CrossMark

A multilevel structural equation modelling approach to study segregation of deprivation: an application to Bolivia

Filippo Temporin[1]

Abstract

The study of segregation of deprivation can provide a tool to determine the economic, social and institutional factors associated with spatial unevenness in the distribution of wealth. Segregation is linked to social exclusion, diminished opportunities for human capital development and lower access to public services. In comparison to descriptive measures of poverty segregation, a multilevel structural equation modelling approach allows us to make statistical inferences about segregation, and to assess the extent to which segregation can be explained by contextual variables. Previous research using multilevel models to analyse segregation is extended to handle a continuous latent variable, measured by multiple binary indicators. The proposed approach is used to quantify the extent to which household deprivation is clustered within communities in Bolivia and to explore contextual factors associated with between-community differences in deprivation. Bolivia had one of the worst performances in poverty headcount ratio and chronic malnutrition in Latin America in the first decade of the twenty-first century, according to World Bank data. Bolivia is found to have a high level of segregation, since the main source of variation in deprivation arises from differences across communities, rather than within communities. Ethnicity, education, administrative region, distance to urban centres, and drought-induced migration significantly predict differences in the mean level of deprivation across Bolivian villages. This analysis helps to identify clusters of deprivation and highlights crucial sectors to be developed in order to reduce unevenness in the distribution of deprivation.

Keywords Segregation of deprivation · Bolivia · Multilevel model · Structural equation model

Abbreviations

DHS   Demographic and Health Survey
GPS   Global positioning system
PCA   Principal component analysis
SEM   Structural equation model

✉ Filippo Temporin
   f.temporin@lse.ac.uk

1   Department of Social Policy, London School of Economics and Political Science (LSE), Room 1.19, Houghton Street, London WC2A 2AE, UK

🖄 Springer

1 Introduction

Segregation can be defined as a form of physical separation where population groups are isolated into different neighbourhoods (in case of residential segregation) or schools (in case of educational segregation), "shaping the living environment at the neighbourhoods [or school] level" (Kawachi and Berkman 2003).

Geographical clustering of deprived people is commonly associated with economic, ethnic, or physical segregation, being the consequence of variation in characteristics under study across areas. Segregation of deprivation may be related to social exclusion,[1] with important consequences for social and health policies. Among the effects of social exclusion, we can highlight a diminished access to public services and decreased opportunities for human capital development. In Bolivia, for instance, social exclusion has been identified as a possible mechanism through which individuals belonging to certain ethnic groups reside in areas that tend also to have lower education and income (Gray-Molina et al. 2002). There is some evidence that the opportunities and even the conduct of people residing in certain neighbourhoods is shaped, among other factors, by the characteristics of their neighbourhood (Jencks and Mayer 1990). Geographic and social isolation could therefore be among the factors underlying certain social pathologies among the poor (Greene 1991). The analysis of deprivation and poverty segregation can help to identify the most deprived areas, which are economically and socially isolated from the more developed areas. It can provide a tool to determine the economic, social and institutional factors related to spatial unevenness in the distribution of wealth over the area under investigation. Deprivation and poverty segregation might be particularly suitable for policy interventions related to urban planning at a more local level than the national or regional level (Amarasinghe et al. 2005). Moreover, since a higher mortality rate and higher exposure to infectious diseases is likely to be found in contexts of concentrated deprivation (Fiscella and Franks 1997; Szwarcwald et al. 2002), reducing the differences in deprivation among communities might also be associated with the better health outcomes.

This study builds on the previous use of multilevel modelling to assess social segregation in schools and areas using a single binary or categorical socioeconomic indicators (Goldstein and Noden 2003; Leckie et al. 2012; Jones et al. 2018a). The main contribution of this paper is that the outcome of interest, household deprivation, is treated as a continuous latent variable, measured by a set of multiple correlated indicators. Multilevel structural equation modelling (SEM) allows the simultaneous creation of a latent variable for household deprivation, and its decomposition into between-community and between-household within-community components to measure segregation of deprivation. Moreover, multilevel modelling allows us not only to describe patterns of segregation, but to investigate the contextual factors associated with deprivation segregation, since it might be of interest to examine whether average levels of segregation vary across communities as a function of community characteristics (Bruch and Atwell 2015).

