benjamin et al. (2014) ask directly whether CVs (or marginal rate substitution) based on SWB reflect preference-based CVs. Using residential choices of medical students, they find substantial differences between the CVs implied by anticipated SWB measures and those revealed by choices. However, when searching which of the anticipated SWB matches most closely to the choices, and similar to Benjamin et al. (2012), they found that ladder-type questions and life satisfaction do better than happiness.

Both previous studies uses hypothetical well-being data. In contrast, Perez-Truglia (2015) test the validity of experienced (as opposed to anticipated) well-being data. Perez-Truglia (2015) find that SWB is correlated with objective measures of well-being, such as suicide rates and frequency of smiling. But, life satisfaction performs significantly better than other objective and subjective measures of well-being. In line with Benjamin et al. (2014) and Benjamin et al. (2012), Perez-Truglia (2015)'s evidence suggests that life satisfaction offers some useful information about utility.[7] In general, measures of reported SWB are viewed as valid and reliable empirical proxy to individual's utility (Diener et al., 1995; Welsch and Kühling, 2009).

In the context of SWB studies, economists have presented different views to deal with the ordinal comparability and interpersonal cardinality assumption. Broadly, the ordinal comparability assumption tells us that all individuals that evaluate their well-being by say 4 have about the same satisfaction (i.e., they are on the same indifference curve). According to Van Praag (2007) and Ferrer-i Carbonell and Frijters (2004) this is not an unreasonable assumption for respondents who have about the same cultural and linguistic background. Since this paper focuses entirely on Chile, this assumption seems plausible for this study.

The interpersonal cardinality is one of the strongest assumptions. However, I am interested in the CVs, which are invariant to any monotonic transformation of the SWB function and intrinsically ordinal, thus it is not necessary to assume cardinality (see Levinson, 2012; Frey et al., 2010; Welsch and Kühling, 2009, for similar assumption).[8]

In sum, this approach has its own weakness since it makes strong assumptions

[7]There are some other studies that have also found a strong and positive correlation between stated SWB measures and emotional expressions. See for example Deiner et al. (1999) or Ferrer-i Carbonell and Frijters (2004) and the references cited therein.

[8]As further explained by Adler (2012), since the CV (or the marginal rate of substitution) is invariant to any monotonic transformation of the subjective well-being functions, no cardinal utility function is required. Furthermore, the direct effects or marginal utilities, captured by the coefficients, only shift the level of satisfaction upward or downwards, without changing the marginal rates of substitution between the life dimensions.

about preferences in the sense that it compares stated well-being of different individuals. Therefore, the link between true well-being and subjective well-being is not without shortcomings. However, the set of assumptions of this approach are not stronger than those of other methods used to compute compensating variations (such as the hedonic or contingent evaluation approach), which have their own sets of strong assumptions (Levinson, 2012; Welsch and Kühling, 2009).

## 2.2 Spatial heterogeneity

Equations (1) and (2) assume fixed compensating variation for each attribute across space. However, one might expect that how people value attributes is not spatially stationary, but rather region-specific. Naturally, a question that arises is: why individuals' preferences might vary across geographical space? According to Fotheringham et al. (2003) there are at least three reasons. The first one, and less interesting, is sampling variation, which relates to a statistical artifact and not to any underlying spatial process. A second potential cause is misspecification or important variables that follow a spatial non-stationary process but are omitted from the model. The third possibility, and more theoretical, is that people's preferences for some attributes are intrinsically different across space. That is, individuals may differentiate strongly in terms of valuation of changes in characteristics among some locations or areas producing different reactions to the same stimuli over space. As an example, consider neighborhood amenities such as urban parks or green areas. In this case, spatial heterogeneity in preferences may be due to the spatial configuration of such goods in terms of amount, quality and distance (Lanz and Provins, 2013), and the variability of substitutes (Brouwer et al., 2010) or to residential sorting (Bateman and Willis, 2001; Franceschinis et al., 2016). Because tastes for such goods differ among people, we could expect those with stronger preferences for these amenities to sort themselves into more amenable places (Roback, 1982). This point will be discussed later.

The link between preferences and CV is shown in Figure 1. The top panel shows the indifference curve for a representative individual from region 1, while the bottom panel shows a representative individual from region 2, holding $H$ and $X$ being constant. People in region 2 are more sensitive to changes in neighborhood quality than those to region 1, which is represented by a steeper slope.

The representative individuals in both regions are initially located in point $A$ with income $y_0^1$ and $y_0^2$, respectively, and the same amount of neighborhood quality $z_0$. Now assume that neighborhood quality decreases in both regions by $\Delta z = z_1 - z_0$, which implies a move to $B$ where individuals from both regions are worse off with welfare $W_1^1$ and $W_1^2$, respectively. The compensating variation, which takes the representative individual from both regions back to $W_0^1$ and $W_0^2$ at point $C$ while accepting the worsening of neighborhood quality $z_1$, is represented $\Delta y = y_1 - y_0$ which is positive. However, individuals in region 2 are, in average, more sensitive to changes in neighborhood quality, that is, $(|\Delta y^2| > |\Delta y^1|)$, therefore they need a bigger CV, given the same change in $Z$. We might be able to remedy the foregoing concern by writing Equation (2) as:

$$CV_c = -\frac{\left[\frac{\partial W}{\partial Z}\right]_c}{\left[\frac{\partial W}{\partial Y}\right]_c}, \tag{3}$$

where the marginal effects of $Z$ and $Y$ vary with location $c$. In other words, we can expect different CVs for each geographical location because individuals differentiate in terms of valuation of changes in neighborhood characteristics and/or in terms of different sensitivities to money changes among locations.

In Section 3, I present the econometric approach used in this work to model location-specific compensating variations. However, up to this point, it should be kept in mind that there are some restrictions. First, both the theoretical model and subsequent econometric implementation allow for spatial heterogeneity across locations, but not within locations. However, this assumption is not as restrictive as assuming fixed sensitivities across geographical space on the underlying utility function. Second, due to restrictions in the estimation method further explained in Section 3.3, the calculation of compensating variations assumes that the marginal effect of income is fixed for the continuous case, that is, what makes CVs vary is just different sensitivities to the variable of interest.

# 3 Econometric modeling

## 3.1 Continuous and discrete spatial heterogeneity

I assume that Equation (1) for individual $i$ in commune $c$ can be operationalized by an underlying continuous but latent well-being process $W^*_{ic}$ based on a linear combination of covariates given by:

$$\begin{aligned} W^*_{ic} &= \mathbf{x}'_{ic}\boldsymbol{\beta}_c + \upsilon_c + \epsilon_{ic}, \\ \boldsymbol{\beta}_c &\sim g(\boldsymbol{\theta}), \end{aligned} \tag{4}$$

where $\mathbf{x}_{ic}$ collects all the model's variables $(Y, H, X, Z)$; $\epsilon_{ic} \sim N(0, \sigma^2_\epsilon)$ is the error term, but since the scale of $W^*_{ic}$ is not identified I normalize $\sigma_\epsilon = 1$; $\upsilon_c$ is a commune-specific unobserved factor which is independent of $\epsilon_{ic}$ and can be interpreted as capturing all those effects not captured by the variables at the commune level or as an unobservable random cluster component.

A main feature in formulation (4) is that I assume that the marginal (dis)utility of predictors ($\boldsymbol{\beta}_c$) are not constant, but rather inherently different across geographical space. This implies that all individuals located in the same commune have the same coefficient, but there exists inter-spatial heterogeneity, i.e., marginal (dis)utility vary across communes but not within the communes.

Spatial heterogeneity is modeled by allowing the parameters associated with each observed variable to vary 'randomly' across space according to some distribution $g(\boldsymbol{\theta})$. However, we do not know how the parameters vary across space. All we know is that they vary locally with population probability density function (pdf) $g(\boldsymbol{\theta})$, which is assumed to be well-behaved. Since the distribution is unknown it has to be chosen a priori by the researcher.[9] It is worth mentioning that by assuming spatial random coefficients the CVs—which are the ratios of the estimated parameters—are also random and vary across space.

I compare two competitive ways of modeling spatial heterogeneity. First, I assume that there exists continuous spatial heterogeneity. This is introduced by assuming that $g(\boldsymbol{\theta})$ is a continuous function. This pdf can in principle take any shape, but the researcher must choose one according to his beliefs of the domain and boundedness of the coefficients (Hensher and Greene, 2003). Once the distribution is chosen, the pdf of the spatially random coefficients in the population is $g(\boldsymbol{\beta}_c|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents, for example, the mean and variance of the distribution of $\boldsymbol{\beta}_c$. The goal is to estimate $\boldsymbol{\theta}$, and then get a profile of the distribution. As an illustration, we can assume that one of the coefficients is normally distributed. Therefore, we can write:

$$\beta_c \sim g(\beta_c|\boldsymbol{\theta}) \implies \beta_c = \beta + \sigma_\beta \omega_c,$$

where $\omega_c \sim N(0,1)$, and $\boldsymbol{\theta} = (\beta, \sigma_\beta)'$. By the normality assumption, we make the implicit assumption that the coefficient in each commune can take any real number in the interval $(-\infty, +\infty)$. Obviously, any other distribution with some restrictions in the domain can be used.[10]

Instead of assuming a continuous distribution for the spatially random coefficients, one may also assume that they are distributed across space following a discrete distribution. In this case, spatial heterogeneity is modeled by assuming that there is a discrete number, $Q$, of classes of communes where the communes belonging to the same class share the same coefficients. Formally, the distribution can be written as:

[9] Here it is important to emphasize that, unlike the Geographically Weighted Regression (Brunsdon et al., 1998) which forces a certain degree of correlation through geographical proximity, our approach does not impose any restriction (beyond the distribution) and allows the coefficients to vary freely across space. In this sense, the approach used here is close in spirit to the Random Coefficient Model (Swamy, 1971), but it differs in that it uses data at the individual level in order to identify the parameters at the regional level.

[10] See for example Hensher and Greene (2003) for a discussion about some distributions and their implications.

$$\beta_c \sim g(\beta_q|\boldsymbol{\theta}_q) \implies \beta_c = \beta_q \quad \text{with probability} \quad w_{cq}, \tag{5}$$

where commune $c$ belongs to class $q$ with probability $w_{cq}$, such that $\sum_q w_{cq} = 1$. The researcher does not know from the sample which region is in which class, hence the assignment probability in (5) is unknown. Thus, the number of classes $Q$ and the discrete distribution must be chosen a priori. The most widely used formulation for $w_{cq}$ is the semi-parametric multinomial logit format (Greene and Hensher, 2013; Shen, 2009; Hess, 2014):

$$w_{cq} = \frac{\exp(\mathbf{h}_c'\boldsymbol{\gamma}_q)}{\sum_{q=1}^{Q}\exp(\mathbf{h}_c'\boldsymbol{\gamma}_q)}; \quad q = 1, ..., Q, \boldsymbol{\gamma}_1 = \mathbf{0},$$

where $\mathbf{h}_c$ represents a vector of communes' characteristics that determine the assignment to classes, and $\boldsymbol{\theta}_q = \boldsymbol{\gamma}_q$. The coefficient vector of the first class, $\boldsymbol{\gamma}_1$, is normalized to zero for identification of the model.

Before going further, it is important to discuss some advantages and disadvantages of both methods. One feature of the continuous case is that we need to specify the distribution of the coefficients a priori. If the distribution for the unobserved spatial heterogeneity is not properly selected, then a problem of misspecification might arise. On the contrary, in the discrete case no assumptions about the shape of the spatial heterogeneity in terms of domain and boundedness are made. This frees the researcher from potential problems of misspecification (Shen, 2009; Hess et al., 2011). The only decision about the distribution is the number of classes—which is the same as the number support points. A disadvantage of the discrete case might be the rapid increase in the number of parameters. For example, in a model with 10 variables and 4 classes we will have $(10 \times 4 + 3 =)$ 43 parameters. Furthermore, it is often observed that some parameters collapse to the same value across classes or some classes obtaining very small probabilities, which is highly likely in the case of strongly peaked distributions (Hess, 2014). Another important difference between these two models is the estimation procedure. The continuous case requires the use of simulation-based methods, which can be very costly in terms of computational time, but no simulation is required for the discrete case. As I will discuss later, this is one of the main issued to be considered when comparing both models.

## 3.2 Estimation

Since the observed SWB measure is binary, the latent variable $W_{ic}^*$ is linked to the observed binary variable $W_{ic}$ by the following rule:[11]

[11] I have also run some basic specifications using the original ordered categorical variable. However, I did not find systematic differences between the binary and ordered probit model. These results are available upon request.

$$W_{ic} = \begin{cases} 1 & \text{if } W_{ic}^* > 0 \\ 0 & \text{if } W_{ic}^* \leq 0 \end{cases}.$$

Then, the probability for individual $i$ in commune $c$, given the latent structure in Equation (4) is:

$$\Pr(W_{ic}^* = W_{ic}|\mathbf{x}_{ic}, \boldsymbol{\beta}_c) = \left[\Phi\left(\mathbf{x}_{ic}'\boldsymbol{\beta}_c\right)\right]^{W_{ic}} \left[1 - \Phi\left(\mathbf{x}_{ic}'\boldsymbol{\beta}_c\right)\right]^{1-W_{ic}},$$

where $\Phi(\cdot)$ denotes the cumulative density function of the standard normal distribution. If there are $C$ different communes with $n_c$ individuals in each of them, and assuming that individuals are independent across communes, the joint probability density function, given $\boldsymbol{\beta}_c$ for commune $c$ is:

$$\Pr(\mathbf{w}_c|\mathbf{X}_c, \boldsymbol{\beta}_c, \upsilon_c) = \prod_{i=1}^{n_c} \Pr(W_{ic}^* = W_{ic}|\mathbf{x}_{ic}, \boldsymbol{\beta}_c), \tag{6}$$

where $\mathbf{w}_c$ is the sequence of choices for all individuals in commune $c$. Since $\boldsymbol{\beta}_c$ is common for individuals living in the commune $c$, within each commune, individuals are not independent. Thus, the unconditional pdf of $\mathbf{w}_c$ given $\mathbf{X}_c$ will be the weighted average of the conditional probability (6) evaluated over all possible values of $\boldsymbol{\beta}$, which depends on the parameters of the distribution of $\boldsymbol{\beta}_c$. For the discrete and continuous spatial heterogeneity, the unconditional pdf's are respectively

$$P_c(\boldsymbol{\theta}_q) = \sum_{q=1}^{Q} w_{cq} \left[\prod_{i=1}^{n_c} \Pr(W_{ic}^* = W_{ic}|\mathbf{x}_{ic}, \boldsymbol{\beta}_c, \boldsymbol{\theta}_q)\right], \tag{7}$$

$$P_c(\boldsymbol{\theta}) = \int_{\boldsymbol{\beta}_c} \left[\prod_{i=1}^{n_c} \Pr(W_{ic}^* = W_{ic}|\mathbf{x}_{ic}, \boldsymbol{\beta}_c, \boldsymbol{\theta})\right] g(\boldsymbol{\beta}_c) d\boldsymbol{\beta}_c. \tag{8}$$

