ACCEPTED VERSION

Assessing Colombia's policy of socio-economic stratification: An intra-city study of self-reported quality of life

Published journal article:

Chica-Olmo, J., Sánchez, A. & Sepúlveda-Murillo, F.H. (2020). Assessing Colombia's policy of socio-economic stratification: An intra-city study of self-reported quality of life of life. *Cities, 97*. https://doi.org/10.1016/j.cities.2019.102560

1. Introduction

Colombian cities are geographically divided into six socio-economic strata (from the lowest stratum 1 to the highest stratum 6), in which dwellings are grouped according to their characteristics and the area where they are located. Taking this classification as a reference, municipal governments allocate subsidies, collect municipal taxes, and charge households different rates for the residential services they provide (DANE 2015).

Assuming that the objective of any public policy is to improve citizens' quality of life, we inquire as to whether the classification of Colombian households into socio-economic strata accurately reflects the level of quality of life of households in the city of Medellin. Given the growing interest among economists and policymakers in subjective well-being and its determinants as a more adequate means to measure economic and social progress and monitor well-being in a more comprehensive way (Frey and Stutzer 2017; Odermatt and Stutzer 2018; O'Donnell et al. 2014), this paper examines Colombia's public policy of stratification in Medellin from the perspective of self-reported quality of life.

The study of self-reported quality of life helps to assess the degree to which countries or cities meet the needs of their citizens and to what extent they can flourish or progress in that environment (Veenhoven 2017). Traditional measures of welfare, such as economic growth, have not fulfilled society's need to effectively monitor progress (Frey and Stutzer 2017; Stiglitz et al. 2011; Van den Bergh 2009). Likewise, policymakers increasingly consider the subjective well-being of the population as a possible policy goal (see O'Donnell et al. 2014; Rojas 2016). There is evidence that improvements in subjective well-being have positive effects for both individuals and society as a whole in many regards, including longer life expectancy (Diener and Chan 2011), increased labour productivity (Oswald et al. 2015), greater participation in voluntary and altruistic activities (Meier and Stutzer 2008) and, in the specific case of Latin America, higher voter turnout (Weitz-Shapiro and Winters 2011).

With this in mind, and taking into account the scarcity of studies on the distribution of subjective well-being within Colombia (Hurtado 2016), we aim to test **three hypotheses**. Firstly,

given that the composition of the socio-economic strata in Medellin is not random but takes into account characteristics of the houses and the environment[1] which affect perceived quality of life in Latin America's cities (see Ahumada et al. 2019; Gandelman et al. 2012; Medina et al. 2008), we hypothesize that the structure of households in Medellin is hierarchical or nested since the strata are more likely to be comprised of residents who share similar values and interests. To account for this structure, we perform a cross-tabulation analysis of quality of life and strata and estimate the null model with multilevel modelling. Whether nested structures were confirmed, multilevel modelling would be a suitable approach because standard estimation techniques could lead to incorrect conclusions (see Goldstein 2011; Snijders and Bosker 2012).

Secondly, we hypothesize that, in addition to the nested structure by strata, the inclusion of economic and non-economic factors identified in the subjective well-being literature would result in a ranking of self-reported quality of life by strata that differs greatly from that of the Colombian authorities.

Thirdly, focusing on the economic resources of households as a traditional reference variable of economic and social policies, we hypothesize that changes in economic resources do not have the same effect on self-reported quality of life for households belonging to different strata. Multilevel modelling allows us to check this hypothesis by introducing the stratum slopes of these household resources as random terms and analyse the interaction between stratum and economic resources.

Confirming these hypotheses could have policy implications, especially at the local level, because the Colombian authorities use this stratification to differentially charge for residential public utility services such that higher strata households pay more to subsidize utility rates for households in the lower strata.

To test these hypotheses, we use data from the Quality of Life Survey 2014 conducted by the Medellin Administrative Department of Planning. Medellin is Colombia's second most populated city with 2.5 million inhabitants in 2017. The availability of this dataset at intra-city level constitutes a strength of our study. In the framework of multilevel models, administrative area, region, or country level is usually chosen due to data availability. However, this does not necessarily represent an individual's daily social interactions precisely (Giordano et al. 2011). In our case, as noted by Farrell et al. (2004), the use of socio-economic strata as the unit of analysis,

[1] The stratification methodology applied by the Colombian authorities is analysed in more detail in section 2.2.

rather than the city, region, or country (which are expected to contain heterogeneous neighbourhoods), allows for a more fine-grained investigation.

The rest of this paper is structured as follows. Section 2 focuses on the conceptual framework, where we review the concept of quality of life, the literature on the contextual effects, and the methodology applied by the Colombian authorities to geographically classify the residential housing into socio-economic strata. Section 3 presents and justifies the dataset and variables used in the analysis. Section 4 describes the empirical strategy under the multilevel modelling approach. Section 5 is dedicated to the main results of our analysis. Lastly, conclusions are drawn in section 6.

2. Conceptual framework

2.1. Quality of life and its determinants

The terms 'well-being' and 'quality of life' may be used interchangeably in the broad sense of living a good life (Veenhoven 2017). Quality of life implies two things. Firstly, that the minimum conditions required for humans to thrive are met and, secondly, that there is a sufficient fit between opportunities and capacities (Veenhoven 2000b). In other words, quality of life means that individuals are able to value life's actual outcomes and usefulness, as well as enjoyment. Hence, self-reported quality of life is referred to as the 'overall evaluation of life [which] involves all the criteria figuring in the mind of an individual' (Veenhoven 2017, p. 7).[2] Reported quality of life is measured by asking individuals to provide a global assessment of their life or domains of life, such as economic resources, health, employment, and social relationships, among others (Di Tella and MacCulloc 2006; Dolan et al. 2008; Helliwell and Huang 2014). Individuals are capable of making overall appraisals of their quality of life (Veenhoven 2000a) as they are not an academically constructed concept but one people use and understand (Rojas 2016). In contrast to other evaluative criteria (such as material wealth or good health), self-reported quality of life encompasses nonmaterial aspects of human well-being, such as the influence of social relationships (family and friends), comparisons with others, self-determination, and the absence of insecurity (Bárcena-Martín et al. 2017; Bartolini and Sarracino 2014; Dolan and Metcalfe 2012; Frey and Stutzer 2002; Stutzer and Frey 2010). An additional advantage to self-reported quality of life is that it takes into account individual judgements about aspects relating to outcome or actual quality of life, whereas other measures are indicators of potential quality of life (Frey and Stutzer 2002, 2017). In this regard, the study of self-reported quality of life in

[2]Following Veenhoven (2017), in this paper we use the terms 'subjective well-being', 'happiness', 'life satisfaction' and 'reported quality of life' as synonyms to refer to the degree to which an individual judges the overall quality of his or her life as favourable.

Latin American countries is very important given that the results could imply rethinking subsequent development strategies and implementing reforms (Rojas 2016).

Most empirical works on subjective well-being have traditionally focused on the study of individual determinants, such as socio-demographic characteristics and other economic and social domains. However, several studies have shown that perceived well-being is conditioned not only by individual determinants (level 1), but also by the geographical contexts (level 2) in which people live. In this literature, multilevel (hierarchical or mixed) modelling is used to distinguish individual effects from contextual effects of subjective well-being. Contextual variables are usually generated in two different ways. For example, data can be collected directly at level 2 (region, country, municipality, etc.) using macroeconomic variables or through surveys at the neighbourhood or municipal level. The second way is to generate the variables at level 2 from the variables at level 1. In the review of the literature that follows, we group the studies according to how the contextual variables are generated.

As regards the first option, Inglehart et al. (2008) used pooled data from the World Values Survey on subjective well-being for 52 countries over the period 1981 to 2007. The authors examined how contextual factors (country level) and individual factors shape subjective well-being. As contextual variables, they considered data on GDP per capita and economic growth rate from the World Bank database and a measure of a society's level of democracy from the Polity IV project (University of Maryland). They found that certain types of societies are more conducive to happiness than others. While economic factors have a strong impact on subjective well-being in low-income countries, at higher levels of development, societies that allow people relatively free choice in how to live their lives foster a higher level of subjective well-being. In the same vein, Novak and Pahor (2017) also used data on 40 countries from the sixth wave of the World Values Survey (2010-2014) to analyse the influence of economic development on life satisfaction. More specifically, they considered individuals' characteristics and macroeconomic variables at the country-level using World Bank data. As concerns contextual effects, the authors found that gross national income per capita had a significant and positive impact on life satisfaction, but that the effect of unemployment and inflation rate was not significant.