The proposed multilevel SEM is applied in a study of segregation of deprivation in Bolivia in 2008 using survey data linked to global positioning system (GPS) data. By the end of the first decade of the millennium, Bolivia was one of the poorest countries in South America (Population Reference Bureau 2013), and more than half of the population fell

---

[1] Social exclusion is the mechanism through which members of a certain group are denied full access to resources and opportunities that are available to others, associated for instance with housing, employment, or healthcare, and linked to social integration (Silver 1994).

below the poverty line, mostly in rural areas (World Bank 2014). Bolivian economic inequality is still great, with a Gini coefficient of 51.4 in 2008 (against an average of 49.9 of the other South American countries). The distribution of wealth within the country was not uniform, with considerable geographic and ethnic dissimilarities (Schroeder 2007). First, the extent of segregation of deprivation across Bolivian communities is quantified, and then area-level variables are used to explain the variation across communities, while allowing for segregation due to unmeasured area characteristics. The latent variable for household deprivation can be considered an alternative to previous indices, since it takes into account only items related to housing conditions with a sufficient degree of correlation among them, and which can therefore be considered manifestations of the underlying concept of household deprivation.

2 Approaches to the measurement of poverty and deprivation segregation

2.1 Descriptive segregation measures

The traditional approach in the study of segregation involves the use of descriptive indicators. The most widespread descriptive measure of segregation is the dissimilarity index (Duncan and Duncan 1955), which can be interpreted as the percentage of one of the population groups (for instance, the white population in the case of racial segregation) that would have to move to different areas in order to reproduce a distribution matching that of the larger areas. The dissimilarity index has been widely used in the deprivation and poverty segregation literature (Bibby 1975; Mershrod 1981; Napierala and Denton 2017), including the only study—to the best of our knowledge—on segregation in Bolivia, which investigated residential segregation in ten Bolivian cities (Gray-Molina et al. 2002). A drawback of the dissimilarity index is that it allows us to compute segregation only between two groups. Theil's (1972) information theory index, Bell's (1954) and Lieberson's (1981) isolation indices for multiple populations, and James' (1986) generalized exposure-based segregation index allow the calculation of segregation among multiple groups. Other measures of segregation that are based on the departure of each observation from measures of central tendency are the variance ratio index (Zoloth 1976), the Atkinson's family of segregation indices (Allison 1978), and the square root index (Hutchens 2001). Measures of variation based on the departure of each observation from all other observations, such as the Gini coefficient (Dorfman 1979), can also be interpreted as measures of segregation (Kim and Jargowsky 2009). As with the dissimilarity index, the Gini coefficient is related to the Lorenz—or segregation—curve (Gastwirth 1972). The standardized versions of these indices range from 0 (no segregation, i.e. all areas have the same proportion of population groups) to 1 (complete segregation, i.e. each area is composed of just one of the population groups) (Massey et al. 1996). The descriptive segregation measures described above are aspatial, meaning that they do not take into account the spatial proximity of the observations (Morrill 1991). A recent development in the measure of segregation involves the spatial dimension of segregation, for instance by including the length of shared boundaries (Wong 1993), or by using GPS data (Matthews and Parker 2013). The gradient of spatial segregation can be measured by spatial autocorrelation (Cliff et al. 1973), which has been widely used in the literature (Chakravorty 1996; Dawkins 2007; Amara and El Lahga 2016).