The discrete spatial heterogeneity model with unconditional pdf given by (7) can be estimated using maximum likelihood approach. However, the unconditional pdf of the continuous case given by (8) has no closed form solution, so the log-likelihood function is hard to compute. To overcome this problem, we can simulate this probability and use the simulated maximum likelihood (SML) approach to estimate $\boldsymbol{\theta}$ (Gourieroux and Monfort, 1997; Hajivassiliou and Ruud, 1986; Stern, 1997; Train, 2009). In particular, $P_c(\boldsymbol{\theta})$ is approximated by a summation over randomly chosen values of $\boldsymbol{\beta}_c$. For a given value of the parameters $\boldsymbol{\theta}$, a value of $\boldsymbol{\beta}_c$ is drawn from its distribution. Using this draw of $\boldsymbol{\beta}_c$, $P_c(\boldsymbol{\theta})$ from Equation (8) is calculated. This process is repeated for many draws, and the average over the draws is the simulated probability. Formally, the simulated probability for commune $c$ is

$$\widetilde{P}_c(\boldsymbol{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\prod_{i=1}^{n_c}\widetilde{P}_{icr}(\boldsymbol{\theta})$$

where $\widetilde{P}_{icr}$ is the probability for individual $i$ in commune $c$ evaluated at the $r$th draw of $\boldsymbol{\beta}_c$, and $R$ is the total number of draws. Then, the simulated log-likelihood function is:

$$\log L_s = \sum_{c=1}^{C}\log\left[\frac{1}{R}\sum_{r=1}^{R}\prod_{i=1}^{n_c}\widetilde{P}_{icr}(\boldsymbol{\theta})\right]$$

Lee (1992), Gourieroux and Monfort (1991) and Hajivassiliou and Ruud (1986) derive the asymptotic distribution of the SML estimator based on smooth probability simulators with the number of draws increasing with sample size.[12]

## 3.3 Location-specific compensating variations

In order to compute commune-specific compensating variations we need the ratio between the coefficient of the variable of interest and the coefficient of income. However, some difficulties arise if we assume that both coefficients are random in the continuous case. To show this point, consider expanding the latent welfare function (4) as a linear form of income and variables $\mathbf{x}_{ci}$,

$$W^*_{ci} = \alpha_c \text{income} + \mathbf{x}'_{ci}\boldsymbol{\beta}_c + \epsilon_{ic}.$$

Then, according to Equation (3), CV for variable $k$ in commune $c$ is given by:

$$CV_{ck} = -\frac{\beta_{kc}}{\alpha_c}, \tag{9}$$

where $\beta_{kc} \sim g(\beta_k, \sigma_{\beta_k})$ is the coefficient of the variable of interest for individual $i$ in commune $c$, and $\alpha_c$ is the coefficient for income, which also varies across individuals according to $\alpha_c \sim h(\alpha, \sigma_\alpha)$.

The main shortcoming with the CV given in Equation (9), when both distributions are continuous, is that the ratio of two estimated distributions may not result in a well-specified distribution. For example, the ratio of two normal distributions produces a Cauchy distribution with no finite moments (see for example Daly et al., 2012, for other examples). In order to avoid this problem, I assume that the marginal utility of household income, $\alpha$, is fixed across communes when estimating a continuous spatial heterogeneous model.[13] This allows the distribution

[12] All the estimations conducted in this study are carried out using R software (R Core Team, 2016). In particular, I used the Rchoice package (Sarrias, 2016) to estimate the models using simulated maximum likelihood. The ML algorithm for the discrete case was also coded in the same software.

[13] In order to check this assumption empirically, I have also estimated a model assuming that

for the $k$th compensating variation to be the same as the distribution of the $k$th variable scaled by $\alpha$. For example, if $\beta_{ck} \sim \mathrm{N}(\beta_k, \sigma_k^2)$, then $(1/\alpha)\beta_{ck} \equiv \mathrm{N}(\beta_k/\alpha, \sigma_k^2/\alpha)$. It is important to highlight that this problem does not occur in the discrete case since both $\alpha$ and $\beta$ will vary across classes. This could be interpreted as another advantage of the discrete over the continuous case.

If we would like to know where each commune-specific compensating variation lie within the population heterogeneity distribution, we can move from the unconditional to the conditional distribution using Bayes' theorem (Revelt and Train, 1999). Formally, if $h(\boldsymbol{\beta}_c)$ is a function of $\boldsymbol{\beta}_c$ (as the CVs in Equation (9)), then:

$$f(h(\boldsymbol{\beta}_c)|\mathbf{w}_c, \mathbf{X}_c, \boldsymbol{\theta}) = \frac{f(\mathbf{w}_c|\mathbf{X}_c, \boldsymbol{\beta}_c)g(\boldsymbol{\beta}_c|\boldsymbol{\theta})}{\int_{\boldsymbol{\beta}_c} f(\mathbf{w}_c|\mathbf{X}_c, \boldsymbol{\beta}_c)g(\boldsymbol{\beta}_c|\boldsymbol{\theta})d\boldsymbol{\beta}_c},$$

where $f(h(\boldsymbol{\beta}_c)|\mathbf{w}_c, \mathbf{X}_c, \boldsymbol{\theta})$ is the distribution of $h(\boldsymbol{\beta}_c)$ conditional on the observed choices of individuals in commune $c$, and $g(\boldsymbol{\beta}_c|\boldsymbol{\theta})$ is the unconditional distribution. Thus, one can infer information about any function of $\boldsymbol{\beta}_c$ for each commune by conditioning on the individuals' observed relationship between SWB and the other covariates in that commune.

The conditional expectation of $h(\boldsymbol{\beta}_c)$ is thus given by:

$$\mathbb{E}\left[h(\boldsymbol{\beta}_c)|\mathbf{w}_c, \mathbf{X}_c, \boldsymbol{\theta}\right] = \frac{\int_{\boldsymbol{\beta}_c} h(\boldsymbol{\beta}_c) f(\mathbf{w}_c|\mathbf{X}_c, \boldsymbol{\beta}_c)g(\boldsymbol{\beta}_c|\boldsymbol{\theta})d\boldsymbol{\beta}_c}{\int_{\boldsymbol{\beta}_c} f(\mathbf{w}_c|\mathbf{X}_c, \boldsymbol{\beta}_c)g(\boldsymbol{\beta}_c|\boldsymbol{\theta})d\boldsymbol{\beta}_c}. \tag{10}$$

The expectation in Equation (10) gives us the conditional mean of the distribution of $h(\boldsymbol{\beta}_c)$, which can also be interpreted as the posterior distribution. Simulators for this conditional expectation are presented below for the continuous and discrete, respectively:

$$\widehat{\widehat{h}}(\widehat{\boldsymbol{\beta}}_c) = \widehat{\mathbb{E}}\left[h(\boldsymbol{\beta}_c)|\mathbf{w}_c, \mathbf{X}_c, \boldsymbol{\theta}\right] = \frac{\frac{1}{R}\sum_{r=1}^{R} h(\widehat{\boldsymbol{\beta}}_{cr}) \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cr}, \widehat{\boldsymbol{\theta}})}{\frac{1}{R}\sum_{r=1}^{R} \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cr}, \widehat{\boldsymbol{\theta}})} \tag{11}$$

$$\widehat{\widehat{h}}(\widehat{\boldsymbol{\beta}}_c) = \widehat{\mathbb{E}}\left[h(\boldsymbol{\beta}_c)|\mathbf{w}_c, \mathbf{X}_c, \boldsymbol{\theta}_q\right] = \frac{\sum_{q=1}^{Q} h(\widehat{\boldsymbol{\beta}}_q)\widehat{w}_{cq} \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)}{\sum_{q=1}^{Q} \widehat{w}_{cq} \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)} \tag{12}$$

It is important to highlight the role played by the individuals in the identification of the location-specific CV. Given that the location-specific estimates are conditional on the choices made by individuals in each commune, the asymptotic for consistency of Equations (11) and (12) requires $n_c \to \infty$; whereas $C \to \infty$ is only required for

individuals' income varies across space following a normal distribution. However the standard deviation for this variable turned out to be non-significant. Given this result, it seems that assuming any kind of parametric distribution, at least for this sample, for the income coefficient is not appropriate.

the consistency of $\widehat{\boldsymbol{\theta}}$. Thus, if the number of individuals rises without bound, then the conditional estimates for each commune should converge to the true commune's parameter (Revelt and Train, 1999; Train, 2009). That is:

$$\widehat{\bar{h}}_c \rightarrow h_c \quad \text{as} \quad n_c \rightarrow \infty.$$

The intuition behind this is that if we have more information about the individuals in each commune, we are in better position to say something about the compensating variation for individuals located in that commune.

In order to construct the confidence interval for $\bar{h}(\widehat{\boldsymbol{\beta}}_c)$, we can derive an estimator of the conditional variance from the point estimates as follows (Greene, 2012, chap. 15):

$$\widehat{V}_c = \widehat{\mathbb{E}}\left[h(\boldsymbol{\beta}_c)^2|\mathbf{y}_c, \mathbf{X}_c, \boldsymbol{\theta}\right] - \widehat{\mathbb{E}}\left[h(\boldsymbol{\beta}_c)|\mathbf{y}_c, \mathbf{X}_c, \boldsymbol{\theta}\right]^2. \tag{13}$$

An approximate normal-based 95% confidence interval can be then constructed as $\bar{h}(\widehat{\boldsymbol{\beta}}_c) \pm 1.96 \times \widehat{V}_c^{1/2}$.

# 4 Data and results

## 4.1 Data

I use data from the 2013 National Socioeconomic Characterization Survey (CASEN) from Chile. CASEN is a national population based survey, which is representative at the communal level and is carried out by the Ministry of Planning (MIDEPLAN) to describe the socioeconomic situation as well as the impact of social programs on the living conditions of the Chilean Population. Our sample in this study corresponds to 16,008 individuals between 15 and 65 years old in 324 communes after cleaning by missing values in the covariates.[14]

The dependent variable in this study is obtained from the response to the question "*How satisfied are you with life?*". The answer ranges from 1 "Completely unsatisfied" to 10 "Completely satisfied". As I reviewed in Section 2.1, this SWB question is considered as a better proxy of welfare, compared to other SWB questions. For sake of simplicity of interpretation and estimation, the response is recoded into a binary variable. Namely, the respondents are classified as being "Satisfied" if they replied 8, 9, or 10, and being "Unsatisfied or neither satisfied nor unsatisfied" if they

[14]One of the variables with more missing values is satisfaction with life. This is because the question is answered only for the household head. Table A1 shows the differences in mean between the individuals who answer this questions and those who do not. In general, the results are as expected. For example, the proportion of inactive individuals is higher in the sample that do not answer the SWB question (no households head), while older individuals are more likely to answer the question. No significant differences were observed in the variables about neighborhood perception.

replied with an integer between 1 and 7. Table 1, which reports summary statistics, shows that almost half of the sample report being satisfied with life.[15]

I control for several individual characteristics ($X$) commonly used in the literature of SWB. Table 1 shows that the average schooling is $\approx 10$ years, and males account for almost 33% of the total sample. Unemployed individuals account just for a small part (5%) of the sample, whereas individuals out of the labor force account for almost a 40%. In terms of marital status, 60% of the sample corresponds to married or cohabiting individuals. $Y$ is approximated by the household income per capita. The mean of household income per person is about $\exp(11.521) \approx \$164,226$ Chilean pesos.[16]

The component $H$ in Equation (1) is captured by three variables. *Disability* equals 1 if the respondent has some kind of long-term disability, such as blindness, deafness, and so on. The variable *treatment* indicates whether the individual is currently taking some medical treatment for the following illness: hypertension, diabetes, depression, heart attack or cancer. Finally, *accident* equals 1 if the individual have had an accident or sickness in the last 3 months. Table 1 shows that 6.7% of the sample has some long-term disability, while 30% is taking some type of treatment and 21% have had some kind of accident or sickness.

In addition, I use a set of dummy variables indicating the perception of neighborhood attributes ($Z$). The first set of indicators are related to self-perception of pollution and environmental problems. Specifically, the dummy variables are obtained from the response to the question: "*What problems related to pollution and environmental degradation do you identify in your neighborhood or location?*" Based on the answer, dummy variables were created for the following problems: 1) noise pollution, 2) air pollution, 3) water pollution, and 4) garbage in the neighborhood. The proportion of individuals reporting these problems are 16%, 17%, 0.7% and 21% respectively. Similarly, individuals were asked *"What problems related to public security do you identify in your neighborhood or location?"* For this case, the following dummy variables were created: 1) robberies and assaults on people, homes and/or vehicles (*Robbery*), 2) drug traffic (*Drug traffic*), 3) outbreaks of alcohol or drugs on the street (*Street drugs & alcohol*), and 4) insufficient police surveillance (*Poor surveillance*).[17]

[15]The estimations were also carried out using other dichotomizations, however no significant differences were observed. The results are available upon request.

[16]This is equivalent to $235 USD.

[17]It is possible that self-perception of neighborhood characteristics may suffer of measurement error, which may bias the estimates. However, the empirical literature has shown that subjective variables related to neighborhood characteristics might in some cases perform better than the objective ones (Day, 2007). For example, Van Praag and Baarsma (2005) found that crucial information regarding to noise pollution is not the objective, but the subjective perception of it. The argument is that when measuring the effect of location-specific amenities on subjective well-

Finally, the following controls at the commune level are used: logarithm of median income as a proxy for commune's development, logarithm of density and a dummy variable indicating if the individual is currently living in a urban area.

## 4.2 Fixed coefficients

In this section I comment on the results of the model with fixed parameters (the benchmark model), which is presented in the first three columns of Table 2. The first column gives the estimated coefficients, while the second column and third column display the standard errors and the compensating variation, respectively. While the magnitude of the estimated coefficients does not have a direct interpretation for binary Probit models, their sign indicates the direction of the relationship.