Most studies that generate contextual variables (level 2) from individual variables (level 1) in a multilevel framework have focused on social capital since it is often considered a collective concept. Neira et al. (2018) studied how contextual effects (at regional level) of social capital affected subjective well-being in the European Union in 2012. The authors developed five indicators of social capital using principal components analysis where regional social capital variables were defined as the average values of individual indicators. They found a strong contextual effect of living in societies with a more developed civil society, that is, where people

use more formal networks or are more involved in social organizations. Nevertheless, given that the average of individual measures has some drawbacks (see Oshio 2017, p. 770), other studies have estimated the contextual variables of social capital using a multilevel framework. More specifically, they fit random intercept models with social capital as the dependent variable. The residuals, which indicate the degree to which social capital in an area (context) differs from the mean over the entire sample, have also been used as explanatory variables in multilevel models for subjective well-being. In this line, Oshio (2017) analysed the association between four kinds of social capital and perceived happiness in Japan in 2011, distinguishing between individual-level and municipality-level social capital. He observed that both individual-level and municipality-level social capital had a positive and strong association with perceived happiness when they were used separately to predict perceived happiness. Furthermore, for all social capital measures, municipality-level social capital showed a much weaker association with perceived happiness than individual-level social capital. Following the same method for estimating contextual variables and using data from the Seoul Welfare Panel Study in 2008, Han (2015) considered social capital at three levels: individual, household and administrative areas of Seoul. He found that a relatively small percentage of happiness was attributed to the administrative-area level compared to the household level, which implies that a household context is more important for understanding variation in individual happiness.

2.2. Method of socio-economic stratification for residential housing in Colombia

Unlike the studies reviewed in the previous section, the main goal of this paper is not to identify the (individual or contextual) determinants of quality of life in Medellin. Rather, we are interested in assessing the effect on the city of Colombia's public policy of geographical stratification from the perspective of self-reported quality of life as a more comprehensive concept of social progress. That is, we aim to determine whether the classification of households in socio-economic strata accurately reflects the level of quality of life of the households. To do so, it is necessary to first provide a brief review of the method used by the Colombian authorities for classifying residential housing into socio-economic strata.

The socio-economic stratification of residential housing in Colombia was carried out from 2012 to 2014 by municipal governments throughout the country. The stratification methodology was designed by the National Administrative Department of Statistics (DANE, 2015).

The objective of this stratified, socio-economic division of the city of Medellin was to group together dwellings with similar characteristics, as well as the streets and geographical areas where they are located. To this end, six socio-economic strata were established: 1 low-

low, 2 low, 3 medium-low, 4 medium, 5 medium-high and 6 high. On the basis of this classification, the municipal government charges households different rates for the residential services it provides,[3] allocates subsidies, and collects municipal taxes. Prior to the stratification process, the dwellings of the most unprotected population, such as indigenous settlements in rural areas, housing for victims of forced displacement, and free housing (subsidized at 100%) were classified as stratum 1 without undergoing the stratification process. The rest of the dwellings were classified according to the following variables: (1) topography (sloped ground or not); (2) type of road (paved, unpaved, untracked); (3) public services available to the household (complete or incomplete); (4) land uses (institutional, residential, agricultural, mixed); and (5) housing characteristics, such as type of building (apartment or house), floor area, and number of rooms.

Firstly, physically homogenous zones were established using the first four variables mentioned above as a reference. Secondly, geoeconomically homogenous zones were determined by grouping together dwellings of a similar price on the real estate market. To do so, the physically homogenous zones were taken into account and the characteristics of the dwellings were incorporated (variables 5). Based on these two classifications, a search was made of the spatial intersection in order to establish the spatial stratification units (SSU; UEE in Spanish). In these units, adjoining dwellings with similar physical and economic characteristics but which differ with respect to adjacent areas were grouped together. In the last step, the six strata were obtained from the SSU. For this purpose, each SSU was assigned a value that represents its quality, and each of the six strata were obtained using minimum variance methods (see DANE 2015, pp. 50–54).

Therefore, the procedure designed by the DANE to obtain the strata only considers the characteristics of the dwellings, the streets and the surrounding areas, which in turn affect housing prices.

3. Data and variables

3.1. Data

The empirical analysis in this study is based on data drawn from the Quality of Life Survey 2014 (Encuesta de Calidad de Vida de 2014, hereafter ECV2014) conducted by the Medellin Administrative Department of Planning using simple random sampling of households by the six socio-economic strata. Due to both sample constraints and missing data, the final number of

[3]Residential public utilities include water supply, sewerage, sanitation, electricity, and gas.

observations we used is 8,884 (heads of household). The ECV2014 was chosen because it includes data on private households related to several of the dimensions which are of interest to the study of quality of life.

3.2. Variables

Appendix 1 lists the variables used in the study. Table 1 shows the descriptive statistics for all the variables. In what follows, we present the variables used in our models and justify their inclusion in the empirical analysis.

Insert Table 1 here

3.2.1. Self-Reported Quality of Life

The survey contains data about individuals' perceived quality of life. More specifically, we use the responses to the question: 'On a scale from 1 to 5, rate the quality of life of the members of your household'. The response options were 1 *Very bad*, 2 *Bad*, 3 *Acceptable*, 4 *Good* and 5 *Very good*. As mentioned above, a question of this type refers to individuals' own criteria and overall evaluation of life (Veenhoven 2017, p. 7). As a result, we used the responses to this question as a proxy of self-reported quality of life. The Colombian authorities are interested in knowing whether the citizens of Medellin consider they have a good quality of life in order to focus their efforts on more disadvantaged citizens. In reference to this, we collapsed the ordinal variable into a binary that takes the value of 1 if the respondent perceives his or her quality of life as good or very good, and 0 otherwise. Some studies have concluded that when the ordinal scale is collapsed into a binary, the logistic regression yields similar results which implies only a slight reduction in power (Armstrong and Sloan 1989; Manor et al. 2000). In the happiness literature, several studies use the dependent variable in dichotomous form (see Medina et al. 2008; Pedersen and Schmidt 2011). We denote this variable as *Quality of life*.

3.2.2. Explanatory variables

Socio-economic characteristics

We control for the socio-economic characteristics that are common in the subjective well-being literature: *Male* (gender), *Age, Race, Living partner, Illiteracy, Secondary, Tertiary, Good health,* and *Permanence in employment*. With regard to the variable *Permanence in employment*, it must be highlighted that this is a continuous variable referring to the number of months the respondent has been working for a company or is self-employed, either in the formal or the informal sector. In our sample, the average number of months that individuals have been

working in the same company is 47, and ranges from 0 (unemployed) to 632 months of employment (Table 1).

Economic resources

Larger economic resources are expected to be associated with greater well-being due to the benefits of higher prosperity. However, the relationship between economic resources and self-reported quality of life is not as straightforward as initially thought (for a review, see Clark et al. 2008; Ferrer-i-Carbonell 2005; Inglehart et al. 2008). For the specific context of Latin America, several studies have shown that economic resources are positively associated with quality of life but are not the most important determinant (see Medina et al. 2008; Rojas 2011).

We used monthly household consumption expenditure as a proxy of economic resources because the ECV2014 does not include a specific question about income. The question is: *What are your total monthly household consumption expenses? (in dollars)*. Household expenses (or income) rather than personal expenses are normally used because they are a better indicator of an individual's real access to economic resources (Ferrer-i-Carbonell 2005; Oshio 2017). In order to control for differences in household size and economies of scale, we have applied an equivalence scale recommended by OECD in which we consider equivalent consumption expense as the household consumption expenses divided by the square root of the number of household members.

Subjective safety

The subjective evaluation of security has an impact on the evaluation of subjective well-being (Dolan et al. 2008; Wills-Herrera et al. 2011). It is convenient to distinguish two dimensions of perceived safety when the scope of study is a city: neighbourhood safety and personal safety. Thus, we consider two variables: *Neighbourhood safety* and *Forced displacement*. The former is a dummy variable that takes the value of 1 if citizens feel *Very safe* or *Safe* living in the neighbourhood, and 0 otherwise. A positive association between *Neighbourhood safety* and quality of life is expected (Powell and Sanguinetti 2010). When people perceive that they live in a safe neighbourhood free of crime and violence they tend to report a higher level of quality of life because the neighbourhood provides a stable living environment. In contrast, living in an insecure neighbourhood causes anxiety and feelings of unease, thus leading to lower perceived subjective well-being (Chong et al. 2017).

The second variable within the subjective safety dimension is *Forced displacement.* This is also a dummy variable that takes the value of 1 if respondents have had to move from their former municipality of residence for any of the following public order causes: extortion,

kidnapping, pressure from armed groups, or threat of common delinquency. The rationale for including this variable is that in 2016, Colombians were the second largest group in the world (after Syrians) with 7.7 million people forcibly displaced by conflict and violence. Most displacements occurred within the country (7.2 million), generally from rural areas to larger urban areas.[4] Medellin, along with Bogotá and Cartagena, became the main destinations for involuntary migration flows given that victims seek big cities in order to remain anonymous as a way to ensure greater security. Forced displacement has negative effects on quality of life because people's family and social ties break down, they lose their possessions, and they also have to re-enter the labour market. This implies that displaced people are more exposed to the risk of poverty (Sánchez Mojica 2013).