The above-mentioned indices are descriptive, meaning they are based on observed proportions of population groups that include the effect of random sampling variation (Allen et al. 2015). In other words, they fail to take into account the probabilistic component resulting from the sampling process; stochastic variation due to population sampling can bias segregation measurement, especially when small numbers are involved (Kish 1954; Leckie and Goldstein 2015). For instance, Leckie et al. (2012) pointed out that the dissimilarity index, which is based on observed rather than underlying proportions, has sources of bias depending on the size of the areas and on the underlying proportions; when analysing small areas, the dissimilarity index systematically overestimates segregation, suffering from the upward bias of the null (Allen et al. 2015). Brunch and Mare (2006) highlighted that indices of segregation based on the division of the population into categories based on some threshold, such as the dissimilarity index, are sensitive to changes in the choice of the thresholds. Finally, it is not possible to investigate the factors associated with deprivation segregation when descriptive measures are used (Owen 2015).

2.2 Multilevel modelling for studying segregation

A multilevel model approach overcomes the above-mentioned limitations, by separating the component of the observed proportion that is due to sampling variation. Segregation can be measured by estimating the higher-level variance parameter in the multilevel model (Goldstein and Noden 2003). This allows the assessment of the proportion of variation in the characteristic of interest that is due to the grouping of individuals within areas: the larger it is, the more segregated the neighbourhoods or schools are. By estimating standard errors, a statistical inference on segregation can be made (ibid). Moreover, multilevel models can be used to explore sources of segregation by including contextual covariates in the models (Leckie et al. 2012).

The first paper in this stream of literature is by Goldstein and Noden (2003), who measured the evenness of the distribution of disadvantaged students across English schools in the period 1994–1999, using a binary variable as the outcome, namely students' eligibility for free school meals. Since then, a growing number of studies using a multilevel approach have appeared in the literature. Three-level models were first used by Leckie et al. (2012) to study social segregation in schools, with students nested within schools nested within London local authorities. They were followed by other researchers, who applied the models to the study of the ethnic distribution within cities (Jones et al. 2015; Leckie and Goldstein 2015; Manley et al. 2015; Johnston et al. 2016; Jones et al. 2018a), allowing simultaneous estimation of the micro-, meso- and macro-effects of segregation. Leckie and Goldstein (2015) and Manley et al. (2015) extended the multilevel binomial logistic regression used in previous work to a multilevel multinomial logistic regression to model segregation by a categorical variable. A multilevel approach in the computation of the dissimilarity index has been developed by Harris (2017) and Harris and Owen (2017) when studying the residential segregation of students in England. Moreover, multilevel models can be extended to take into account the spatial proximity of areas, by including spatial weights (Jones and Subramanian 2014) and dependencies between areal units (Dong and Harris 2015; Jones et al. 2018b).

The present analysis involves a continuous latent dependent variable measured by multiple binary indicators as an outcome, and therefore requires an extension in a SEM framework of the multilevel models used in previous work. An application to Bolivia is proposed in the last sections of the paper, in order to quantify the extent of segregation in the country

and to explore contextual factors associated with differences in the mean deprivation across communities.

3 Statistical methods

3.1 Latent variable model for household deprivation

An index measuring deprivation (or wealth) is an alternative to monetary measures such as income or expenditure, which are often unavailable or unreliable in low- or middle-income countries (Filmer and Kinnon 2012). Deprivation can be considered as a concept underlying certain characteristics of living standards and can therefore be derived from a set of observable items.

A key point in the creation of a composite index of deprivation is the choice of weights to be assigned to the observed items. Many approaches exist in the literature, ranging from the simple sum of the owned items to more sophisticated data-driven techniques that take into account the extent to which each item discriminates between households' deprivation (Vandemorteele 2014). Among these composite indicators, the DHS wealth index, built from principal component analysis (PCA), is probably the most widespread (Rutstein and Johnson 2004). In the following sections, a critique of the construction of the DHS wealth index is presented, and a latent variable approach is proposed.