As expected, the results indicate a significant positive relationship between household income pc[18] and well-being; in other words, people with higher household income are more likely to report higher subjective well-being. Similarly, more educated individuals are more satisfied with their life. It can also be observed that men have, on average, higher levels of life satisfaction than women, whereas unemployed, and older people are more likely to be less satisfied with their lives. In line with the previous literature (see for example Clark and Oswald, 2002; Frey et al., 2010; Bruni, 2007), I also found that being out of labor force, living with a partner and bigger household size also increases the probability of reporting higher level of well-being. All the variables related to health status have the expected negative sign, and are highly significant: having an accident, some kind of disability or medical treatment increases the probability of reporting lower levels life satisfaction. Table 2 also shows that perception of noise pollution correlates negatively (strongly) with life satisfaction. I do not find sufficient evidence to reject that the rest of the variables measuring perceived pollution in the neighborhood (air, water and garbage pollution) are not correlated with SWB. The results also show a negative and statistically significant relationship between neighborhood problems and individual life satisfaction. For example, robbery, poor surveillance, street alcohol and drug trafficking problems in the neighborhood are negative and statistically different from zero.

The only variable at the commune level that enters significantly is urban. The estimated coefficient implies that individuals living in urban areas are on average more satisfied with their life than those living in rural areas. In general, the sign of the coefficients are consistent with the applied literature.

It is important to highlight that these estimates should not be overemphasized since they require the strong assumption of interpersonal cardinality. Having said

being it is important that they are linked to the individual at the level at which she/he experience them.

[18]Due to frequent use, 'per capita' is abbreviated as 'pc'.

that, it is more appropriate and conservative to analyze compensatory variations.

Compensating variations for the spatially stationary model are presented in column 3. The figures are computed as minus the ratio between the coefficient for the variable in interest and the coefficient for the logarithm of household income per capita taken from column 1 in Table 2.[19] Note that when changes in the variable imply a worsening (improvement) of well-being, then individuals require an increase (decrease) in household income pc to compensate for this loss (gain), and we should observe a positive (negative) sign for the ratio. For example, an individual with perceived noise pollution in the neighborhood would require, in average, a relative increase of 0.321 (or an increase of 32.1%) in his household income pc in order to remain in the respective satisfaction level with the externality.[20] An important observation is that a disability situation is the detrimental effect that requires the highest compensation: a disabled person requires a relative increase of income of about .95. This is followed by unemployment with a relative increase of .92. This last result is consistent with the finding from previous studies on the non-pecuniary cost of unemployment on an individual's subjective well-being (see Winkelmann and Winkelmann, 1998; Clark and Oswald, 2002; Blanchflower and Oswald, 2004). For example, Blanchflower and Oswald (2004)'s results suggest that to compensate men for unemployment would require a rise in income of \$60,000 USD per annum. Among the perceived neighborhood problems, drug trafficking is the situation that requires the highest compensation. For example, on average, an individual living in a neighborhood with drug trafficking would require an increase of about 30% of the household income pc in order to be unaffected by this problem. How much do people value living in urban areas? The results show that moving from urban to rural communes would require, in average, a relative increase of $\Delta Y/Y_0 = .33$. Strictly speaking, an individual living in a rural area with income $Y_0 + \Delta Y/Y_0$ would have the same level of well-being as an individual living in a urban area with income $Y_0$. Thus, $\Delta Y/Y_0$ can be considered as the average shadow price of urbanism.

Although the main objective of this study is to highlight the potential spatial heterogeneity that may be present when evaluating compensatory variations, it is important to emphasize the potential endogeneity of some of the variables used in this study. In particular, endogeneity problems may arise if individuals with higher

[19] The standard errors are computed using the Delta Method.

[20] Let $W^* = W(Z, \log(Y))$, where $Y$ is household income pc. By total differentiation, setting $W^* = 0$ and holding all the other variables constant yields

$$\frac{d\log(Y)}{dZ} = \frac{\partial W/\partial Z}{\partial W/\partial \log(Y)}.$$

Note that $d\log(Y)/dZ = (dY/Y)/dZ$. Then, the ratio measures the 'relative change' in $Y$ needed to bring the individual to his original level of well-being given a marginal change in $Z$, other things equal.

satisfaction with life are more productive and hence earn higher wages (reverse causality) or if individuals with stronger preferences for neighborhood amenities sort themselves into more amenable places (household sorting). Regardless of the reason, both income and environmental/neighborhood amenities may suffer of endogeneity producing biased CVs. However, the direction of the bias of CVs is not clear. Separate studies using IV estimates indicates that both the income (Powdthavee, 2010a; Luttmer, 2005) and pollution (Luechinger, 2010) coefficients should be larger. Therefore, it is not clear whether CV are overstated or understated. This will depend on which coefficient is more biased. For example, consider our fixed estimates for noise pollution ($0.076/0.236 = 0.322$): if noise pollution's coefficient is more biased than income's coefficient, say that the true coefficients are 0.1 and 0.280, then the CV will be downward biased (True CV = 0.357); whereas if noise pollution's coefficient is less biased than income's coefficient, say that the true coefficients are 0.09 and 0.35, then the CV will be upward biased (True CV = 0.285). In light of the foregoing, the results of this study should be considered with caution. However, one can think of this study as adopting very strong identification assumption, but calling attention to the importance of spatial heterogeneity when analyzing compensating variations.

## 4.3 Continuous spatial heterogeneity

The coefficients in the previous section reflect the average marginal impact of the variables and their respective CVs for some average individual in the sample. However, they obscure the potential latent heterogeneity across communes. The second model of Table 2 shows the results for a model where all the parameters, except for the logarithm of household income pc, are assumed to be independently normally distributed across space. That is, I assume that the coefficients vary randomly across communes and therefore they can take positive or negative values.[21] As explained before, the household income pc is held fixed to prevent the non-existence of moments for the ratio when computing the CVs. The coefficient vector is expressed as $\boldsymbol{\beta}_c = \boldsymbol{\beta} + \mathbf{L}\boldsymbol{\omega}_c$, where $\mathbf{L}$ is a diagonal matrix whose elements are standard deviations and $\boldsymbol{\omega} \sim \mathrm{N}(0,1)$. For the simulation procedure 100 Halton draws were used for each commune and parameter in each specification.[22] For each parameter, the mean and standard deviation of the distribution is estimated. Subsequent models

[21] I have also tried other distributions in the specification such as the triangular, uniform and Johnson $\mathrm{S}_b$ distribution, and other different combinations. However, models showed comparative lower fit than the models presented in Table 2.

[22] Good performance of SML requires a very large number of draws. However, the maximization of SML can be very time consuming when estimating large and complex models. Researchers have gained speed with no degradation in simulation performance through the use of smaller number of Halton draws (Bhat, 2001; Train, 2000). Bhat (2001)'s Monte Carlo analysis found that the precision of the estimated parameters was smaller using 100 Halton draws than 1000 pseudo-random number in the context of Mixed Logit. In this study, I found that beyond 100 Halton draws did not lead to significant changes in the estimated parameters.

allow correlation among the coefficients and specify log-normal and truncated normal distribution for some of the coefficients.

The results are broadly robust. Both the mean coefficients and the CVs are very close in magnitude, vis-a-vis, to those for the benchmark (stationary) model. The standard deviations for some of the variables are highly significant, indicating that those parameters do indeed vary across communes, and focusing on the central tendency alone veils useful information. The LR statistic comparing model 1 and 2 is $\chi^2_{22} = 139.04$, therefore we reject the null hypothesis that the standard deviations are jointly zero.

The parameters with significant standard deviation—which measure the degree of spatial heterogeneity—are age, male, inactive, married, medical treatment, accident, air pollution, robbery, poor surveillance, log of communal median income and urban. Since the parameters are normally distributed we can get an approximate proportion of communes with positive and negative coefficients (see column 7).[23] For example, the mean and standard deviation of the schooling parameter imply that almost 100% of the communes have a positive correlation between schooling and well-being. Furthermore, the standard deviation of this parameter is not significant, implying that the coefficient does not vary across space, but rather it is fixed and unique for all the communes. We explore this option later. The standard deviation of the commune-specific unobserved factor $\upsilon$, which is assumed to be normally distributed is significant. This unobserved factor is intended to capture all those effects not captured by the variables at the commune level included in the model.

In the previous model, the coefficients are specified to be independently distributed. However, we might expect some degree of correlation between the coefficients. To allow this feature, I specify $\boldsymbol{\beta}_c = \mathrm{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma})$ for a general $\boldsymbol{\Sigma}$. The matrix $\mathbf{L}$ is now a lower-triangular Chokeski factor of $\boldsymbol{\Sigma}$, such that $\mathbf{LL}' = \boldsymbol{\Sigma}$. In extensive form, the coefficient vector can be expressed as:

$$\begin{pmatrix} \beta_{1r} \\ \beta_{2r} \\ \vdots \\ \beta_{Kr} \end{pmatrix} = \begin{pmatrix} \beta_{1} \\ \beta_{2} \\ \vdots \\ \beta_{K} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \dots & a_{KK} \end{pmatrix} \begin{pmatrix} \omega_{1r} \\ \omega_{2r} \\ \vdots \\ \omega_{Kr} \end{pmatrix}.$$

Thus, each single coefficient can be written as:

$$\beta_{kr} = \beta_k + \sum_{h=1}^{K} a_{kh}\omega_{hr}.$$

The third model in Table 2 presents the results for the correlated spatially random

[23] The proportion of communes with a positive coefficient for some attribute is given by $\Phi(\widehat{\beta}/\widehat{\sigma}_\beta)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

parameter model. The standard deviations of the parameters are computed by diag $\left(\widehat{\boldsymbol{\Sigma}}^{1/2}\right)$, and their standard errors are computed using Delta Method. In terms of relative quality of the correlated parameter model, it can be observed that both measures of information–AIC and BIC–are higher compared with the model where all parameters are assumed to be independently normally distributed. In other words, the model that assumes independent spatial heterogeneous parameters is favored by both criteria. An important result is that all the standard deviations —except for male—are highly significant.

One important limitation of the previous models is that the communes' coefficients can take positive and negative values, whether this is true or not. For example, according to the results for the correlated parameter model for 20% of the communes increasing individuals' education reduces satisfaction with life. However, it is not clear whether this is in fact true or is simply an artifact of the model specification (normal assumption) or data quality. Similarly, for some variables such as health problems, or problems in the neighborhood, it is reasonable to expect a negative sign for all the communes—but different intensities. The fourth model in Table 2 abandons this possibility for some of the variables. In particular, I assume that schooling, unemployed, disability, medical treatment, accident, robbery, poor surveillance, street drugs, and drug trafficking follow a log-normal distribution. In this case the coefficient can be written as $\beta_c = \exp(\beta + \sigma_\beta \eta_c)$ where $\eta_c \sim \mathrm{N}(0, 1)$. The population parameters $\beta$ and $\sigma_\beta$, which represent the mean and standard deviation of $\log(\beta_c)$, are estimated. The mean and standard deviation of $\beta_c$ are $\exp(\beta + \sigma^2_\beta/2)$ and mean $\times \sqrt{\exp(\sigma^2_\beta) - 1}$, respectively.

Note that the log-normal distribution has its domain in the positive real line $(0, \infty)$. For variables where a negative sign is expected I include the negative of the variable. This allows the coefficients to be negative without imposing a sign change in the estimation procedure.

For the pollution variables, I assume a truncated normal distribution, whose domain is also $(0, \infty)$. The parameter for each region is created as $\beta_c = \max(0, \beta + \sigma_\beta \eta_c)$ where $\eta_c \sim \mathrm{N}(0, 1)$. The share of communes massed at zero—i.e. share of communes for which the relationship is assumed to be null—is equal to $\Phi(-\beta/\sigma_\beta)$.

One result worth mentioning for the last model in Table 2 is that the mean parameters of the variables related to problems in the neighborhood are now statistically significant. In terms of commune-level variables, we can see that the mean of density is negative and statistically significant. Finally, both AIC and BIC indicate that the fit is not good enough compared to previous models.

## 4.4 Discrete Spatial Heterogeneity

Table 3 presents two models that assume that coefficients follow a discrete distribution across space. The first model assumes that there exist 2 classes, whereas the second assumes three classes. In both models, I assume that the probability of belonging to some class depends on the geographic location of each commune represented by the geographical coordinates of the centroids $x$ and $y$, which represent the longitude (higher values implies a movement from west to east) and latitude (higher values implies a movement from north to south) respectively.[24]

A first important result is that both models present a better statistical fit compared to all the previous models, while the model with 2 classes presents a weak superiority over the model with 3 classes by presenting a lower BIC. In other words, the model with 3 classes is not able to retrieve significant amount of additional spatial heterogeneity when compared with the model with 2 classes. Given this, I will comment on the results of the model with 2 classes.

To give a more intuitive interpretation of the coefficients for the class-assignment, panel A and B of Figure 2 show the geographical distribution of the conditional (posterior) probability of belonging to each class across communes.[25] For example, panel A shows that communes located further south and west are more likely to belong to the first class, whereas panel B shows that communes located further north and east are more likely to belong to the second class. Thus, including geographical coordinates allows the identification of classes based on the spatial proximity of the communes.

In general, the following can be noted from the model:

- Individual characteristics ($X_i$): Education is more valued in communes belonging to the second class. For example, a reduction of one year of education requires an increase of household income pc of approximately 15% (in average) compared to 1% of the first class, ceteris paribus. Both are statistically significant. Age is also valued negatively in both classes, however the compensation is higher for the first class (2.7% vs. 1.5%). While gender differences seem to be not important for the first class, being female in communes belonging to class 2 would require an increase in household income

[24] I have also estimated versions of the continuous model where the mean of the random parameters varies according to functions of the geographical coordinates. However, due to the complexity of the estimation procedure, the models in some cases did not converge or I encountered flat regions of the simulated likelihood function producing a singular Hessian.

[25] Specifically, the conditional probability of commune $c$ of belonging to class $q$ is computed as:

$$\widehat{\pi}_{cq} = \frac{\widehat{w}_{cq}\prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)}{\sum_{q=1}^{Q}\widehat{w}_{cq}\prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)}.$$

pc of 35% to offset gender differences in well-being. Being unemployed is detrimental in both classes. However, the CV is higher in the second class. For example, individuals living in communes in the first class require an increase of their household income pc of about 88% in order to be as well-off as someone who is not unemployed compare with the 80% of the second class. Household size has similar CV in both classes, whereas being married is more valued in the second class.

- Health conditions ($H_i$): In terms of health conditions, disability is the detrimental effect that requires the highest monetary compensation among the three health conditions in both classes (68% in class 1 vs. 100% in class 2).[26] Having some kind of medical treatment is significant in both classes, however individuals in communes of class 2 are more sensitive requiring a higher compensation compare to class 1 (33% vs. 91%). Finally, having an accident is, in average, only detrimental for communes in the second class requiring an increase of household income pc of about 42% in order to keep the same well-being as someone without an accident. Since the second class is mostly representative of communes located in north (see panel B of Figure 2), these results could reflect that the spatial configuration in terms of amount (number of health services) and quality (spread of diseases) is lower in these communes.