Social and cultural capital

Muffels and Headey (2013) distinguished between social capital, defined as the level of trust in other people and the capacity of people to build a social network, and cultural capital, defined as no materialistic values that influence people's achievements and outcomes. We use membership in associations and organizations as a proxy of social capital. We incorporate the dummy variable *Social capital*, which takes the value of 1 if the respondent is a member of at least one of the organizations in a list of 11, and 0 otherwise.[5] As shown in Table 1, 11.1% of our sample is involved in such associations. Taking into account that social connections might enable individuals to access valuable resources, such as affective support, information on employment opportunities, or possibilities for association to develop productive projects, several studies have found positive effects of social capital on subjective well-being (Bárcena-Martín et al. 2017; Bartolini and Sarracino 2014; Han 2014; Neira et al. 2018; Oshio 2017) and also in Latin American countries (see Ateca-Amestoy et al. 2014; Wills-Herrera et al. 2011).

Given that perceived freedom is a value that influences people's achievements and outcomes (Sen 1999), we consider perceived freedom to express thoughts and political ideas as a proxy of cultural capital. We use the dummy variable *Freedom* which takes the value of 1 if the respondent perceives that *There is a lot of freedom*, *There is freedom* and *Acceptable*, and 0 if the respondent perceives *Little freedom* and *Very little freedom*. The freedom to express thoughts and political ideas could encourage individuals to become involved in national politics and environmental protection, which constitute cultural capital. How much freedom or

[4]Estimates provided by the Internal Displacement Monitoring Centre (http://www.internal-displacement.org/); the source used by the World Bank in the displacement statistic.

[5]These organizations include parents' associations, corporations, local administration boards, women's associations, youth groups, citizen oversight committees, community action boards, clubs for the elderly, neighbourhood assemblies, communal budget councils, and community health participation committees.

perceived freedom people feel is positively associated with subjective well-being (Inglehart et al. 2008; Okulicz-Kozaryn 2015; Rahman and Veenhoven 2018).

Location

The conditions of the location where people live, such as housing and neighbourhood attributes, are associated with subjective well-being (Farrel et al. 2004; Gandelman et al. 2012). To analyse how location attributes affect self-reported quality of life, we include three variables. The dummy variable *Long-term resident* indicates whether the respondent has always lived in the city of Medellin. The variable *Time in neighbourhood* denotes the number of years that the respondent has lived in the same neighbourhood. Lastly, the dummy variable *Owner* refers to families who are outright home owners. A positive association between these variables and quality of life would be expected, since they would positively contribute to the promotion of neighbourhood cohesion (Robinson and Wilkinson, 1995). Additionally, home ownership has a positive effect on subjective well-being because there is a positive feeling of living in one's own place, which is perceived as better than living in rented accommodation (D'Ambrosio and Frick 2012).

4. Empirical approach

In order to assess the socio-economic stratification policy of Medellin as a guideline for the public provision of residential utility services, the implementation of differential strata-based utility rates, the allocation of subsidies, and tax collection, we test three hypotheses.

The **first hypothesis** is that since the composition of the socio-economic strata in Medellin is not random but, as stated above, has been established based on certain characteristics that affect perceived quality of life, the structure of data on self-reported quality of life in Medellin could be hierarchical or nested into the six strata identified by the Colombian government. We consider the possibility that two individuals randomly selected from the same stratum will report a more similar level of quality of life than two individuals selected randomly from different strata. In other words, we assume no independence among respondents belonging to the same stratum. To test this hypothesis, multilevel models should be used. In a classical one-level model it is assumed that the observations are independent and the error is treated as noise, so the estimate should minimize the error. However, when the data is nested, the correlation between observations within a group could be different than the correlation between groups, resulting in two types of errors. An advantage of multilevel models is that they analyze what part of the random error is due to the effect of level 2 (socio-economic strata) and what part is due to level 1 (individuals, that would be the classical regression) (see Goldstein

2011; Snijders and Bosker 2012). That is, multilevel modelling will allow us to know what part of the variability in the self-reported quality of life would be explained by socio-economic strata. Following specification 1, we estimate the null model (Model 0), which does not include any explanatory variables.

The **second hypothesis** is that the strata ranking (from 1 low-low to 6 high) might not accurately reflect the standard of living of citizens in Medellin. Therefore, if the Colombian authorities guide their policy of subsidies and municipal taxes based on the current socio-economic strata, the principle that better-off families would contribute to improving the situation of the most disadvantaged households is not fully upheld. From the literature review we know that in addition to the factors taken into account for the stratification, other economic and non-economic factors are also determinants of individuals' quality of life. The multilevel approach allows us to empirically contrast this hypothesis by estimating random intercept models in which we incorporate these previously identified factors (models 1 to 5 with specification 1).

The **third hypothesis** is that changes in households' economic resources do not have the same effect on self-reported quality of life for households belonging to different strata. Moreover, if hypothesis 2 is confirmed, it may not be the case that the higher the socio-economic stratum in the classification, the less effective the policies to increase families' economic resources (i.e. the increases in self-reported quality of life would be lower). In this case, we estimate a multilevel random intercept and random slope model (model 6 with the specification 2).

Given the discrete nature of our dependent variable *Quality of life*, we estimate logit multilevel models (LMLM) which take the binomial distribution as a reference.

Specification 1: Multilevel random intercept model

We consider a two-level structure where individuals *i* (level 1) are nested into strata *j* (level 2). The random intercept model accounts for stratum differences in quality of life. In this specification, the intercept varies randomly between the different strata, but the slope is the same for all strata. Let us consider QL^*_{ij} as a latent variable of quality of life that indicates the probability that an individual *i* of stratum *j* will report a good or very good quality of life (*QL* = 1), where $i \in \{1,..., 8{,}884\}$ and $j \in \{1,...,6\}$. For each observation located in the j-stratum the model can be written as follows:

$$QL^*_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij} \quad (1)$$

where $\beta_{0j}=\beta_0+u_j$, $x_{ij} \in X$, being X a nXm-dimensions matrix of observed explanatory variables at the individual level, and β_1 its associated parameters. For stratum *j*, the intercept is β_{0j}, which may be smaller or larger than the intercept of population β_0. The individual residuals (with nXm dimensions) are denoted by e_{ij} and the stratum random effects are denoted by u_j. The residuals u_j are assumed to have a normal distribution of zero mean and variance σ_u^2. In order to identify the fixed and random parts of the model, Equation (1) can be written as:

$$QL_{ij}^{*} = \beta_0 + \beta_1 x_{ij} + e_{ij} + u_j \tag{2}$$

In this equation, the fixed part of the model shows the relationship between the mean of *QL* and the explanatory variables ($\beta_0+\beta_1 x_{ij}$ with parameters β_0, β_1), and the random part captures the residuals from different levels ($e_{ij}+ u_j$ with parameters σ_e^2 , σ_u^2).

Following this specification, we estimate the null model or Model 0 (without explanatory variables) and Models 1 to 5 which successively include the different groups of explanatory variables. More specifically, Model 1 includes socio-economic characteristics (*Male*, Age, *Race, Living partner, Illiteracy, Secondary, Tertiary, Good health, and Permanence in employment*). In addition to the variables in Model 1, Model 2 includes economic resources (*Consumption*). In addition to the variables in Model 2, Model 3 includes variables that capture reported safety (*Neighbourhood safety* and *Forced displacement*). In addition to the variables in Model 3, Model 4 includes social and cultural capital (*Social capital* and *Freedom*). Lastly, Model 5 includes variables of location (*Time in neighbourhood*, *Long-term resident*, and *Owner*) in addition to the variables in Model 4.

Specification 2: Random slope model for *Consumption* variable

The specification of this model is an extension of the random intercept model which also considers that the slope for the variable *Consumption* varies randomly among the different strata. Let QL^{*}_{ij} a latent variable of quality of life that indicates the probability that an individual *i* of stratum *j* will report a good or very good quality of life (*QL* = 1), where $i \in \{1,..., 8{,}884\}$ and $j \in \{1,...,6\}$. For each observation located in the j-stratum the model can be written as follows:

$$QL_{ij}^{*} = \beta_{0j} + \beta_{1j} c_{ij} + \beta_2 x_{ij} + e_{ij} \tag{3}$$

where $\beta_{0j}=\beta_0+u_{0j}$ and $\beta_{1j}=\beta_1+u_{1j}$; $x_{ij} \in X$, being X a nXm-dimensions matrix of observed explanatory variables at the individual level, and β_2 its associated parameters. The variable *Consumption* is denoted by *c*. The average regression for *Consumption* has slope β_1 and the slope for each stratum is β_{1j}. The random errors u_{0j} and u_{1j} are assumed to have a normal distribution of zero

mean and variance σ^2_{u0} and σ^2_{u1}, respectively. Model 6 is estimated following this specification. Developing equation 3, we can identify the fixed and random parts of the model:

$$QL^*_{ij} = \beta_0 + \beta_1 c_{ij} + \beta_2 x_{ij} + e_{ij} + u_{0j} + c_{ij} u_{1j} \quad (4)$$

In this equation (4), the fixed part of the model shows the relationship between the mean of *QL* and the explanatory variables ($\beta_0 + \beta_1 c_{ij} + \beta_2 x_{ij}$ with parameters $\beta_0, \beta_1, \beta_2$), and the random part captures the residuals from different levels ($e_{ij}+u_{0j}+c_{ij}u_{1j}$ with the parameters σ^2_e, σ^2_{u0}, σ^2_{u1}; where $c_{ij}u_{1j}$ is the interaction between the stratum and *Consumption*).