3.2 Critique of the DHS wealth index

The DHS wealth index is constructed by means of PCA, a technique that transforms a set of observed correlated items into a set of linearly uncorrelated principal components by means of an orthogonal transformation (Jolliffe 1986). PCA's major limitation is that it does not take into account the categorical nature of the observed indicators, treating them as continuous, which is analogous to using an OLS regression for the analysis of a categorical outcome (Howe et al. 2008). The wealth index scores are built from the first principal component, which often explains only a low proportion of the total variation in the observed items (Kolenikov and Angeles 2004). Moreover, since the correlation between the observed indicators has not been investigated before the analysis, the linear dependence between the items could lead to incorrect estimates of the wealth index (ibid.). Finally, using the DHS wealth index as a measure of deprivation in further analyses ignores the measurement error that arises from constructing an index from a set of items.

3.3 Rationale for the construction of a latent variable for household deprivation

SEM is a latent variable approach that incorporates a model for the relationship between a continuous latent variable and a set of observed items, considered as the manifestation of the latent variable (Bartholomew and Knott 2011). In this case, for instance, a set of observed items relating to housing conditions and living standards are combined into a latent variable for household deprivation.

A SEM is composed of a measurement model and a structural model, estimated simultaneously. The measurement model describes the relationship between the observed items and the latent variable. The structural model is a regression of the latent variable on a set of covariates (Bartholomew and Knott 2011). In contrast to PCA, the items included in the measurement

model of SEM can be binary or polytomous (ibid.). Weights are assigned to the items depending on their ability to discriminate between households' scores on the latent variable. By estimating standard errors, SEM also allows testing hypotheses involving parameters of both the measurement and structural models. An important feature of SEM is that it takes into account the measurement error which may bias the estimates of the level of segregation within communities. Latent variables do not have measurement error associated with them, since they are not directly measured, therefore the association between them and other covariates can be estimated without any bias (Muthén and Muthén 2010).

In comparison to the DHS wealth index, a further development of the proposed approach is the selection of the observed items, which is based on the correlation matrix of all items. Only items relating to the latent concept of deprivation are included in the measurement model, as explained later.

3.4 Measurement model

The measurement model specifies the relationship between the latent variable and the observed items. Denote by y_{rjk} the rth item ($r = 1, \ldots, p$) of household $j (j = 1, \ldots, n_k)$, nested within community $k (k = 1, \ldots, K)$. Then the logit of the probability that household j in community k owns item r is:

$$\text{logit}\big(P\big(y_{rjk} = 1|\eta\big)\big) = \text{logit}\big(\pi_{rjk}(\eta)\big) = \alpha_{r1}\eta_{jk} - \alpha_{r0}, \tag{1}$$

where $\eta_{jk} \sim N\big(0, \sigma_\eta^2\big)$ is the latent variable for household deprivation and α_{r0} and α_{r1} are, respectively, the difficulty and the discrimination parameters. The difficulty parameter α_{r0} indicates how "difficult" an item is to be owned, while the discrimination parameter α_{r1} indicates how well the rth item discriminates between households with different scores for deprivation. In order to identify the model, some constraint must be imposed on the item parameters. It is common to constrain one of the α_{r1}s to 1, which sets the scale of the latent variable to be equal to the scale of the chosen item.

3.5 Multilevel structural model

In this paper, the multilevel structural models specify the partitioning of the variance into a between-community component and a within-community between-household component. Of particular interest is the extent to which community variation can be explained by the community-level covariates described earlier. An important characteristic of multilevel SEM is that the creation of the latent outcome variable and the analysis of its between- and within-community components is done simultaneously, while accounting for measurement error (Muthén and Muthén 2010).

The structural model specifying the decomposition of the latent variable η_{jk} into its within- and between-community components is:

$$\eta_{jk} = \beta_k + u_{jk}^{(hh)} \tag{2}$$

$$\beta_k = \gamma_{00} + u_k^{(PSU)}, \tag{3}$$

where $u_{jk}^{(hh)} \sim N\big(0, \sigma_u^{2(hh)}\big)$ is the household residual and $u_k^{(PSU)} \sim N\big(0, \sigma_u^{2(PSU)}\big)$ is the community-level random effect. They represent, respectively, the within-community and the

between-community components of household deprivation, and their variances $\sigma_u^{2(hh)}$ and $\sigma_u^{2(PSU)}$ are the within-community and the between-community variances.