- Neighborhood characteristics ($Z_i$): In terms of perceived pollution in the neighborhood, water and trash pollution are not important. Those results are in line with those for the continuous spatial heterogeneity assuming correlated parameters and lognormal and truncated normal distribution. The detrimental effect of noise pollution is negative and significant for the second class, with a monetary compensation of about 69% increased in household income. It can be noticed that dis-amenities in the neighborhood such as robbery, poor surveillance and street drug problems are only important for communes in the first class: individuals living in those communes need an increase of about 73%, 30% and 37% in household income pc respectively. There are two variables that present some counterintuitive results for class 1: both air pollution and drugs trafficking present a positive and significant coefficient (and hence a negative CV). This could be reflecting an omitted-variable problem or a different error variance for class 1.

[26] To put in context the compensatory variation for disability for the second class, it can be considered that an average Chilean family would require an increase of household income pc of approximately $900 USD. This value is not unrealistic if one considers valuations made for similar long-term health problems. For example, Groot and Maassen van den Brink (2006) estimate that the monetary value for cardiovascular diseases is approximately £49564 for men and £17503 for women.

- Variables at the commune level: Density reduces subjective well-being of individuals in class 1: an increase on 1% on density would require an increase of 0.14% of household income to keep the individuals at the same level of subjective well-being. Log of median income at the commune level (our proxy for development) presents different results for both classes. For the first class, the log of median income presents a positive correlation with individuals' subjective well-being, whereas in the second class the correlation is negative. In other words, individuals living in communes of the first class enjoy better development so that they should be compensated by an increase of about 68% of household income per capita if median income is reduced by 1%. Oppositely, inhabitants of communes belonging to class 2 dislike higher levels of development; therefore they should be compensated with an increase of 59% of household income if median income increases by 1%. Finally, urbanism matters only in the second class. The benefit of moving from rural to urban communes is worth around a relative increase of 0.6 in household income for the second class.

## 4.5 Spatial Heterogeneity in CV

In this Section I analyze the distribution of the conditional mean of compensating variations across communes estimated using Equations (11) and (12). Since the number of estimated parameters is large, it is almost impossible to comment in detail each coefficient. Given this restriction and the purpose of this study, I discuss some general aspects of the results, and then I will focus in CVs for certain variables.[27]

Figure 3 show the distribution (boxplot) of the conditional mean of CVs for some selected coefficients estimated using continuous and discrete heterogeneity.[28] In particular, I plotted the distribution for the following models:

- C1: Continuous model with all parameters as normally distributed,
- C2: Continuous model with correlated random parameters (LN and CN),
- D1: Discrete model with $Q = 2$,
- D2: Discrete model with $Q = 3$.

The idea behind this exercise is to graphically detect differences in the CVs using both methods. The optimal scenario would be if we could observe similar distributional patterns of the compensating variations. This would indicate that the

[27] In cases where the research question is more direct, one could focus on particular communes if there are reasons to do so or describe the distribution of specific quantiles of the location-specific CVs as shown in Daziano and Achtnicht (2014).

[28] The conditional mean of CVs for the rest of variables are shown in Figure A1.

results are robust to any kind of spatial heterogeneity and assumption about the underlying distribution of the CVs. As we will see, this is not the case for some variables in this particular sample.

A first important result can be observed by taking a quick look to the distributions. Some communes show positive and negative compensating variations when the parameters are assumed to be normally distributed. This result can be true or just an artifact of the normal distribution. For some of the variables, the latter seems to be true. For example, it is difficult to argue against the fact that having and accident has a detrimental effect on well-being, and therefore the compensatory variation in any geographical location must be positive. However, panel D shows that when the parameter is assumed to be normal distributed, there exists some communes with negative CV. The results are as (theoretically) expected when a discrete distribution with 2 classes is assumed or when it is assumed that the parameter is continuously distributed according to a log-normal distribution. For other variables, such as schooling, unemployment and disability, the results are as expected and robust to any kind of distribution.

In terms of variability, it can be observed that, in general, the log-normal distribution shows greater variability. This is not unsurprising since the log-normal is characterized by having very long tail resulting in implausible compensating variations for some communes. For example, considering the compensating variation for employment, the log-normal distribution determines that there exists a commune that requires a relative increase of income of about 6.5.

Another important aspect is whether the location-specific CVs are significant. For illustration purposes, Figure 4 shows the compensating variations for disability for each commune along with the 95% confidence intervals (CI). The blue dashed lines correspond to CV and 95% confidence interval computed assuming fixed coefficients. The standard errors used to construct the CIs were computed using Equation (13). When looking at the continuous case, it can be observed that the compensating variations for some communes are not statistically significant since they include zero. In fact, 1% and 72% of the communes have non-significant CV, when assuming the normal and log-normal distribution respectively. The lack of conditional estimates with significant values when using the log-normal distribution has been also documented by Daziano and Achtnicht (2014). For the discrete case we observed that none of the CIs includes zero, implying that individuals in all the communes should be compensated for the detrimental effect of disability. In general, the communes' estimates are close to the CV computed using the benchmark model. However, the fact that the different communes have a value close to the model with fixed parameters does not imply that the preferences are actually spatially stationary. This only reveals that the heterogeneity is low compared to the average.

Since non-pecuniary cost of unemployment on individual's subjective well-being is a very important topic on the SWB literature (Winkelmann and Winkelmann, 1998), Figure 5 plots the commune-specific estimates for this variable. The patter is closely similar to Figure 4. Again, confidence intervals using the log-normal distribution are quite wide. In effect, only 1% of the communes have significant CV. When imposing the normal distribution, the proportion is 17%. In the model with two classes all the communes have a significant CV respect to unemployment; whereas with three classes the proportion of communes with significant CV is roughly 82%. In terms of the magnitude of the CVs, the discrete cases show a lower range than the continuous specifications. This is also corroborated by Figure 3.

The lack of significance of location-specific CV under the continuous approach is striking, however it can be explained by considering the complexity of the estimation procedure. The continuous approach, unlike the discrete approach, requires the simulations of the probability due to no close form solution. Therefore, in addition to the small-sample bias, the SML adds another layer of bias due to the simulation. In a Monte Carlo study, Sarrias (2015) analyzes the ability of the conditional mean of the distribution of the spatially random parameters to retrieve the true spatial representation of the parameters under both methods. One of the main important results is that the continuous case requires more spatial units for the SML estimates to achieve acceptable levels of bias when compared with the ML estimates used in the discrete approach. In other words, given a small sample, the estimates from the discrete approach are more reliable than those from continuous approach.

Another important issue is the number of parameters estimated under the more complex model. When assuming correlated parameters, the continuous model estimate 276 parameters compare with 73 parameters in the model with three classes. This might also have an impact on the variance of the estimated parameters.

Finally, I have applied a test for nonnested models proposed by Ben-Akiva and Swait (1986) to discriminate between the discrete and continuous specification. The intuition behind this test is as follows. Suppose there are two nonnested models $A$ and $B$. Model $A$ explains choices using $K_A$ coefficients, while model $B$ explains the same data using $K_A$ coefficients. Assume that the two models have different functional forms. Define the fitness measure for model $j, j = A, B$:

$$\rho_j^2 = 1 - \frac{L_j - K_j}{L(0)},$$

where $L_j$ is the log-likelihood at convergence and $L(0)$ is the log-likelihood for constants only. Ben-Akiva and Swait (1986) show that under the null hypothesis that Model $A$ is the true model, the probability that the fitness measure for model $B$ will be greater than of model $A$ is asymptotically bounded by:

$$\Pr\left(\left|\rho_B^2-\rho_A^2\right|\geq z\right)\leq\Phi\left(-\left[-2zL(0)+(K_B-K_A)\right]^{1/2}\right)$$

where $z=\rho_B^2-\rho_A^2$. Therefore, this equation sets an upper bound for the probability of rejecting the null hypothesis that Model $A$ is the correct model. Using this procedure, I assume that the true models are those using using continuous distribution (those estimated in columns 2, 3 and 4 in Table 2) and Model $B$ is the model with discrete distribution using 2 classes. The results are the following:

- $H_0$: All normals, $H_1$: Latent class $Q=2$: $p=0.002$.
- $H_0$: Correlated, all normals, $H_1$: Latent class $Q=2$: $p=0.000$.
- $H_0$: Correlated with transformations of normals, $H_1$: Latent class $Q=2$: $p=0.000$.

In all cases, the continuous distribution is rejected. Then, based on the goodness of fit measures (AIC and BIC) and the results from the Ben-Akiva and Swait (1986)'s test, there exists evidence that the discrete approach fits better the data than the continuous case .

# 5 Discussion and Conclusion

Using SWB equations to estimate compensating variations has become a very simple but powerful tool for policy makers, especially for evaluating intangible goods that have no market price. Many of the theoretical and applied studies have helped us to expand our understanding of how people value different health conditions, environmental externalities, and their social capital. This study goes a step further and tries to add a new dimension to the debate by calling attention to the importance of spatial heterogeneity when analyzing compensating variations.

To this end, I compare two different models, continuous and discrete, to incorporate unobserved spatial heterogeneity into the estimation of CV. Both approaches assume that the coefficients (and therefore compensatory changes) follow a distribution which is unknown to the investigator. In the first case, the spatial unobserved heterogeneity of coefficients is determined by a continuous distribution; therefore parameters may take any value within the domain of the chosen distribution. In the second case, spatial units (communes in this study) belong to different groups, and each group has the characteristic of having the same sensitivities to certain variables. Thus, there will be as many parameters for the same variable as classes of regions. In this study, the probability of belonging to a particular class is determined by the geographical location captured by the geographical coordinates.

Important findings emerged from this analysis applied to the Chilean case. The results support the evidence of substantial heterogeneity in the CV for several variables. For example, individuals perceiving poor surveillance in the neighborhood need in average a relative increase of household income pc of about 0.225 to be as well-off as someone who does not perceive this problem (figure based on the fixed model). However, when I allow for different compensating variations for each spatial unit I found that there are some communes where individuals should be compensated for much less, while for others communes more compensation is required. In other words, a similar negative characteristic of the neighborhood has varying compensating variation across communes. A non-stationary model would have disguised this remarkable geographical variation and led to broad generalizations ignoring these local differences in terms of CVs.

Although both approaches unveil spatial heterogeneity, I do find some differences on the results of both approaches. In this study, the discrete approach has a slight advantage over the continuous approach in terms of model fit: models with 2 and 3 classes present a better statistical fit compared to the non-stationary model and all the models with continuous spatial heterogeneity. This superiority of the discrete specification is also found in similar studies that compare both methods in the context of Mixed and Latent Class Logit models (See for example Greene and Hensher, 2013; Shen, 2009; Hess et al., 2011). Furthermore, although the medians for the CVs with both approaches are relatively close, I noticed more mass at the extremes of the CV distributions when a log-normal distribution is assumed. This is partially explained by the domain of this distribution. The main drawback of the log-normal distribution is that it has a very long right-hand tail, resulting in communes having implausible and extreme coefficients.

An important aspect of spatial analysis is whether CVs are significant for each location. Nevertheless, analyzing the evidence on CIs is challenging because of remarkably differences between both models. In general, more communes with significant CVs are found assuming a discrete rather than a continuous distribution. I speculate that this is due to the complexity of SML approach. In fact, Sarrias (2015) performs a Monte Carlo experiment in order to understand the ability of both approaches to retrieve the true representation of the spatially varying process under small sample size. He finds that the data requirement to achieve lower bias in the continuous case is substantial compared with the discrete case. He also finds that the precision to identify each locational-specific estimate improves as the number of individuals per region increases. However, the discrete case is able to retrieve the true spatial heterogeneity surface with lower bias and better coverage for small samples.

So, the key question is which one should be used? Overall, the results of this study suggest that, at least for the sample used, the discrete approach provides better statistical fit and easier interpretation. In addition, the appeal of the discrete approach is that the analyst does not need to specify a particular distribution for the unobserved spatial heterogeneity. As I reviewed, assuming some continuous distribution might lead to unreasonable coefficients. In other words, when the distributions of the parameters are less constrained may lead to significantly better results. Furthermore, the SML approach used in the continuous case is very costly in terms of computational times. For example, the more complex model (model 4 of Table 2) took 1 hour and 29 minutes to converge, whereas the model with three classes took about 15 minutes. Nonetheless, this evidence is not conclusive and it can be sample-specific. Therefore, the need for ongoing research on the comparison between these two specifications with different samples is fundamental.

Finally, this work is not without shortcomings. First, the SWB approach needs very strong assumptions to be valid to proxy for individuals' utility. However, the literature reviewed in Section 2.1 suggests that using satisfaction with life (instead of other measures of SWB) is a more valid and reliable measure for measuring true well-being. Second, the results should be interpreted cautiously because of the potential endogeneity of some variables. Given this, I confine myself to talk about spatial heterogeneous relationships instead of causation acknowledging this potential bias. Nevertheless, one can think of the study as adopting the same strong identification assumptions as previous studies, but exploring how we can push their findings further by exploring spatial heterogeneity. Third, the methods used in this study allow us to detect the apparent spatial variation of CVs, but they do not explain why CVs are different at any given location.

Despite these limitations, this study has important implications. I argue that when analyzing CVs explicit recognition of spatial heterogeneity is required. I have shown that the response of individuals' underlying utility to covariates is not fixed over space, but rather it varies among space in terms of intensity. From a policy perspective, an understanding of how certain geographical areas need more or less compensation can be helpful in designing effective geographical interventions.

# References


Adler, M. D. (2012). Happiness Surveys and Public Policy: What's the Use. *Duke LJ*, 62:1509.

Ali, K., Partridge, M. D., and Olfert, M. R. (2007). Can Geographically Weighted Regressions Improve Regional Analysis and Policy Making? *International Regional Science Review*, 30(3):300–329.

Bateman, I. J. and Willis, K. G. (2001). *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford University Press on Demand.

Ben-Akiva, M. and Swait, J. (1986). The Akaike Likelihood Ratio Index. *Transportation Science*, 20(2):133–136.

Benjamin, D., Heffetz, O., Kimball, M., and Rees-Jones, A. (2014). Can Marginal Rates of Substitution Be Inferred From Happiness Data? Evidence from Residency Choices. *The American economic review*, 104(11):3498–3528.

Benjamin, D. J., Kimball, M. S., Heffetz, O., and Rees-Jones, A. (2012). What Do You Think Would Make You Happier? What Do You Think You Would Choose? *The American economic review*, 102(5):2083.

Berry, B. J. and Okulicz-Kozaryn, A. (2011). An Urban-Rural Happiness Gradient. *Urban geography*, 32(6):871–883.