In the framework of multilevel models, the marginal R-squared (R^2m) represents the variance explained by fixed factors of the model (or individual characteristics) and the conditional R-squared (R^2c) represents the variance explained by fixed and random factors (see Nakagawa and Schielzeth 2013). The difference between the corresponding R^2c and R^2m values reflects the amount of variability in the random effects or across strata (contextual effects).

5. Results

5.1. Cross-tabulation analysis of quality of life and strata

In order to empirically test whether self-reported quality of life can be determined by the strata to which individuals belong (Hypothesis 1), it is useful to analyse the relationship between quality of life and strata by cross tabulation prior to estimating the multilevel regression model. Firstly, a chi-square test was performed ($X^2 = 276.875$, df = 5, $p < 0.001$) to determine the presence of statistical independence between *Quality of life* and the socio-economic strata to which individuals belong. The test shows that there is significant dependence between *Quality of life* and strata given that the null hypothesis that the two categorical variables are independent is rejected.

Secondly, Table 2 shows the cross-tabulation analysis for both variables. Depending on the stratum, the number of surveys varies from 201 for stratum 6 to 2,957 for stratum 2. Stratum 6 registers the highest *Quality of life*, with 96% of citizens reporting very good or good quality of life, while stratum 1 registers the lowest, with 35.9% of citizens reporting very bad, bad, or acceptable quality of life. The table also shows the contribution of the chi-square of each stratum to the total chi-square ($X^2 = 276.875$). As can be observed, the highest contribution corresponds to respondents in stratum 1 who report a low quality of life ($X^2 = 71.539$). This provides strong evidence that the number of dissatisfied respondents in stratum 1 differs from the expected number, as well as a stronger relationship between *Quality of life* and stratum. In contrast, the smallest chi-square contribution corresponds to satisfied respondents in stratum

3 (X^2 = 1.389), which in this case indicates the weakest relationship between *Quality of life* and stratum.

Insert Table 2 here

Third, as a complement to the cross-tabulation analysis, Figure 1 shows *Quality of life* for each stratum in graphic form. In the figure, the width of each rectangle corresponds to the number of surveys and the two heights in different colours represent the number of respondents that perceive quality of life as good or very good (*Quality of life* = 1) or as very bad, bad, or acceptable (*Quality of life* = 0), respectively. The colours denote the *p*-value of the Pearson test to determine whether there is independence between the stratum to which an individual belongs and the quality of life he or she reports. Residual Pearson's cells below -2.0 and above 2.0, respectively, indicate that there are fewer or more observations in that cell than those expected under the null hypothesis (independence between variables).[6] As can be seen, except for stratum 3, there is no independence between the variables in the rest of the strata. Similarly, there is strong evidence of dependence between *Quality of life* and stratum in strata 1 and 5 (the *p*-values are further from -2 and 2, respectively, and the colours are the most intense).

Insert Figure 1 here

These findings justify that perceived quality of life must be analysed taking into account the stratum to which individuals belong.

5.2. Multilevel analysis

Based on the previous results, the use of multilevel models may be suitable to account for the differential effects of the strata on *Quality of life* (Hypothesis 1), to analyse the effect that other variables identified in the happiness literature might have on *Quality of life* (Hypothesis 2), and also to study whether changes in *Consumption* have a different effect on *Quality of life* across strata (Hypothesis 3). The results of the estimations for the random intercept models (Model 0 to Model 5) and random slope model (Model 6) are shown in Table 3. The table also indicates the effect of each variable on the response probability, which is given by the odds ratios. Several indicators of the model fit are reported at the bottom of the table.

Insert Table 3 here

We start by estimating the null model to explain quality of life taking into account only the effect of the strata (Table 3, Model 0). This is a LMLM which does not include any explanatory

[6] The heuristic for choosing the cut-off point equal to 2 (-2) is that the Pearson residuals are approximately standard normal, which implies that the individual residuals of the highlighted cells are significant at approximately the $\alpha = 0.05$ level.

variables. We can conclude that there is a significant variation in citizens' perceived quality of life across strata because the likelihood ratio test (X^2 = 268.0), which contrasts the LMLM against the classical one-level logit model, is significant for testing the null hypothesis that σ_u^2 = 0. In other words, there are differences across strata and the LMLM, which would be more suitable. In the same line, the intra-class correlation coefficient (ICC = 0.167) indicates that approximately 16.7% of the variability in reported quality of life is attributable to differences across strata.

The random intercept effects by strata can also be examined graphically. Figure 2 (Model 0) displays the estimated average level of reported quality of life in each stratum with 95% confidence intervals (points and horizontal lines). When explanatory variables are not taken into account, the perceived quality of life of people in strata 1, 2 or 3 is below the overall average of the one-level model (vertical line equal to 0), while it is above average for people in strata 5 and 6. However, people in stratum 4 seem to perceive quality of life in line with the overall average. In Figure 2 it can also be observed that the 95% confidence intervals corresponding to strata 2 and 3 are very low compared to the confidence interval of stratum 6, thus indicating a less significant dispersion in the estimation of random effects in strata 2 and 3 than in stratum 6.

Insert Figure 2 here

Models 1 to 5 progressively incorporate the different groups of explanatory variables analysed in this study. The results of Model 1 reflect that, with the exception of *Age* and gender (*Male)*, the other socio-economic variables are statistically significant. The signs of their estimated parameters are similar to those of previous studies for Latin America (Ateca-Amestoy et al. 2014; Cid et al. 2008; Garcia et al. 2007; Hurtado 2016). *Good health* is the factor which has the strongest positive effect on perceived quality of life. A positive association was also found between quality of life and marital status due to partners' support. Having attained a higher level of education is associated with a greater probability of reporting a very good and good level of quality of life, while being illiterate yields precisely the opposite result. The longer people have been employed, the greater the likelihood of reporting high levels of quality of life, although the odds ratio close to a value of 1 indicates that the quantitative significance of the coefficient is low. And, lastly, mixed race, Afro-Colombian, and Afro-descendant citizens are less likely to report good or very good quality of life. These same findings are reached in all the estimated models.

The result of the likelihood ratio (LR) test shows that Model 1 is an improvement over Model 0. The values of marginal R-squared (R^2m) and conditional R-squared (R^2c) indicate that if we control only for socio-economic characteristics, random effects or variability across strata explain approximately 40% ([0.157 - 0.094]/0.157) of the variance of self-reported quality of life.

Model 1 and Model 0 in Figure 2 reflect the same successive strata ranking from 1 (lowest) to 6 (highest).

Model 2 (Table 3) controls for both socio-economic characteristics and the variable *Consumption*. This new variable shows high quantitative significance and its sign is positive, thus, as in the literature reviewed for Latin America, an increase in household consumption expenditure increases the probability of reporting good or very good quality of life. In addition, it significantly improves the specification of Model 1, as shown by the statistical LR test. Both R^2m and R^2c have increased with respect to Model 1. However, in Figure 2 (Model 2) it can be observed that citizens in stratum 5 perceive a higher level of quality of life than those in stratum 6 (the intercept of stratum 5 is larger than the intercept of stratum 6). In addition, the perceived quality of life of citizens in strata 1 and 2 continues to be significantly below the overall average in contrast to those in stratum 5. The same behaviour was maintained in Models 2 to 5. Consequently, the inclusion of the variable *Consumption* as a proxy of economic resources changes the order of perceived quality of life of the citizens in strata 5 and 6.

The successive incorporation of new variables related to subjective safety, social and cultural capital, and location improves the estimates: the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) decrease, while R^2m and R^2c increase (see Table 3, Model 3, Model 4, and Model 5). In these models, *Consumption* remains significant and the estimated value of its parameter is high with odds ratio greater than one. The two variables related to subjective safety (*Neighbourhood safety* and *Forced displacement*) are also statistically significant, with high quantitative significance and the signs of their estimated parameters are as expected. The safer citizens feel, the more likely they are to report good or very good quality of life. In contrast, when individuals have had to change their place of residence due to problems of public order, they are less likely to report good or very good quality of life. Moreover, people who relate to others and are involved in associations (*Social capital*) and feel free to express political ideas (*Freedom*) are more likely to report good or very good quality of life. As regards location, it is worth noting that living all life in Medellin does not affect self-reported quality of life. However, the longer people remain in the same stratum, the less likely they are to report good or very good quality of life. In any case, the odds ratio of the variable *Time in neighbourhood* is very close to 1, indicating that the quantitative importance of its parameter is low. Consistent with previous studies, being a homeowner has a positive effect on self-reported quality of life.