Segregation of deprivation is strictly related to variation across communities. In fact, the higher the between-community variation of the level of deprivation in a country, the higher the level of grouping of deprived people within geographical areas. On the other hand, no between-community variation indicates that no segregation is present in a country (Bulle 2016).

The models are fitted by maximum likelihood, and likelihood ratio tests can be used to compare the fit of nested models. The analyses have been carried out using the *gsem* function in the Stata software (StataCorp 2013).

4 An application to Bolivia

4.1 Potential explanations for geographical segregation of deprivation in Bolivia

An application of the SEM models explained earlier is here proposed to explain the segregation of deprivation in Bolivia, by looking at the potential factors associated with the between-community variation in deprivation. Among these, ethnic composition, education, distance to urban centres and drought-induced rural–urban migration can have a central role.

The first factor that may affect the segregation of deprivation is ethnicity. The Bolivian population is mainly indigenous, and the ethnic distribution is not uniform, with indigenous populations more concentrated in certain areas—mainly the *Altiplano* (high plateau) and *Valle* (valley) regions. Almost the whole indigenous population (97.5%) of rural areas is found to be chronically poor (Castellanos 2007), since the lack of social welfare programmes leads to a high vulnerability to shocks such as droughts, floods and hailstorms (Buzaglo and Calzadilla 2009).

Education can play a role in explaining between-community variation in the level of deprivation in the country. The link between parental education and the socioeconomic status of a household is well established (Cornia 2014; King and Hill 1993). Education can also be a contextual factor in determining the unevenness of the distribution of deprivation across Bolivian communities. The average degree of education in the community can set the context for a wide set of socioeconomic factors, including economic disadvantage (Wight et al. 2006) which lead to the geographical segregation of deprivation.

Distance to urban centres might also explain deprivation segregation. Social segregation studied by Gray-Molina et al. (2002) in Bolivian urban environments, can be extended to rural areas. The main activity in rural areas is farming: peasants are vulnerable to shock linked to climate change such as drought (Castellanos 2007), and lack of roads might affect peasants' access to the market (Buzaglo and Calzadilla 2009). Rural areas are also associated with a lack of infrastructure (Andersen 2002) and basic services like sanitation and availability of clean water (Coa and Ochoa 2009), creating a setting of a higher mean level of deprivation.

Finally, Bolivia has been subject to natural disasters over the last decades. In particular, prolonged droughts have affected the South-West part of the country (Kessler and Stroosnijder 2006). Agriculture and livestock rely strongly on vegetation resources, the availability of which can be jeopardized by these events: it has been calculated that, in the period 1953–1993, Bolivia lost 30% of its agricultural productivity, and one of the main reasons

is related to soil erosion (Benton 1993). Droughts have fostered migration towards the cities. Bolivia faced a rapid process of urbanization, either temporary or permanent, between the 1980s and the 2000s (World Bank 2015). Drought-driven rural–urban migration can lead to the uneven residential sorting of rural migrants within cities, which leads to a rise in the level of urban residential segregation. Moreover, there is some evidence of a recent trend towards migration differentiated by age-group. The main mechanism is related to the fact that young men are gradually excluded from access to agricultural soil, due to the increased unavailability of land (Balderrama 2011). Lands are usually distributed among the children, but there is evidence of the tendency of migrant young men to refuse their share of the inheritance (Michels 2011). This selective migration (Borjas and Tienda 1987) can therefore be another explanation for the segregation of deprivation in Bolivia.