Bhat, C. R. (2001). Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B: Methodological*, 35(7):677–693.

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being Over Time in Britain and the USA. *Journal of public economics*, 88(7-8):1359–1386.

Bockstael, N. E. and Freeman, A. M. (2005). Welfare Theory and Valuation. In Mler, K.-G. and Vincent, J. R., editors, *Valuing Environmental Changes*, volume 2 of *Handbook of Environmental Economics*, pages 517 – 570. Elsevier.

Brereton, F., Clinch, J. P., and Ferreira, S. (2008). Happiness, Geography and the Environment. *Ecological Economics*, 65(2):386–396.

Brouwer, R., Martin-Ortega, J., and Berbel, J. (2010). Spatial Preference Heterogeneity: A Choice Experiment. *Land Economics*, 86(3):552–568.

Bruni, L. (2007). *Handbook on the Economics of Happiness*. Edward Elgar Publishing.

Bruni, L. and Porta, P. L. (2005). *Economics and Happiness: Framing the Analysis*. Oxford University Press.

Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443.

Campbell, D., Hutchinson, W. G., and Scarpa, R. (2009). Using Choice Experiments to Explore the Spatial Distribution of Willingness to Pay for Rural Landscape Improvements. *Environment and Planning A*, 41(1):97–111.

Clark, A. E. and Oswald, A. J. (2002). A Simple Statistical Method for Measuring How Life Events Affect Happiness. *International Journal of Epidemiology*, 31(6):1139–1144.

Daly, A., Hess, S., and Train, K. (2012). Assuring Finite Moments for Willingness to Pay in Random Coefficient Models. *Transportation*, 39(1):19–31.

Day, R. (2007). Place and the Experience of Air Quality. *Health & Place*, 13(1):249–260.

Daziano, R. A. and Achtnicht, M. (2014). Accounting for Uncertainty in Willingness to Pay for Environmental Benefits. *Energy Economics*, 44:166–177.

Deiner, E., Suh, E., Lucas, R. E., and Smith, H. L. (1999). Subjective Well-being: Three Decades of Progress. *Psychological Bulletin*, 125(2):276–302.

Di Tella, R. and MacCulloch, R. (2006). Some Uses of Happiness Data in Economics. *The Journal of Economic Perspectives*, pages 25–46.

Diener, E., Suh, E. M., Smith, H., and Shao, L. (1995). National Differences in Reported Subjective Well-being: Why Do They Occur? *Social Indicators Research*, 34(1):7–32.

DAcci, L. (2014). Monetary, Subjective and Quantitative Aproaches to Assess Urban Quality of Life and Pleasantness in Cities (Hedonic Price, Willingness-to-Pay, Positional Value, Life Satisfaction, Isobenefit Lines). *Social Indicators Research*, 115(2):531–559.

Eid, M. and Larsen, R. J. (2008). *The Science of Subjective Well-Being*. Guilford Press.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal*, 114(497):641–659.

Ferrer-i Carbonell, A. and van Praag, B. (2002). The Subjective Costs of Health Losses due to Chronic Diseases. An Alternative Model for Monetary Appraisal. *Health economics*, 11(8):709.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2003). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons.

Franceschinis, C., Scarpa, R., Thiene, M., Rose, J., Moretto, M., and Cavalli, R. (2016). Exploring the Spatial Heterogeneity of Individual Preferences for Ambient Heating Systems. *Energies*, 9(6):407.

Frey, B. S. (2008). Happiness: A Revolution in Economics. *MIT Press Books*, 1.

Frey, B. S., Luechinger, S., and Stutzer, A. (2010). The Life Satisfaction Approach to Environmental Valuation. *The Annual Review of Resource Economics*, 2:139–60.

Frey, B. S. and Stutzer, A. (2002). What Can Economists Learn from Happiness Research? *Journal of Economic literature*, pages 402–435.

Glaeser, E. L., Gottlieb, J. D., and Ziv, O. (2016). Unhappy Cities. *Journal of Labor Economics*, 34(S2):S129–S182.

Goetzke, F. and Islam, S. (2016). Testing for Spatial Equilibrium Using Happiness Data. *Journal of Regional Science*.

Gourieroux, C. and Monfort, A. (1991). Simulation Based Inference in Models with Heterogeneity. *Annales d'Economie et de Statistique*, (20-21):69–107.

Gourieroux, C. and Monfort, A. (1997). *Simulation-based Econometric Methods*. Oxford University Press.

Greene, W. H. (2012). *Econometric Analysis*. Prentice Hall, 7 edition.

Greene, W. H. and Hensher, D. A. (2013). Revealing Additional Dimensions of Preference Heterogeneity in a Latent Class Mixed Multinomial Logit Model. *Applied Economics*, 45(14):1897–1902.

Groot, W. and Maassen van den Brink, H. (2006). The Compensating Income Variation of Cardiovascular Disease. *Health Economics*, 15(10):1143–1148.

Hajivassiliou, V. A. and Ruud, P. A. (1986). Classical Estimation Methods for LDV Models Using Simulation. *Handbook of Econometrics*, 4:2383–2441.

Hensher, D. A. and Greene, W. H. (2003). The Mixed Logit Model: The State of Practice. *Transportation*, 30(2):133–176.

Hess, S. (2014). Latent Class Structures: Taste Heterogeneity and Beyond. *Handbook of Choice Modelling*, page 311.

Hess, S., Ben-Akiva, M., Gopinath, D., and Walker, J. (2011). Advantages of Latent Class Over Continuous Mixture of Logit Models. *Institute for Transport Studies, University of Leeds. Working paper.*

Lanz, B. and Provins, A. (2013). Valuing Local Environmental Amenity with Discrete Choice Experiments: Spatial Scope Sensitivity and Heterogeneous Marginal Utility of Income. *Environmental and Resource Economics*, 56(1):105–130.

Lee, L.-F. (1992). On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Response Models. *Econometric Theory*, 8(04):518–552.

Levinson, A. (2012). Valuing Public Goods Using Happiness Data: The Case of Air Quality. *Journal of Public Economics*, 96(9):869–880.

Lin, C.-H. A., Lahiri, S., and Hsu, C.-P. (2014). Happiness and Regional Segmentation: Does Space Matter? *Journal of Happiness Studies*, 15(1):57–83.

Luechinger, S. (2010). Life Satisfaction and Transboundary Air Pollution. *Economics Letters*, 107(1):4–6.

Luttmer, E. F. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly journal of economics*, 120(3):963–1002.

Moore, S. and Shepherd, J. P. (2006). The Cost of Fear: Shadow Pricing the Intangible Costs of Crime. *Applied Economics*, 38(3):293–300.

Moro, M., Brereton, F., Ferreira, S., and Clinch, J. P. (2008). Ranking Quality of Life Using Subjective Well-being Data. *Ecological Economics*, 65(3):448–460.

Perez-Truglia, R. (2015). A Samuelsonian Validation Test for Happiness Data. *Journal of Economic Psychology*, 49:74–83.

Powdthavee, N. (2008). Putting a Price Tag on Friends, Relatives, and Neighbours: Using Surveys of Life Satisfaction to Value Social Relationships. *The Journal of Socio-Economics*, 37(4):1459–1480.

Powdthavee, N. (2010a). How Much Does Money Really Matter? Estimating the Causal Effects of Income on Happiness. *Empirical Economics*, 39(1):77–92.

Powdthavee, N. (2010b). *The Happiness Equation: The Surprising Economics of our Most Valuable Asset*. Icon Books.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rehdanz, K. and Maddison, D. (2008). Local Environmental Quality and Life-Satisfaction in Germany. *Ecological economics*, 64(4):787–797.

Revelt, D. and Train, K. (1999). Customer-Specific Taste Parameters and Mixed Logit. *University of California, Berkeley*.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of political economy*, 90(6):1257–1278.

Sarrias, M. (2015). Continuous and Discrete Spatial Heterogeneity: Modeling Strategies and Simulation. Technical report.

Sarrias, M. (2016). Discrete Choice Models with Random Parameters in R: The Rchoice Package. *Journal of Statistical Software, Articles*, 74(10):1–31.

Shen, J. (2009). Latent Class Model or Mixed Logit Model? A Comparison by Transport Mode Choice Data. *Applied Economics*, 41(22):2915–2924.

Stern, S. (1997). Simulation-Based Estimation. *Journal of economic Literature*, 35(4):2006–2039.

Swamy, P. (1971). *Statistical Inference in Random Coefficient Regression Models*. Number 55. Springer Berlin.

Train, K. (2000). Halton Sequences for Mixed Logit. *Department of Economics, UCB*.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge university press.

Van Praag, B. and Baarsma, B. E. (2005). Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. *The Economic Journal*, 115(500):224–246.

Van Praag, B. M. (2007). Perspectives from The Happiness Literature and the Role of New Instruments for Policy Analysis. *CESifo Economic Studies*, 53(1):42–68.

Welsch, H. and Kühling, J. (2009). Using Happiness Data for Environmental Valuation: Issues and Applications. *Journal of Economic Surveys*, 23(2):385–406.

Winkelmann, L. and Winkelmann, R. (1998). Why Are The Unemployed so Unhappy? Evidence from Panel Data. *Economica*, 65(257):1–15.

# Appendix A Figures

Figure 1: Compensating variation for two regions

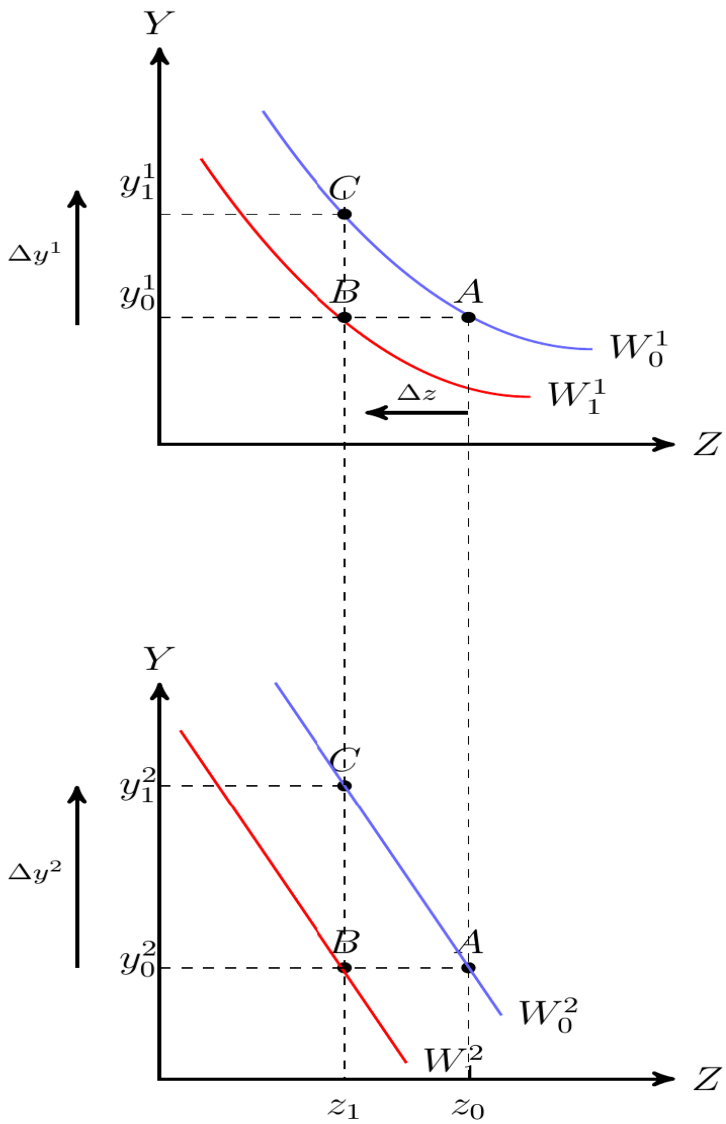
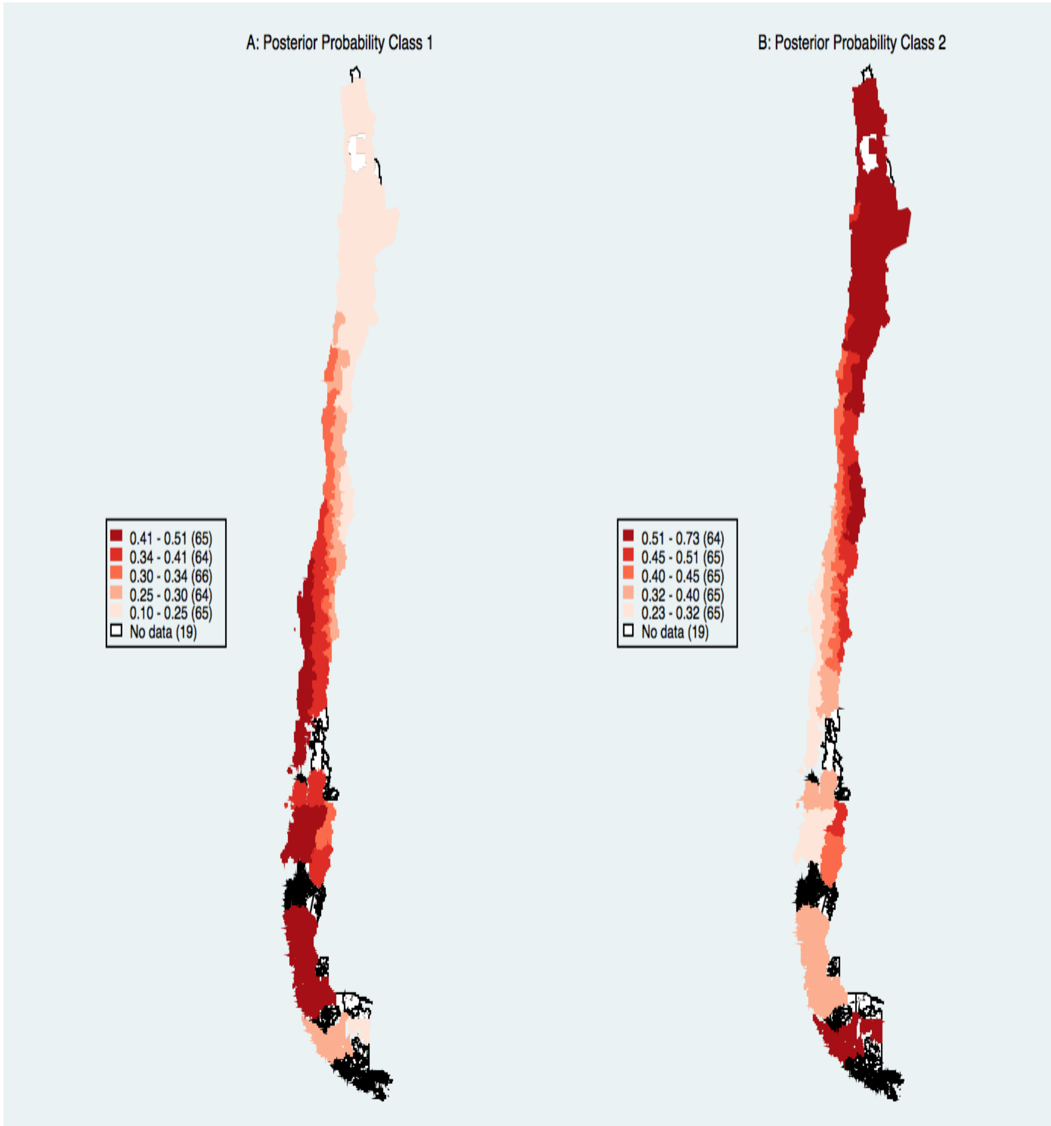
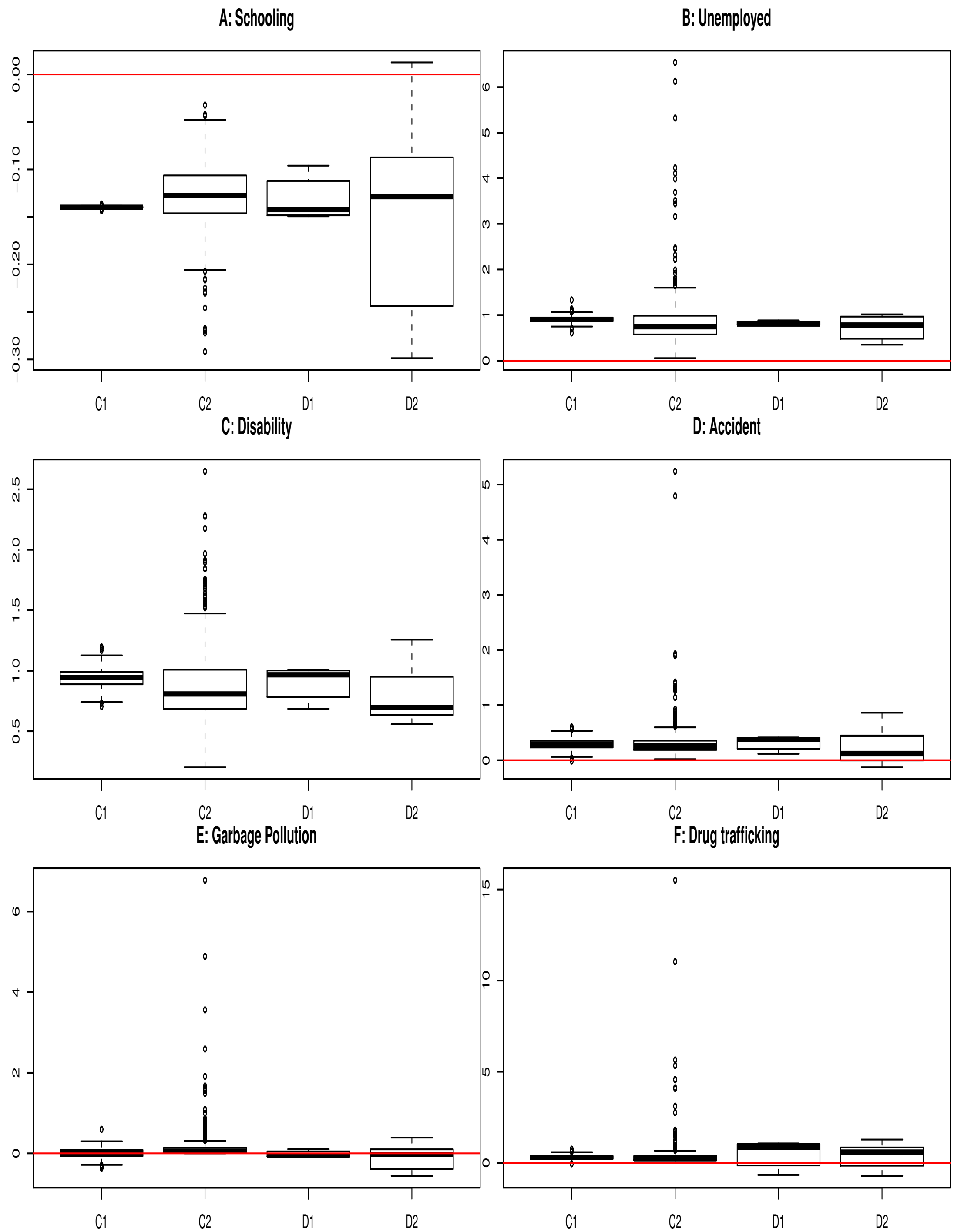