5.3. Relationship between *Quality of Life* and *Consumption* by stratum

In order to check Hypothesis 3, Figure 3 shows the relationship between *Quality of life* and *Consumption* for each of the six strata. The logit models are shown in blue and represent the probability of being satisfied with life considering the level of monthly equivalent consumption expenditure per household (*Consumption*). The graphs in the figure indicate that the probability of being satisfied with life rises with an increase in *Consumption*. However, as can be observed, this behaviour is not the same for all six strata. In fact, the odds ratio of strata 1 and 2 are two times higher than the values of the rest of the strata, thus indicating that a change in *Consumption* does not have the same effect on *Quality of life* if a citizen belongs to a high or a low stratum. Therefore, if *Consumption* is used as a proxy for household economic resources, a change in this factor for the inhabitants in the lowest strata (1 or 2) would have a greater effect on their perceived quality of life than if this change occurred in the other strata. Hence, it would be interesting to analyse *Quality of life* taking into account not only the stratum to which an individual belongs (Specification 1), but also the interaction with the variable *Consumption* (Specification 2).

Insert Figure 3 here

In addition, as Figure 2 (model 6) indicates, the effect of *Consumption* on *Quality of life* is not the same if the individual belongs to a low or a high stratum. Therefore, to determine whether *Consumption* could affect each of the strata differently, we estimated a random slope model for each stratum (Table 3, Model 6). This model is more robust than Model 5, since the AIC value is lower and R^2c and R^2m are higher. Likewise, the LR test indicates that there are significant differences with respect to Model 5. In Model 6, the explanatory variables show the expected behaviour, which is similar to that of Model 5. Random effects or variability across strata explain approximately 10.26% of the variance of self-reported quality of life.

As can be observed in Model 6-intercept of Figure 2, the perceived quality of life of citizens in strata 1 and 2 is significantly lower than the overall average, while it is above average for those in stratum 4. Furthermore, Model 6- *Consumption* in Figure 2 shows that the effect of a variation in *Consumption* on the perceived quality of life of individuals in strata 1 and 2 is above the overall average, while the effect of this variation lowers the overall average for individuals in stratum 4. Nevertheless, these effects are not significantly different from the overall average for the individuals in strata 3, 5 and 6. These three strata show a higher dispersion in the 95% confidence intervals. In other words, in strata 1 and 2, where the likelihood of reporting good or very good quality of life is lower, an increase in household consumption expenditure would have a stronger effect on self-reported quality of life. Conversely, individuals in stratum 4 are more likely to report good or very good quality of life; however, an increase in household expenditure would produce the lowest increase in the level of self-reported quality of life. In addition, as

shown in Figure 2, the order of the strata has changed with regard to the previous order maintained in Model 2 to Model 5.

6. Conclusions and discussion

The main objective of this paper was to analyse whether the Colombian government's classification of households in socio-economic strata accurately reflects the level of quality of life of households in Medellin. To this end, we formulated three hypotheses, which have been verified in the analysis. Thus, the answer to our initial research question is that the socio-economic strata do not accurately reflect citizens' quality of life. This finding is particularly important given that the socio-economic stratification system of Colombia is used to guide public policies at the local level, namely, to differentially charge for residential public utility services such that households in the higher strata pay more to subsidize utility rates for households in the lower strata.

In what follows, we provide an overview of the main conclusions arising from the three hypotheses. We tested for the existence of a stratum effect such that families living in the same stratum report more similar levels of quality of life (Hypothesis 1). That is, the factors taken into account by the Colombian authorities to carry out the stratification (characteristics of the dwellings and the environment, as well as housing prices) affect citizens' quality of life. However, when considering the individual factors reported in the literature on subjective well-being, we find a significant variation in citizens' perceived quality of life across strata (Hypothesis 2). The ranking of self-reported quality of life according to the strata established by the Colombian authorities ranges from 1 low-low to 6 high, whereas in our study the strata rank from low to high self-reported quality of life in the following order: 1, 2, 6, 3, 5, 4. In this regard, some studies have identified a cluster with high homicide rates in stratum 6 corresponding to the city centre of Medellin, as well as the highest rates of assault and motor vehicle and property theft in areas with the highest socio-economic status (Gaviria et al. 2010; Medina et al. 2008). Conversely, these studies have identified a cluster of low homicide rates in stratum 4, which occupies the highest position in our quality-of-life ranking. Likewise, in the case of the two most depressed strata (1 and 2), our results indicate that public policies targeted at increasing families' level of consumption expenditure would be very effective in improving citizens' quality of life (Hypothesis 3). It is important to recall that these two strata host people internally displaced by violence; a collective which is more exposed to the risk of poverty.

Focusing on the determinants of quality of life in Medellin, our study identifies several non-economic factors, such as perceived freedom, subjective safety, and social capital, which are particularly significant in the context of a city in a country that has suffered armed conflict

for decades. Given that the conflict in Colombia is a political one where individuals fear expressing their political views (Wills-Herrera et al. 2011), it follows that citizens' perceived freedom to express political ideas is associated in a positive and very significant way with quality of life. Neighbourhood safety is positively correlated with self-reported quality of life, while personal insecurity (measured as forced displacement) is negatively correlated. Belonging to an association that fosters social contacts is also positively associated with perceived quality of life. Similarly, several studies have argued that these factors are complementary to each other in the sense that a safe neighbourhood provides a space of trust for people to interact with one another, thus leading to higher levels of social capital. In turn, these trusting relationships with neighbours could increase the perceived safety of the neighbourhood (see Chong et al. 2017; Wills-Herrera et al. 2011).

In summary, our findings are highly consistent with the concept of quality of life presented in the previous section and lead us to argue that if the objective of public policy is to create an environment for citizens to lead satisfying lives and enhance their quality of life, then public policy measures should take a different direction. Firstly, given that there are striking differences between the current socio-economic stratification system and our ranking of strata in Medellin, it would seem that the system needs to be revised in such a way as to design municipal public policies that truly contribute to improving the situation of the most disadvantaged households. Secondly, in order to improve the quality of life in Medellin, public policies should focus on promoting the traditional determinants of quality of life (consumption, good health, and education), as well as other social resources (subjective safety, perceived freedom, and social capital), which will undoubtedly require more innovative measures.

The empirical and graphical analysis method followed in this study could be extrapolated to other cities in different countries since the underlying idea is that the exchange of relationships and social support between people living in close proximity shapes quality of life. In an additional way, it would be key to identify the determinants of quality of life according to the history, socio-economic situation, cultural values, etcetera, of each country. In this vein, a limitation of our work is that due to the absence of statistical information, we have not included variables referring to psychological capital or personality traits of respondents.

For future studies, we will focus more deeply on the analysis of the association between spatial relationships and quality of life with spatial econometrics techniques. We could study the existence of spatial dependence in quality of life, so that a portion of the quality of life of a respondent might be explained not only by his or her drivers but also by the quality of life of his or her neighbourhood. The presence of spatial clusters in high and low quality of life could be also studied. If spatial dependence on the quality of life were confirmed, public policies should

explicitly incorporate spatial information and be targeted to account for personal inequalities in quality of life.

Table 1. Descriptive statistics of quality of life in Medellin, 2014

	Mean	SD	Minimum	Maximum
Quality of Life	0.762	0.426	0	1
Male	0.517	0.500	0	1
Age	53.778	16.446	17	103
Race	0.036	0.185	0	1
Living partner	0.531	0.499	0	1
Illiteracy	0.027	0.162	0	1
Secondary	0.460	0.498	0	1
Tertiary	0.110	0.312	0	1
Good health	0.768	0.422	0	1
Permanence in employment	47.067	92.831	0	632
Consumption	0.661	0.691	0.016	15
Neighbourhood safety	0.887	0.317	0	1
Forced displacement	0.046	0.211	0	1
Social capital	0.111	0.314	0	1
Freedom	0.921	0.269	0	1
Time in neighbourhood	29.098	22.424	0	100
Long-term resident	0.730	0.444	0	1
Homeowner	0.545	0.498	0	1

Note: N = 8,884. Adapted from Administrative Department of Planning of Medellin, Quality of Life Survey of 2014 (Encuesta de Calidad de Vida de 2014).

Table 2. Cross-tabulation analysis of quality of life and socio-economic stratum in Medellin, 2014

		Socio-economic stratum						
Quality of life		1	2	3	4	5	6	Row total
0	n	420	828	643	153	63	8	2,115
	X^2	71.539	21.853	4.445	27.914	52.019	33.189	
	%	35.9	28.0	21.9	15.6	9.9	4.0	
1	n	751	2129	2292	829	575	193	6,769
	X^2	22.353	6.828	1.389	8.722	16.253	10.370	
	%	64.1	72.0	78.1	84.4	90.1	96.0	
Column total		1,171	2,957	2,935	982	638	201	8,884

Note: *Quality of life* equal to 1 indicates that respondents perceive their quality of life as good or very good; *Quality of life* equal to 0 indicates that respondents perceive their quality of life as very bad, bad, or acceptable. Cell contents: numbers of surveys (n), chi-square contribution (X^2) and n/column total *100 (%). Total observations: 8,884.

Table 3. Determinants of perceived quality of life and socio-economic stratum effect in Medellin, 2014

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects							
Intercept	1.6105***	0.2625	-0.1702	-0.6955*	-0.9641**	-0.9260**	-1.004**
Socio-economic characteristics							
Male	--	0.1149 (1.122)	0.0911 (1.095)	0.0873 (1.091)	0.0970 (1.102)	0.0949 (1.099)	0.0853 (1.089)
Age	--	-0.0056 (0.9944)	-0.0069 (0.993)	-0.0062 (0.994)	-0.0067 (0.993)	-0.0105 (0.990)	-0.0110 (0.989)
Age2	--	0.0001 (1.000)	0.0001 (1.000)	0.0001 (1.000)	0.0001 (1.000)	0.0001 (1.000)	0.0001 (1.000)
Race	--	-0.3934** (0.675)	-0.3802** (0.684)	-0.3585** (0.699)	-0.3582** (0.699)	-0.3432** (0.709)	-0.3485** (0.706)
Living partner	--	0.2188*** (1.245)	0.2264*** (1.254)	0.2275*** (0.699)	0.2292*** (1.258)	0.2247*** (1.252)	0.2309*** (1.260)
Illiteracy	--	-0.4805*** (0.619)	-0.4658** (0.628)	-0.4567** (0.6334)	-0.4464** (0.640)	-0.4559** (0.634)	-0.4303** (0.650)
Secondary	--	0.3305*** (1.392)	0.2954*** (1.344)	0.2794*** (1.322)	0.2749*** (1.316)	0.2774** (1.320)	0.2682*** (1.308)
Tertiary	--	0.8557*** (2.3530)	0.7057*** (2.025)	0.7091*** (2.032)	0.6882*** (1.990)	0.6510*** (1.917)	0.6836*** (1.981)
Good health	--	1.0280*** (2.796)	1.0190*** (2.770)	0.9798*** (2.664)	0.9770*** (2.656)	0.9651*** (2.625)	0.9665*** (2.628)
Permanence in employment	--	0.0011*** (1.001)	0.0010*** (1.001)	0.0010** (1.001)	0.0011** (1.001)	0.0011** (1.001)	0.0011** (1.001)
Economic resources							
Consumption	--	--	0.6061*** (1.833)	0.6079*** (1.837)	0.6081*** (1.837)	0.6279*** (1.874)	0.7913*** (2.206)
Subjective safety							
Neighbourhood safety	--	--	--	0.6892*** (1.992)	0.5980*** (1.818)	0.6096*** (1.840)	0.6138*** (1.847)
Forced displacement	--	--	--	-0.4544*** (0.635)	-0.4643*** (0.629)	-0.4678*** (0.626)	-0.4594*** (1.583)
Social and cultural capital							
Social capital	--	--	--	--	0.2565** (1.292)	0.2429** (1.275)	0.2462** (1.279)
Freedom	--	--	--	--	0.3931*** (1.482)	0.3823*** (1.466)	0.3849*** (1.469)
Location							
Time in neighbourhood	--	--	--	--	--	-0.0048** (0.995)	-0.0049*** (0.995)
Long-term resident	--	--	--	--	--	0.0421 (1.043)	0.0519 (1.053)
Owner	--	--	--	--	--	0.3463*** (1.414)	0.3544*** (1.425)
Random effects (variance)							
Intercept – Social stratum	0.661	0.246	0.055	0.059	0.060	0.057	0.160
Slope – Expenses	--	--	--	--	--	--	0.179
Model fit							
AIC	9486.9	8980.5	8929.0	8839.2	8817.7	8783.6	8772.0
BIC	9501.1	9065.6	9021.2	8945.6	8938.3	8925.4	8928.1
R^2m	--	0.094	0.156	0.169	0.173	0.183	0.210
R^2c	--	0.157	0.170	0.183	0.187	0.197	0.234
logLik	-4741.4	-4478.3	-4451.5	-4404.6	-4391.9	-4371.8	-4364.0
LR test (X^2)	268.0***	526.3***	53.54***	93.80***	25.46***	40.11***	15.56***

Note: N = 8,884. Regression logit multilevel. Entries show parameter estimates with odds ratio in parentheses. *p < 0.1. **p < 0.05. ***p < 0.001.

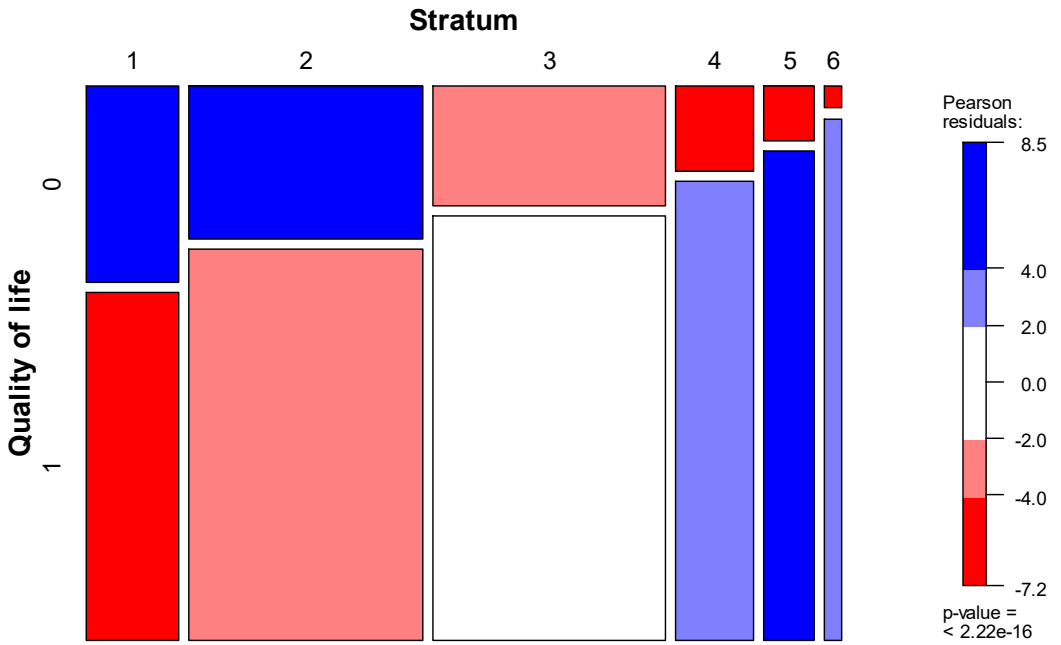

Figure 1. Mosaic chart of cross-tabulation between *Quality of life* and socio-economic stratum. *Quality of life* equal to 1 indicates that respondents perceive quality of life as good or very good. *Quality of life* equal to 0 indicates that respondents perceive quality of life as very bad, bad or acceptable. The Pearson's test was performed to determine independence between *Quality of life* and stratum. Residual Pearson's cells between (-2, 2) indicate independence.