Figure 2: Spatial distribution of conditional probability of class assignment



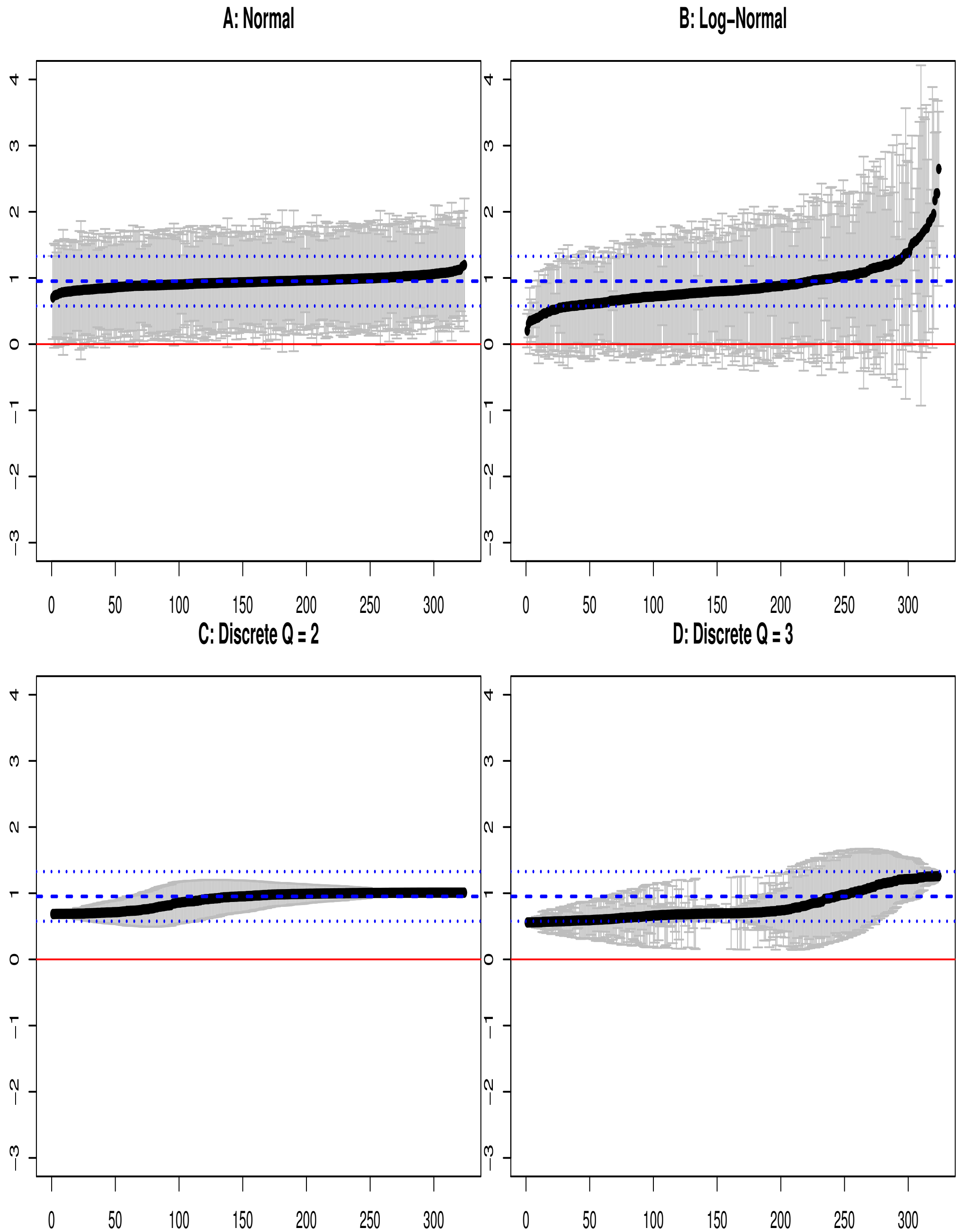

*Notes: This graph shows the commune-specific probabilities of belonging to each class. The estimates are computed as* $\widehat{\pi}_{cq} = \frac{\widehat{w}_{cq} \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)}{\sum_{q=1}^{Q} \widehat{w}_{cq} \prod_i f(W_{ci}|\mathbf{x}_{ci}, \widehat{\boldsymbol{\beta}}_{cq}, \widehat{\boldsymbol{\theta}}_q)}$*, where* $\widehat{\boldsymbol{\theta}}_q$ *correspond to the estimates of the model with two classes.*

Figure 3: Distribution of CVs for selected covariates



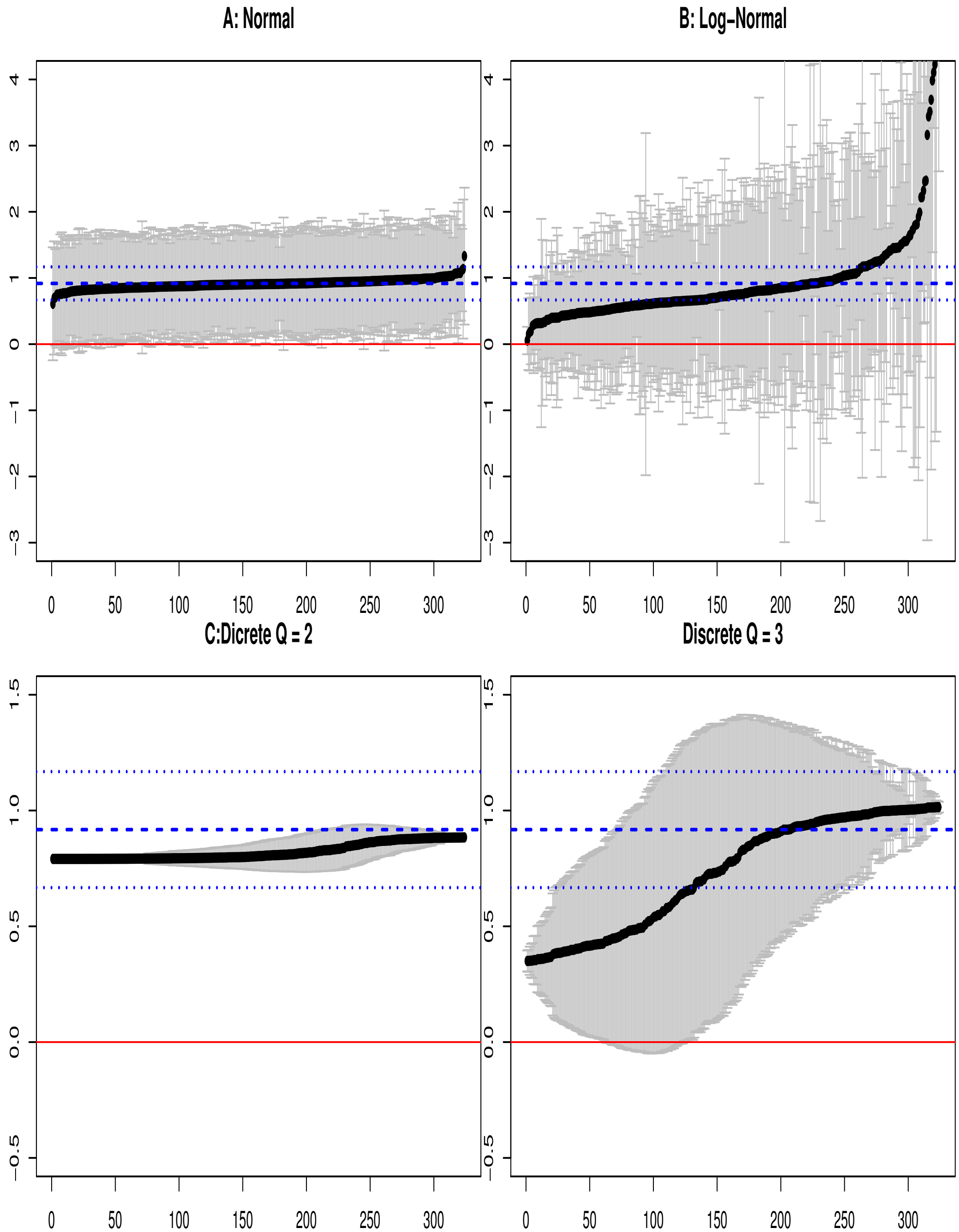

*Notes:* Each point correspond to the estimate of conditional expectation of compensating variation for each commune using Equation (11) and (12). C1 correspond to CVs based on estimates using model 2 (All Normals) from Table 2; C2 correspond to CVs based on estimates using model 4 (Correlated with LN and CN distributions) from Table 2; D1 and D2 correspond to CVs based on estimates using discrete distributions with 2 and 3 classes from Table 3.