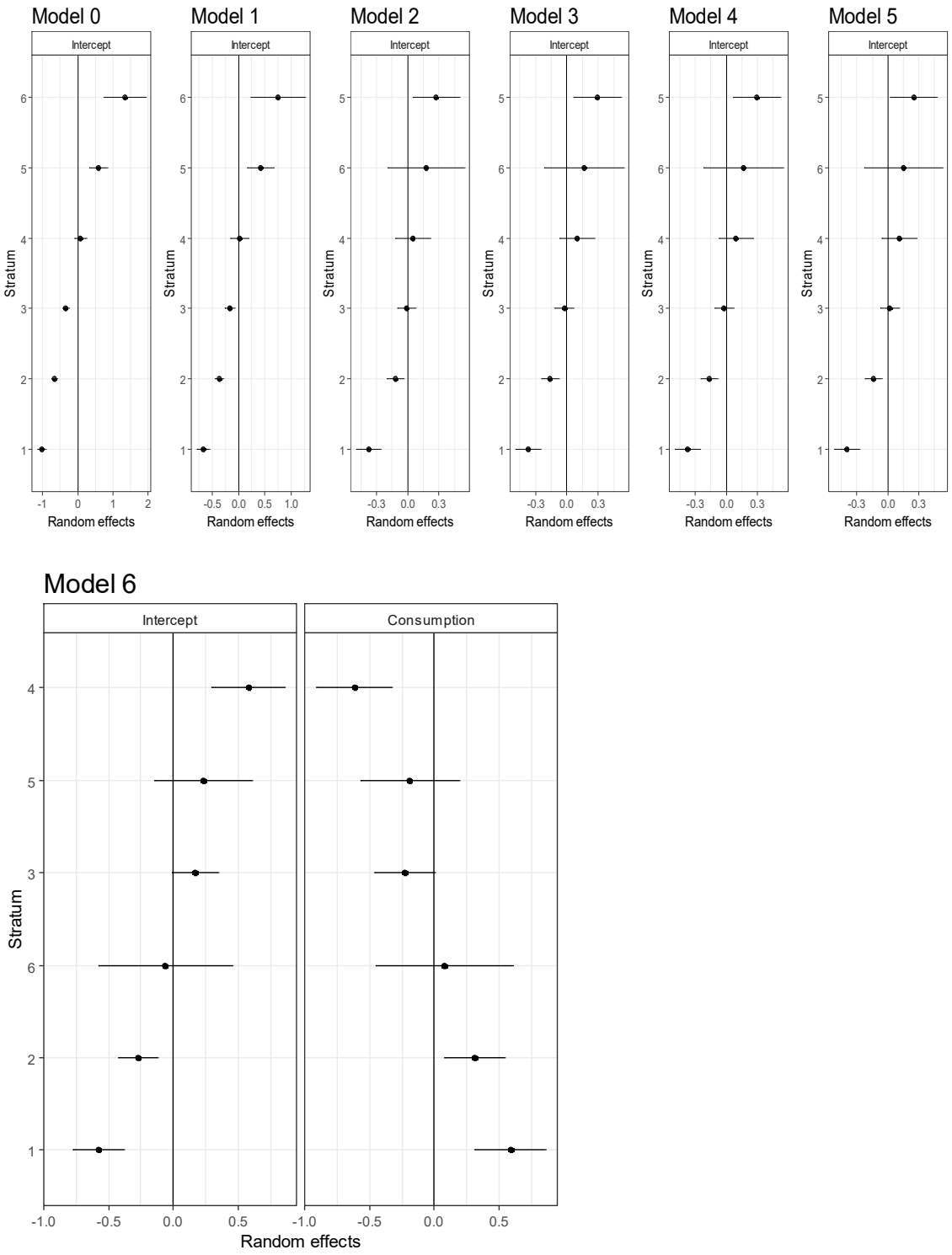

Figure 2. Conditional modes of the random effects. Model 0 to Model 5 are estimated random intercept models. Model 6 is an estimated random slope model for the variable *Consumption*. Horizontal lines represent the confidence interval (95%). The vertical line are the overall average for all the surveys or the intercept in the logit model.

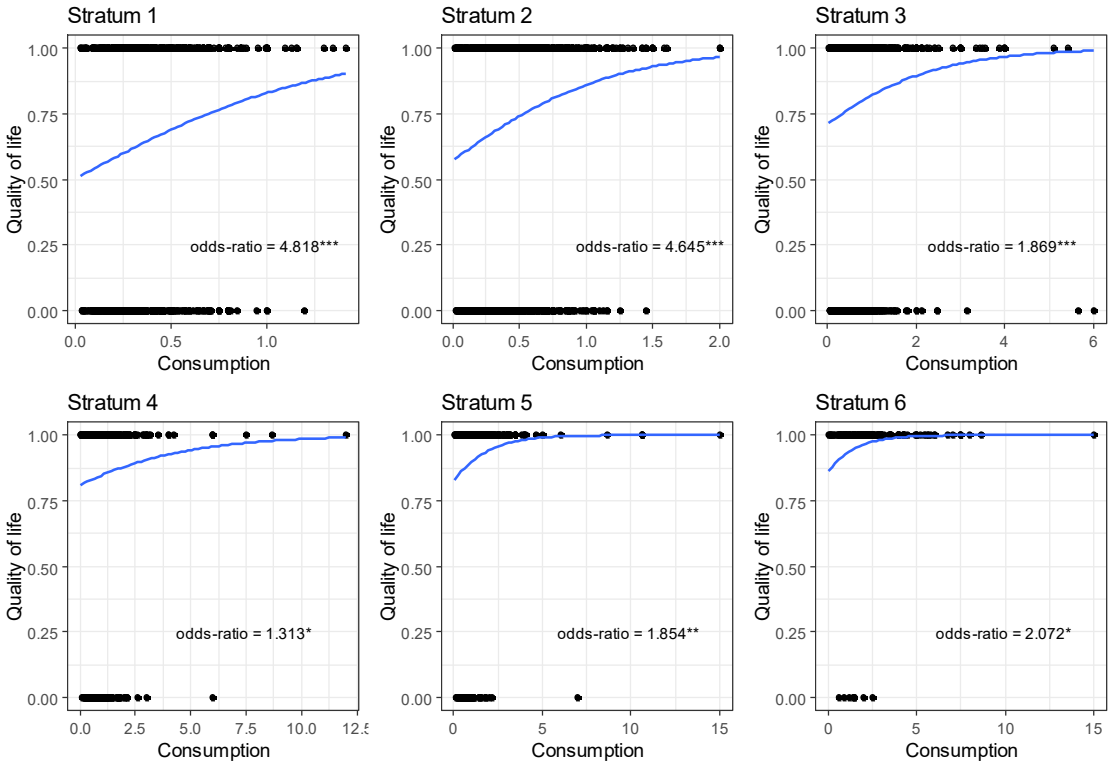

Figure 3. Relationship between *Quality of life* and *Consumption* by stratum. The points in the figure represent the quality of life variable for each respondent (1 or 0) and the blue line represents the estimated logit model with *Consumption* as explanatory variable. *p < 0.1. **p < 0.05. ***p < 0.001.

Appendix 1. Variables to analyse quality of life in Medellin, 2014

Variable	Cases where dummy takes the value of 1
Quality of life	Respondent perceived his/her quality of life as very good or good
Male	1 = Male, 0 = Female
Age	Age in years
Race	Respondent is black, mulatto, Afro-Colombian or Afro-descendant
Living partner	Respondent lives with a partner
Illiteracy	Respondent is illiterate
Secondary	The last approved level of study is secondary
Tertiary	The last approved level of study is tertiary
Good health	Respondent perceives his/her state of health as good or very good
Permanence in employment	Number of months respondent has been working for a company or self-employed (formal and informal sectors)
Consumption	Monthly equivalent expenditure on household consumption
Neighbourhood safety	Respondent feels very safe or safe living in the neighbourhood
Forced displacement	Respondent has moved for any public order causes
Social capital	Respondent is involved in any organization on a list of 11
Freedom	Respondent considers that there is either a lot of freedom or that there is freedom to express political opinions
Time in neighbourhood	Number of years living in the neighbourhood
Long-term resident	Respondent has lived all his or her life in the same municipality
Owner	The house is owned and fully paid

Note: Adapted from Administrative Department of Planning of Medellin, Quality of Life Survey of 2014 (Encuesta de Calidad de Vida de 2014).

Funding acknowledgement

This work was supported by the University of Medellin (Colombia), Project "Modelización Espacial de los Determinantes de la Calidad de Vida en los Hogares de Medellín", under the Grant code 881/2017.

References

Ahumada, G., Iturra, V. &, Sarrias, M. (2019). We Do Not Have the Same Tastes! Evaluating Individual Heterogeneity in the Preferences for Amenities. *Journal of Happiness Studies*. DOI 10.1007/s10902-019-00081-2.

Armstrong, B., & Sloan, M. (1989). Ordinal regression models for epidemiologic data. *American Journal of Epidemiology, 129*(1), 191-204.

Ateca-Amestoy, V., Cortés Aguilar, A., & Moro-Egido, A.I. (2014). Social Interactions and Subjective Well-Being: Evidence from Latin America. *Journal of Happiness Studies, 15*(3), 527–554.

Bárcena-Martín, E., Cortés-Aguilar, A., & Moro-Egido, A. (2017). Social Comparisons on Subjective Well-Being: The Role of Social and Cultural Capital. *Journal of Happiness Studies, 18*(4), 1121-1145.

Bartolini, S., & Sarracino, F. (2014). Happy for how long? How social capital and economic growth relate to happiness over time? *Ecological Economics, 108*, 242-256.

Chong, S.T., Koh, D., Fauziah, I., & Samsudin, A.R. (2017). Neig-hbourhood Social Capital and Neighbourhood Safety in Predicting the Subjective Well-being of Young Malaysians. *Pertanika Journal of Social Sciences and Humanities, 25*, 155-163.

Cid, A., Ferres, D., & Rossi, M. (2008). Subjective Well-being in the Southern Cone: Health, Income and Family. *Documentos de Trabajo* No. 13/08, Departamento de Economía, Universidad de la República.