Figure 4: Compensating Variation for Disability with 95% CI




*Notes:* Each point in the graph corresponds to a commune. The confidence intervals are computed using Equation (13). The graph from panel A corresponds to CVs estimated using coefficients from model 2 (All Normals) in Table 2. Panel B corresponds to CVs estimated using coefficients from model 3 (Correlated with LN and CN distributions) in Table 2 . Panel C and D correspond to CVs estimated using coefficient from models assuming discrete spatial heterogeneity. In particular, Panel C corresponds to model 1 in Table 3 and Panel D corresponds to model 2 in Table 3. The blue dashed lines correspond to CV and 95% confidence interval computed assuming fixed coefficients

Figure 5: Compensating Variation for Unemployed with 95% CI




*Notes:* Each point in the graph corresponds to a commune. The confidence intervals are computed using Equation (13). The graph from panel A corresponds to CVs estimated using coefficients from model 2 (All Normals) in Table 2. Panel B corresponds to CVs estimated using coefficients from model 3 (Correlated with LN and CN distributions) in Table 2 . Panel C and D correspond to CVs estimated using coefficient from models assuming discrete spatial heterogeneity. In particular, Panel C corresponds to model 1 in Table 3 and Panel D corresponds to model 2 in Table 3. The blue dashed lines correspond to CV and 95% confidence interval computed assuming fixed coefficients

# Appendix B Tables

Table 1: Summary Statistics

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Satisfied with life* | 16,008 | 0.528 | 0.499 | 0 | 1 |
| *Log(Household Income pc)* | 16,008 | 12.009 | 0.785 | 0.000 | 16.611 |
| *Schooling* | 16,008 | 9.967 | 3.980 | 0 | 22 |
| *Age* | 16,008 | 42.564 | 13.476 | 15 | 65 |
| *Male* | 16,008 | 0.330 | 0.470 | 0 | 1 |
| *Unemployed* | 16,008 | 0.043 | 0.202 | 0 | 1 |
| *OLF* | 16,008 | 0.409 | 0.492 | 0 | 1 |
| *Married* | 16,008 | 0.601 | 0.490 | 0 | 1 |
| *Household size* | 16,008 | 3.610 | 1.618 | 1 | 17 |
| *Disability* | 16,008 | 0.067 | 0.250 | 0 | 1 |
| *Treatment* | 16,008 | 0.301 | 0.459 | 0 | 1 |
| *Accident* | 16,008 | 0.214 | 0.410 | 0 | 1 |
| *Noise pollution* | 16,008 | 0.163 | 0.369 | 0 | 1 |
| *Air pollution* | 16,008 | 0.167 | 0.373 | 0 | 1 |
| *Water pollution* | 16,008 | 0.073 | 0.260 | 0 | 1 |
| *Garbage pollution* | 16,008 | 0.142 | 0.349 | 0 | 1 |
| *Robbery* | 16,008 | 0.309 | 0.462 | 0 | 1 |
| *Poor surveillance* | 16,008 | 0.316 | 0.465 | 0 | 1 |
| *Street drugs & alcohol* | 16,008 | 0.280 | 0.449 | 0 | 1 |
| *Drug trafficking* | 16,008 | 0.145 | 0.352 | 0 | 1 |
| *Log(Density)* | 16,008 | 3.813 | 2.354 | −1.514 | 9.558 |
| *Log(Median income)* | 16,008 | 12.302 | 0.264 | 11.667 | 13.816 |
| *Urban* | 16,008 | 0.670 | 0.470 | 0 | 1 |

Note: Individual covariates and variables based on income come from the 2013 National Socioeconomic Characterization Survey (CASEN) from Chile. Density comes from the Chilean National Institute of Statistic (INE).

## Table 2: Continuous Spatial Heterogeneity

| | Fixed Parameters | | All Normals | | | Correlated | | | Correlated (LN and CN) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | CV | Est. | CV | % > 0 | Est. | CV | % > 0 | Est. | Mean | SD | CV |
| | | | | | | *A: Mean Parameters* | | | | | | |
| *Log(Income)* | 0.236*** | | 0.239*** | | | 0.261*** | | | 0.260*** | | | |
| | (0.016) | | (0.017) | | | (0.013) | | | (0.012) | | | |
| *Constant* | -2.970*** | | -2.970*** | | | -2.491*** | | | -3.122*** | | | |
| | (0.532) | | (0.708) | | | (0.638) | | | (0.617) | | | |
| *Schooling* | 0.032*** | -0.137*** | 0.033*** | -0.140*** | 100% | 0.032*** | -0.125*** | 91% | -3.501*** | 0.034*** | 0.017*** | -0.130*** |
| | (0.003) | (0.019) | (0.003) | (0.019) | | (0.003) | (0.015) | | (0.125) | (0.003) | (0.005) | (0.018) |
| *Age* | -0.005*** | 0.020*** | -0.005*** | 0.022*** | 3% | -0.005*** | 0.019*** | 32% | -0.005*** | | | 0.021*** |
| | (0.001) | (0.004) | (0.001) | (0.004) | | (0.001) | (0.004) | | (0.001) | | | (0.004) |
| *Male* | 0.045* | -0.189* | 0.043* | -0.178* | 69% | 0.042* | -0.161* | 74% | 0.037 | | | -0.144 |
| | (0.024) | (0.101) | (0.024) | (0.103) | | (0.025) | (0.098) | | (0.025) | | | (0.096) |
| *Unemployed* | -0.217*** | 0.917*** | -0.217*** | 0.905*** | 2% | -0.225*** | 0.862*** | 17% | -2.012*** | 0.225*** | 0.304 | -0.866 |
| | (0.052) | (0.236) | (0.053) | (0.240) | | (0.056) | (0.226) | | (0.051) | (0.053) | (0.102) | (0.806) |
| *Inactive* | 0.082*** | -0.346*** | 0.084*** | -0.352*** | 84% | 0.091*** | -0.348*** | 81% | 0.088*** | | | -0.337*** |
| | (0.024) | (0.099) | (0.024) | (0.102) | | (0.025) | (0.097) | | (0.025) | | | (0.096) |
| *Married* | 0.194*** | -0.820*** | 0.195*** | -0.817*** | 99% | 0.205*** | -0.785 | 90% | 0.207*** | | | -0.798*** |
| | (0.022) | (0.108) | (0.023) | (0.111) | | (0.024) | (0.099) | | (0.023) | | | (0.098) |
| *Household Size* | 0.030*** | -0.127*** | 0.029*** | -0.120*** | 100% | 0.034*** | -0.132 | 70% | 0.036*** | | | -0.137*** |
| | (0.007) | (0.028) | (0.007) | (0.029) | | (0.008) | (0.029) | | (0.007) | | | (0.028) |
| *Disability* | -0.225*** | 0.952*** | -0.226*** | 0.944*** | 1% | -0.229*** | 0.878 | 17% | -1.699*** | 0.223*** | 0.155 | -0.857** |
| | (0.042) | (0.191) | (0.044) | (0.196) | | (0.048) | (0.188) | | (0.345) | (0.045) | (0.102) | (0.393) |
| *Medical Treat.* | -0.152*** | 0.643*** | -0.154*** | 0.645*** | 4% | -0.168*** | 0.643 | 8% | -2.397*** | 0.133** | 0.143*** | -0.514*** |
| | (0.025) | (0.114) | (0.026) | (0.116) | | (0.027) | (0.108) | | (0.319) | (0.026) | (0.044) | (0.169) |
| *Accident* | -0.070*** | 0.298*** | -0.069** | 0.287** | 23% | -0.078*** | 0.299 | 35% | -3.156*** | 0.085*** | 0.144 | -0.326 |
| | (0.026) | (0.110) | (0.027) | (0.114) | | (0.028) | (0.109) | | (0.694) | (0.027) | (0.093) | (0.359) |
| *Noise* | -0.076*** | 0.321 | -0.070** | 0.294** | 23% | -0.066** | 0.252 | 41% | -0.228 | | | 0.878 |
| | (0.029) | (0.124) | (0.030) | (0.129) | | (0.032) | (0.124) | | (0.320) | | | (1.213) |
| *Air Poll.* | -0.014 | 0.060 | -0.009 | 0.039 | 46% | -0.003 | 0.011 | 49% | -0.181 | | | 0.696 |
| | (0.028) | (0.119) | (0.030) | (0.125) | | (0.031) | (0.120) | | (0.732) | | | (2.706) |
| *Water Poll.* | -0.044 | 0.187 | -0.044 | 0.184 | 33% | -0.077* | 0.296 | 41% | -0.720 | | | 2.772 |
| | (0.040) | (0.168) | (0.042) | (0.175) | | (0.046) | (0.178) | | (1.192) | | | (4.488) |
| *Trash Poll.* | -0.003 | 0.012 | -0.003 | 0.013 | 49% | -0.001 | 0.005 | 50% | -1.050 | | | 4.043 |
| | (0.030) | (0.128) | (0.032) | (0.132) | | (0.034) | (0.131) | | (0.816) | | | (3.172) |
| *Robery* | -0.042* | 0.179* | -0.042* | 0.177* | 33% | -0.036 | 0.139 | 44% | -3.820*** | 0.053** | 0.115 | -0.203 |
| | (0.023) | (0.099) | (0.025) | (0.104) | | (0.026) | (0.100) | | (1.039) | (0.024) | (0.086) | (0.331) |
| *Poor surveillance* | -0.053** | 0.225** | -0.052** | 0.218** | 27% | -0.050* | 0.191 | 41% | -4.056*** | 0.054** | 0.156 | -0.206 |
| | (0.023) | (0.099) | (0.024) | (0.102) | | (0.026) | (0.098) | | (1.089) | (0.023) | (0.110) | (0.423) |
| *Alcohol* | -0.056** | 0.236** | -0.051* | 0.213* | 30% | -0.065** | 0.251 | 41% | -3.823*** | 0.058** | 0.140 | -0.222 |
| | (0.025) | (0.109) | (0.027) | (0.114) | | (0.029) | (0.111) | | (1.140) | (0.025) | (0.116) | (0.447) |
| *Trafficking* | -0.071** | 0.301** | -0.070** | 0.293** | 23% | -0.048 | 0.186 | 45% | -3.714*** | 0.087*** | 0.295 | -0.334 |
| | (0.032) | (0.135) | (0.034) | (0.142) | | (0.036) | (0.138) | | (1.133) | (0.032) | (0.301) | (0.447) |
| *Log(Density)* | -0.007 | 0.031 | -0.010 | 0.041 | 0% | -0.008 | 0.030 | 41% | -0.011* | | | 0.043* |
| | (0.005) | (0.021) | (0.006) | (0.026) | | (0.006) | (0.024) | | (0.006) | | | (0.023) |
| *Log(Income)* | -0.004 | 0.016 | -0.004 | 0.017 | 18% | -0.069 | 0.263 | 31% | -0.014 | | | 0.054 |
| | (0.046) | (0.193) | (0.060) | (0.250) | | (0.054) | (0.204) | | (0.052) | | | (0.199) |
| *Urban* | 0.078*** | -0.323*** | 0.074*** | -0.309*** | 83% | 0.082*** | -0.313 | 61% | 0.068** | | | -0.261* |
| | (0.024) | (0.104) | (0.025) | (0.108) | | (0.031) | (0.105) | | (0.026) | | | (0.102) |
| | | | | | | *B: Standard Deviations* | | | | | | |
| $\upsilon_c$ | | | 0.065** | | | 2.062* | | | 1.631 | | | |
| | | | (0.028) | | | (1.174) | | | (1.305) | | | |
| *Schooling* | | | 0.001 | | | 0.024*** | | | 0.481*** | | | |
| | | | (0.002) | | | (0.006) | | | (0.118) | | | |
| *Age* | | | 0.003*** | | | 0.010*** | | | 0.008*** | | | |
| | | | (0.000) | | | (0.002) | | | (0.002) | | | |
| *Male* | | | 0.086** | | | 0.065 | | | 0.059 | | | |
| | | | (0.038) | | | (0.044) | | | (0.050) | | | |
| *Unemployed* | | | 0.101 | | | 0.237** | | | 1.020** | | | |
| | | | (0.133) | | | (0.108) | | | (0.433) | | | |
| *Inactive* | | | 0.085** | | | 0.102*** | | | 0.118*** | | | |
| | | | (0.039) | | | (0.039) | | | (0.044) | | | |
| *Married* | | | 0.078*** | | | 0.161*** | | | 0.175*** | | | |
| | | | (0.028) | | | (0.043) | | | (0.046) | | | |
| *Household Size* | | | 0.011 | | | 0.065*** | | | 0.061*** | | | |
| | | | (0.008) | | | (0.012) | | | (0.013) | | | |
| *Disability* | | | 0.096 | | | 0.243*** | | | 0.627* | | | |
| | | | (0.094) | | | (0.090) | | | (0.357) | | | |
| *Medical Treat.* | | | 0.087** | | | 0.121*** | | | 0.875*** | | | |
| | | | (0.038) | | | (0.045) | | | (0.205) | | | |
| *Accident* | | | 0.092* | | | 0.196*** | | | 1.171*** | | | |
| | | | (0.053) | | | (0.049) | | | (0.437) | | | |
| *Noise* | | | 0.097 | | | 0.300*** | | | 0.478* | | | |
| | | | (0.062) | | | (0.053) | | | (0.258) | | | |
| *Air Poll.* | | | 0.089** | | | 0.180*** | | | 0.204 | | | |
| | | | (0.045) | | | (0.052) | | | (0.496) | | | |
| *Water Poll.* | | | 0.098 | | | 0.356*** | | | 0.789 | | | |
| | | | (0.161) | | | (0.080) | | | (0.772) | | | |
| *Trash Poll.* | | | 0.106 | | | 0.320*** | | | 0.829* | | | |
| | | | (0.078) | | | (0.057) | | | (0.457) | | | |
| *Robbery* | | | 0.095** | | | 0.226*** | | | 1.323** | | | |
| | | | (0.040) | | | (0.039) | | | (0.553) | | | |
| *Poor surveillance* | | | 0.086* | | | 0.220*** | | | 1.502*** | | | |
| | | | (0.045) | | | (0.041) | | | (0.524) | | | |
| *Alcohol* | | | 0.098 | | | 0.275*** | | | 1.391** | | | |
| | | | (0.068) | | | (0.043) | | | (0.610) | | | |
| *Trafficking* | | | 0.097 | | | 0.364*** | | | 1.593** | | | |
| | | | (0.092) | | | (0.051) | | | (0.626) | | | |
| *Log(Density)* | | | 0.002 | | | 0.034*** | | | 0.027** | | | |
| | | | (0.005) | | | (0.010) | | | (0.012) | | | |
| *Log(Income)* | | | 0.004* | | | 0.136*** | | | 0.117 | | | |
| | | | (0.002) | | | (0.090) | | | (0.100) | | | |
| *Urban* | | | 0.079** | | | 0.305*** | | | 0.305*** | | | |
| | | | (0.035) | | | (0.038) | | | (0.038) | | | |
| LL | -10560 | | -10500 | | | -10280 | | | -10330 | | | |
| # parameters | 23 | | 45 | | | 276 | | | 276 | | | |
| $N$ | 16008 | | 16008 | | | 16008 | | | 16008 | | | |
| AIC | 21176 | | 21081 | | | 21100 | | | 21206 | | | |
| BIC | 21352 | | 21426 | | | 23220 | | | 23326 | | | |

Note: Simulation based on 100 Halton draws. Standard errors in parentheses. Standard errors for CV were computed using Delta Method. Mean and SD corresponds to the mean and standard deviation for the variables assumed to be log-normal distributed. $AIC = -2\log L + 2K$, $BIC = -2\log L + \log(N)K$. $^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.001$

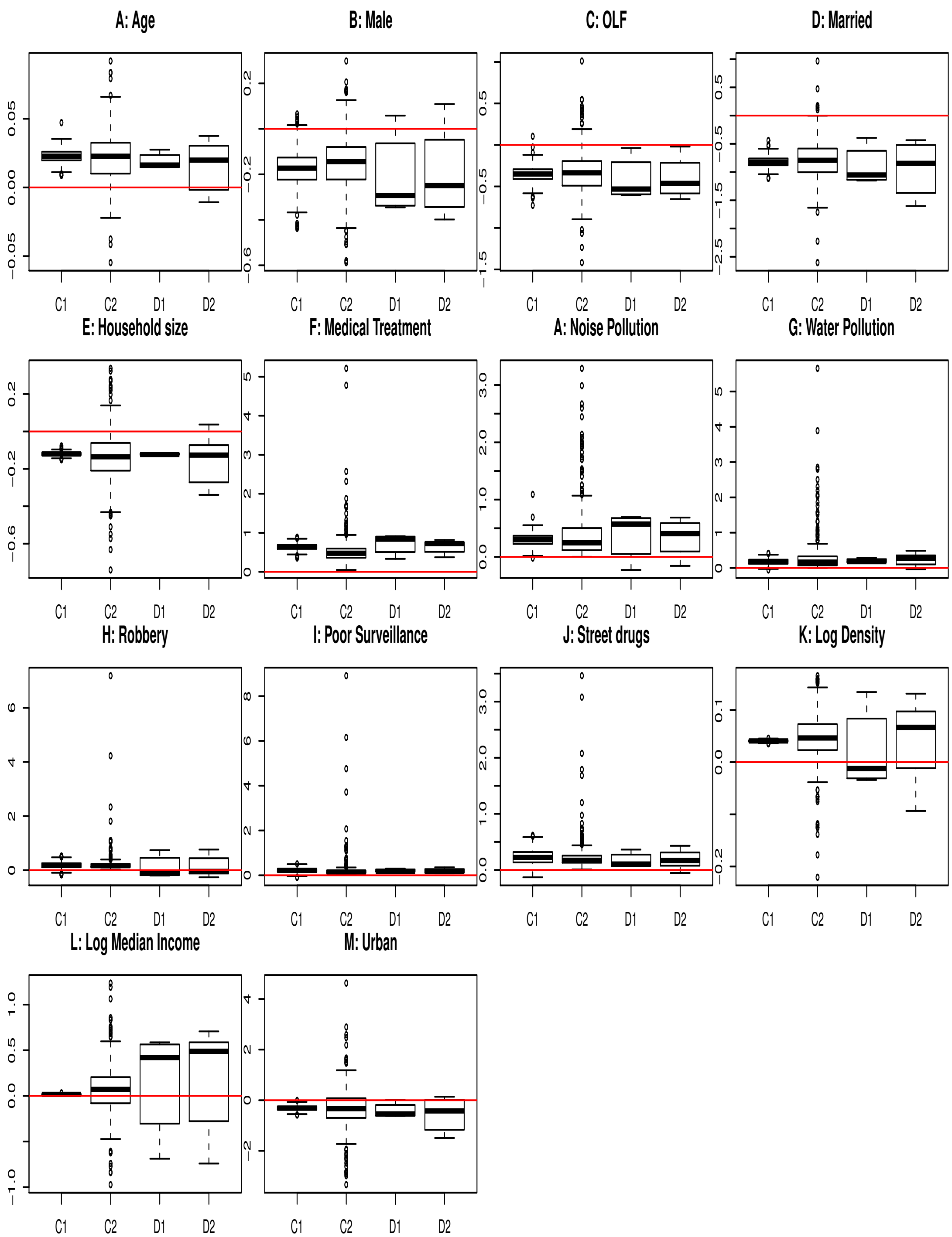
## Table 3: Discrete Spatial Heterogeneity

| | Model 1 | | | | Model 2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***Class 1*** | | ***Class 2*** | | ***Class 1*** | | ***Class 2*** | | ***Class 3*** | |
| | *Est.* | *CV* | *Est.* | *CV* | *Est.* | *CV* | *Est.* | *CV* | *Est.* | *CV* |
| *Log(Income)* | 0.365*** | | 0.196*** | | 0.365*** | | 0.177*** | | 0.266*** | |
| | (0.037) | | (0.014) | | (0.039) | | (0.017) | | (0.040) | |
| *Constant* | -7.385*** | | -1.187* | | -7.435*** | | -1.602* | | -0.791 | |
| | (1.317) | | (0.700) | | (1.407) | | (0.903) | | (1.607) | |
| *Schooling* | 0.035*** | -0.096*** | 0.029*** | -0.149*** | 0.033*** | -0.090*** | 0.053*** | -0.299*** | -0.003 | 0.013 |
| | (0.007) | (0.023) | (0.004) | (0.026) | (0.007) | (0.024) | (0.006) | (0.049) | (0.008) | (0.028) |
| *Age* | -0.010*** | 0.027*** | -0.003** | 0.015** | -0.011*** | 0.031*** | 0.002 | -0.011 | -0.010*** | 0.038*** |
| | (0.002) | (0.006) | (0.001) | (0.006) | (0.002) | (0.006) | (0.002) | (0.010) | (0.002) | (0.009) |
| *Male* | -0.021 | 0.058 | 0.068** | -0.345** | -0.040 | 0.109 | 0.070 | -0.398 | 0.064 | -0.241 |
| | (0.050) | (0.138) | (0.031) | (0.162) | (0.053) | (0.144) | (0.045) | (0.258) | (0.057) | (0.219) |
| *Unemployed* | -0.322*** | 0.883*** | -0.155** | 0.791** | -0.370*** | 1.015*** | -0.062 | 0.350 | -0.268** | 1.006** |
| | (0.112) | (0.332) | (0.067) | (0.355) | (0.119) | (0.355) | (0.099) | (0.570) | (0.120) | (0.496) |
| *Inactive* | 0.013 | -0.036 | 0.119*** | -0.606*** | 0.007 | -0.019 | 0.115** | -0.652** | 0.121** | -0.453** |
| | (0.050) | (0.138) | (0.031) | (0.162) | (0.053) | (0.144) | (0.045) | (0.256) | (0.057) | (0.214) |
| *Married* | 0.144*** | -0.395*** | 0.225*** | -1.150*** | 0.158*** | -0.435*** | 0.283*** | -1.601*** | 0.121** | -0.454** |
| | (0.047) | (0.136) | (0.029) | (0.172) | (0.049) | (0.144 ) | (0.042) | (0.285) | (0.053) | (0.208) |
| *Household Size* | 0.043*** | -0.117*** | 0.024*** | -0.123*** | 0.027* | -0.074* | 0.060*** | -0.339*** | -0.010 | 0.038 |
| | (0.016) | (0.041) | (0.009) | (0.045) | (0.016) | (0.043 ) | (0.013) | (0.077) | (0.016) | (0.063) |
| *Disability* | -0.250*** | 0.685*** | -0.198*** | 1.009*** | -0.254*** | 0.696*** | -0.099 | 0.559 | -0.335*** | 1.258*** |
| | (0.087) | (0.248) | (0.057) | (0.300) | (0.092) | (0.261) | (0.083) | (0.475) | (0.101) | (0.425) |
| *Medical Treat.* | -0.122** | 0.333** | -0.179*** | 0.912*** | -0.136** | 0.373** | -0.131*** | 0.743*** | -0.218*** | 0.818*** |
| | (0.054) | (0.151) | (0.033) | (0.181) | (0.056) | (0.159) | (0.048) | (0.279) | (0.059) | (0.255) |
| *Accident* | -0.042 | 0.115 | -0.082** | 0.421** | -0.046 | 0.126 | 0.022 | -0.123 | -0.230*** | 0.863*** |
| | (0.055) | (0.152) | (0.034) | (0.176) | (0.058) | (0.160) | (0.049) | (0.281) | (0.062) | (0.266) |
| *Noise* | 0.082 | -0.226 | -0.135*** | 0.691*** | 0.058 | -0.158 | -0.121** | 0.687** | -0.099 | 0.373 |
| | (0.062) | (0.170) | (0.039) | (0.205) | (0.065) | (0.178) | (0.055) | (0.321) | (0.072) | (0.273) |
| *Air Poll.* | 0.190*** | -0.520*** | -0.087** | 0.445** | 0.212*** | -0.583*** | -0.042 | 0.240 | -0.150** | 0.561** |
| | (0.061) | (0.176) | (0.037) | (0.193) | (0.065) | (0.188) | (0.053) | (0.303) | (0.068) | (0.265) |
| *Water Poll.* | -0.102 | 0.278 | -0.033 | 0.168 | 0.013 | -0.035 | -0.048 | 0.270 | -0.130 | 0.487 |
| | (0.099) | (0.273) | (0.052) | (0.266) | (0.105) | (0.288) | (0.072) | (0.410) | (0.093) | (0.352) |
| *Trash Poll.* | -0.038 | 0.104 | 0.016 | -0.079 | -0.035 | 0.095 | 0.099* | -0.559* | -0.104 | 0.391 |
| | (0.065) | (0.179) | (0.040) | (0.205) | (0.068) | (0.187) | (0.058) | (0.333) | (0.076) | (0.290) |
| *Robbery* | -0.269*** | 0.739*** | 0.039 | -0.200 | -0.278*** | 0.762*** | 0.015 | -0.087 | 0.071 | -0.265 |
| | (0.051) | (0.157) | (0.031) | (0.157) | (0.054) | (0.166) | (0.044) | (0.249) | (0.056) | (0.215) |
| *Poor surveillance* | -0.110** | 0.302** | -0.032 | 0.164 | -0.129** | 0.354** | -0.015 | 0.084 | -0.051 | 0.190 |
| | (0.049) | (0.139) | (0.031) | (0.156) | (0.051) | (0.147) | (0.044) | (0.250) | (0.055) | (0.209) |
| *Alcohol* | -0.132** | 0.363** | -0.013 | 0.068 | -0.157*** | 0.431*** | -0.030 | 0.167 | 0.014 | -0.053 |
| | (0.054) | (0.153) | (0.034) | (0.173) | (0.056) | (0.163) | (0.049) | (0.276) | (0.062) | (0.232) |
| *Trafficking* | 0.243*** | -0.667*** | -0.208*** | 1.064*** | 0.262*** | -0.718*** | -0.107* | 0.605* | -0.340*** | 1.276*** |
| | (0.070) | (0.202) | (0.041) | (0.225) | (0.074) | (0.215) | (0.058) | (0.337) | (0.074) | (0.339) |
| *Log(Density)* | -0.049*** | 0.135*** | 0.007 | -0.034 | -0.048*** | 0.131*** | -0.013 | 0.072 | 0.025* | -0.094* |
| | (0.011) | (0.033) | (0.007) | (0.034) | (0.012) | (0.035) | (0.009) | (0.048) | (0.013) | (0.051) |
| *Log(Income)* | 0.251** | -0.689** | -0.115* | 0.586* | 0.270** | -0.742** | -0.125* | 0.706* | -0.133 | 0.500 |
| | (0.112) | (0.330) | (0.059) | (0.296) | (0.119) | (0.353) | (0.075) | (0.423) | (0.136) | (0.499) |
| *Urban* | -0.002 | 0.004 | 0.122*** | -0.621*** | -0.041 | 0.112*** | 0.264*** | -1.493*** | -0.037 | 0.138 |
| | (0.051) | (0.139) | (0.032) | (0.171) | (0.053) | (0.146) | (0.047) | (0.297) | (0.057) | (0.212) |
| Class assignment variables | | | | | | | | | | |
| *x* | | | 4.622*** | | | | 7.620*** | | 5.135*** | |
| | | | (0.314) | | | | (0.380) | | (0.389) | |
| *y* | | | -0.079*** | | | | -0.271*** | | -0.225*** | |
| | | | (0.013) | | | | (0.017) | | (0.017) | |
| LL | -10489.63 | | | | -10451.4 | | | | | |
| *# parameters* | 48 | | | | 73 | | | | | |
| *N* | 16008 | | | | 16008 | | | | | |
| *AIC* | 21075 | | | | 21049 | | | | | |
| *BIC* | 21181 | | | | 21210 | | | | | |

Note: Standard errors in parentheses. Standard errors for compensating variations (CVs) are computed using Delta method. $AIC = -2\log L + 2K$, $BIC = -2\log L + \log(N)K$. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.001$.

# Appendix C Additional Figures and Tables

Figure A1: Distribution of CVs for rest of covariates




*Notes:* Each point correspond to the estimate of conditional expectation of compensating variation for each commune using Equation (11) and (12). C1 correspond to CVs based on estimates using model 2 (All Normals) from Table 2; C2 correspond to CVs based on estimates using model 4 (Correlated with LN and CN distributions) from Table 2; D1 and D2 correspond to CVs based on estimates using discrete distributions with 2 and 3 classes from Table 3.

Table A1: Analysis of missing values

| | Answer SWB | | Do not Answer SWB | | |
|---|---|---|---|---|---|
| | Mean | Sample | Mean | Sample | Difference |
| *Log(Income)* | 12.009 | 16008 | 11.871 | 5767 | 0.138*** |
| *Schooling* | 9.967 | 16008 | 9.852 | 3731 | 0.114 |
| *Age* | 42.564 | 16008 | 27.886 | 5767 | 14.677*** |
| *Household Size* | 3.61 | 16008 | 4.424 | 5767 | -0.814*** |
| *Male* | 0.33 | 16008 | 0.571 | 5767 | -0.241*** |
| *Unemployed* | 0.043 | 16008 | 0.039 | 3756 | 0.004 |
| *Inactive* | 0.409 | 16008 | 0.439 | 3756 | -0.030*** |
| *Married* | 0.601 | 16008 | 0.285 | 5767 | 0.316*** |
| *Disability* | 0.067 | 16008 | 0.06 | 5760 | 0.007 |
| *Medical Treat.* | 0.301 | 16008 | 0.158 | 5767 | 0.143*** |
| *Accident* | 0.214 | 16008 | 0.163 | 5767 | 0.051*** |
| *Noise* | 0.163 | 16008 | 0.158 | 5767 | 0.004 |
| *Air Poll.* | 0.167 | 16008 | 0.156 | 5767 | 0.011 |
| *Water Poll.* | 0.073 | 16008 | 0.068 | 5767 | 0.005 |
| *Trash Poll.* | 0.142 | 16008 | 0.136 | 5767 | 0.006 |
| *Robbery* | 0.309 | 16008 | 0.301 | 5767 | 0.008 |
| *Poor surveillance* | 0.316 | 16008 | 0.319 | 5767 | -0.003 |
| *Alcohol* | 0.28 | 16008 | 0.286 | 5767 | -0.006 |
| *Trafficking* | 0.145 | 16008 | 0.14 | 5767 | 0.005 |

*Notes:* $t$-test corresponds to the difference in means for continuous variables and difference in proportion test for dummy variables. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.001$.