Clark, A., Frijters, P., & Shields, M. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature, 46*(1), 95–144.

D'Ambrosio, C., & Frick, J.R. (2012). Individual well-being in a dynamic perspective. *Economica*, *79*, 284-302.

DANE (2015). Metodología de estratificación socioeconómica urbana para servicios públicos domiciliarios.: https://www.dane.gov.co/files/geoestadistica/estratificacion/ManualdeRealizacion.pdf

Di Tella, R., & MacCulloch, R.J. (2006). Some uses of happiness data in economics. *Journal of Economics Perspectives, 20*(1), 25-46.

Diener, E., & Chan, M.Y. (2011). Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. *Applied Psychology: Health and Well-Being, 3*(1), 1-43.

Dolan, P., & Metcalfe, R. (2012). Measuring subjective well-being: Recommendations on measures for use by national governments. *Journal of Social Policy, 41*(2), 409-427.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what make us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology, 29*, 94-122.

Farrell, S., Aubry, T., & Coulombe, D. (2004). Neighborhoods and neighbors. Do they contribute to personal well-being? *Journal of Community Psychology, 32*(1), 9–25.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics, 89*, 997-1019.

Frey, B., & Stutzer, A. (2002). What Can Economists Learn from Happiness Research? *Journal of Economic Literature, 40*(2), 402-435.

Frey, B. & Stutzer, A. (2017). Public Choice and Happiness. CREMA (Center for Research in Economics, Management and the Arts), Working Paper 2017-03. Zürich.

Gandelman, N., Piani, G., & Ferre, Z. (2012). Neighborhood Determinants of Quality of Life. *Journal of Happiness Studies, 13,* 547–563. DOI 10.1007/s10902-011-9278-2

Garcia, J.J., Fuentes, N.C., Borrego, S.A., Tapia, A., & Gutierrez, M.D. (2007). Values and happiness in Mexico: The case of the metropolitan city of Monterrey. In L. Bruni, & L. Porta (Eds.), *Handbook on the economics of happiness* (pp. 407–429). Massachusetts: Edward Elgar Publishing Limited.

Gaviria, A., Medina, C., Morales, L., & Núñez, J. (2010). The Cost of Avoiding Crime: The Case of Bogotá. In R. Di Tella, S. Edwards, & E. Schargrodsky (eds), *The Economics of Crime: Lessons for and from Latin America*. University of Chicago: National Bureau of Economic Research (pp. 101-132).

Giordano, G. N., Ohlsson, H., & Lindström, M. (2011). Social capital and health-purely a question of context? *Health & Place, 17*(4), 946–953.

Goldstein, H. (2011). *Multilevel Statistical Models* (4th ed.). West Sussex, UK: John Wiley & Sons, Ltd.

Han, S. (2015). Social capital and subjective happiness: Which contexts matter. *Journal of Happiness Studies, 16*(1), 241–255.

Helliwell, J. F., & Huang, H. (2014). New measures of the costs of unemployment: Evidence from the subjective well-being of 3.3 million Americans. *Economic Inquiry, 52*(4), 1485–1502.

Hurtado, D.A. (2016). Socioeconomic Disparities in Subjective Well-being in Colombia. In M. Rojas (Ed.), *Handbook of Happiness Research in Latin America* (pp. 343-356). Mexico: Springer.DOI 10.1007/978-94-017-7203-7_1

Inglehart, R., Foa, R., Peterson, C., & Welzel, C. (2008). Development, freedom, and rising happiness: A global perspective (1981–2007). *Perspectives on Psychological Science, 3*, 264–285.

Manor, O., Matthews, S., & Power, C. (2000). Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *International Journal of Epidemiology, 29*(1), 149-157.

Medina, C., Morales, L., & Núñez, J. (2008). Quality of Life in Urban Neighborhoods in Colombia: The Cases of Bogotá and Medellín. *Borradores de Economía, 536*. Banco de la República de Colombia.

Meier, S., & Stutzer, A. (2008). Is Volunteering Rewarding in Itself? *Economica, 75*(297), 39–59.

Muffels, R., & Headey, B. (2013). Capabilities and choices: Do they make Sen'se for understanding objective and Subjective Well-Being? An Empirical Test of Sen's Capability Framework on German and British Panel Data. *Social Indicators Research, 110*, 1159-1185.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142.

Neira, I., Bruna, F., Portela, M., & García-Aracil, A. (2018). Individual Well-Being, Geographical Heterogeneity and Social Capital. *Journal of Happiness Studies, 19* (4), 1067–1090. DOI 10.1007/s10902-016-9840-z

Novak, M., & Pahor, M. (2017). Using a multilevel modelling approach to explain the influence of economic development on the subjective well-being of individuals. *Economic Research-EkonomskaIstraživanja, 30*(1), 705-720. DOI: 10.1080/1331677X.2017.1311229

Odermatt, R., & Stutzer, A. (2018). Subjective Well-Being and Public Policy. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being*. Salt Lake City, UT: DEF Publishers. DOI:nobascholar.com(pp.

OˈDonnell, G., Deaton, A., Durand, M., Halpern, D., & Layard, R. (2014). *Wellbeing and Policy*. London: Legatum Institute.

Okulicz-Kozaryn, A. (2015). Freedom and life satisfaction in transition. *Society and Economy 37*(2), 143-164.

Oshio, T. (2017). Which is More Relevant for Perceived Happiness, Individual-Level or Area-Level Social Capital? A Multilevel Mediation Analysis. *Journal of Happiness Studies, 18*, 765–783. DOI 10.1007/s10902-016-9752-y

Oswald, A.J., Proto, E., & Sgroi, D. (2015). Happiness and Productivity. *Journal of Labour Economics, 33*(4), 789-822.

Pedersen, P.J., & Schmidt, T.D. (2011). Happiness in Europe: Cross-country differences in the determinants of satisfaction with main activity. *The Journal of Socio-Economics, 40*(5), 480-489.

Powell, A., & Sanguinetti, P. (2010). Measuring quality of life in Latin America's urban neighborhoods: A summary of results from the city case studies. In E. Lora, A. Powell, B. M. S. van Praag, & P. Sanguinetti (Eds.), *The quality of life in Latin American cities. Markets and perception* (pp. 31-64). Washington, DC: Inter-American Development Bank and the World Bank.

Rahman, A.A., & Veenhoven, R. (2018). Freedom and Happiness in Nations: A Research Synthesis. *Applied Research in Quality of Life, 13*(2), 435-456. DOI 10.1007/s11482-017-9543-6

Robinson, D., & Wilkinson, D. (1995). Sense of community in a remote mining town: Validating a neighborhood cohesion scale. *American Journal of Community Psychology, 23*, 137–148.

Rojas, M. (2011). Income, Happiness, and Beyond. *Applied Research in Quality of Life*, *6*(3), 265-276.

Rojas, M. (2016). The Relevance of Happiness: Choosing Between Development Paths in Latin America. In M. Rojas (Ed.), *Handbook of Happiness Research in Latin America* (pp. 51-62). Mexico: Springer. DOI 10.1007/978-94-017-7203-7_1

Sánchez Mojica, B.E. (2013). A City Torn Apart: Forced Displacement in Medellín, Colombia. *International Law. Revista Colombiana de Derecho Internacional, 22*, 179-210.

Sen, A. (1999). *Development as freedom*. New York: Alfred A. Knopf.

Snijders, T.A.B., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modelling* (2nd ed.). London: Sage Publishers.

Stiglitz, J., Sen, A., & Fitoussi, J. (2011). *Mismeasuring Our Lives: Why GDP doesn't Add Up*. New York: The New Press.

Stutzer, A., & Frey, S. (2010). Recent Advances in the Economics of individual Subjective Well-Being. *Social Research, 77*(2), 679-714.

Van den Bergh, J. (2009). The GDP paradox. *Journal of Economic Psychology, 30*(2), 117–135.

Veenhoven, R. (2000a). Freedom and Happiness. A comparative study in 46 nations in the early 90's. In E. Diener & E. Suh (ed.), Culture and subjective wellbeing. Cambridge MA: MIT press, 257–288.

Veenhoven, R. (2000b). The Four Qualities of Life. *Journal of Happiness Studies, 1*, 1–39.

Veenhoven, R. (2017). Greater Happiness for a Greater Number: Did the Promise of Enlightenment Come True? *Social Indicators Research, 130*(1), 9-25.

Weitz-Shapiro, R., & Winters, M.S. (2011). The Link Between Voting and Life Satisfaction in Latin America. *Latin America Politics and Society, 53*(4), 101–126.

Wills-Herrera, E., Orozco, L. E., Forero-Pineda, C., Pardo, O., & Andonova, V. (2011). The Relationship Between Perceptions of Insecurity, Social Capital, and Subjective Well-Being: Empirical Evidences from Areas of Rural Conflict in Colombia. *The Journal of Socio-Economics, 40*, 88–96.