4.2 Data and measures

The Demographic and Health Surveys (DHS) collect data on a broad range of aspects related to health and living conditions. In the sampling process, clusters of a standard size of 100 households are identified and mapped in the territory of the country under investigation, and a further selection within each of these selected clusters is made: each of these areal units serves as a primary sample unit (US Aid 2012). In this paper, primary sample units are considered to be proxies for the respondents' communities, as in previous studies (Uthman et al. 2011; Robson et al. 2012).

The 2008 Bolivian DHS dataset contains 19,564 households from 999 communities. Among them, 11,361 household have complete records on the ownership of the items related to housing conditions and on the variable included as predictors in the structural model. The full set of items related to housing conditions, living standards and owned assets available in the DHS dataset includes: availability of electricity, availability of clean water, type of sanitation, material of the floor, type of cooking fuels, and ownership of refrigerator, radio, television, motorbike, car, telephone and bicycle. These are the items used in the construction of the DHS wealth index, a composite measure of a household's cumulative living standard (Rutstein and Johnson 2004). All the observed variables have been dichotomized, in order to simplify the interpretation of the parameters of the models. As noted earlier, there are four main factors that can be linked to the between-community variation in deprivation: ethnicity, education, distance to urban centres and drought-induced migration. These are represented by six explanatory variables listed in Table 1. All of these (except group mean centred years of male education) have been measured at the community level.

The contextual binary variable *Indigenous*, provided by DHS, indicates whether a household lives in a community which has a majority of indigenous or non-indigenous villages. The mean level of male education within each community has been chosen as a contextual variable. When including a contextual variable calculated as the mean of a household-level variable, it is common to include the group mean centred household-level variable, in order to separate the between- and within-community effects (Snijders and Bosker 2012). For households with more than one adult male (5.97% of the total), the mean value of years of schooling of the males registered at that household has been calculated. In general, individual-level male education can better explain the level of deprivation than female education: paternal rather than maternal income is a strong determinant of the wealth status of the household (Cornia 2014; Thomas 1990), and in Bolivian indigenous groups, men are more likely to assume the position of breadwinners (Paulson et al. 1990).

Table 1 List of covariates

Variable	Source	Values
Indigenous village	DHS	Indigenous, non-indigenous
Community-level mean years of male education	DHS	[0.7; 17]
Group mean centred years of male education	DHS	[−13.2; 13.2]
Administrative region	DHS	Beni, Chuquisaca, Cochabamba, La Paz, Oruro, Pando, Potosí, Santa Cruz, Tarija
Distance to the closest municipal capital (km)	GeoBolivia	[0.06; 96.51]
Risk of drought	SINSAAT	Very low, low, medium, high

The distance from the centroid of each DHS cluster to the closest municipal capital has been obtained by linking the DHS GPS dataset and the GeoBolivia dataset (GeoBolivia 2017a), which provides the location of the 339 Bolivian municipal capitals. The distance has been calculated using the Haverisine formula[2] (Robusto 1957). The distance to the closest municipal capital can provide a better measure of the variation between urban and rural environments, approaching the concept of Woods' (2003) "urban–rural continuum". The mean distance of the communities labelled as urban in the DHS variable is 3.88 km, while it is 16.84 km for the rural communities. The variable related to risk of drought has been created by linking the DHS GPS dataset with the 2002 National System for Early Alert of Food Security (Sistema Nacional de Seguridad Alimentaria Alerta Temprana, SINSAAT) (GeoBolivia 2017b). This dataset classifies areas into four levels of drought risk, depending on the frequency of drought over the period 1972–2002. Very low risk is defined as one or no drought every fifth year over the 30-year period, low risk as a drought every fourth year, medium risk as a drought every second year and high risk as four or more droughts every 5 year.

In the most recent DHS surveys, each community is georeferenced during the sample listing process. The GPS readers are in general accurate to less than 15 metres, but the GPS coordinates of each community are randomly displaced due to issues of confidentiality: the error ranges from 0 to 2 km for urban communities and from 0 to 5 km for rural communities (Perez-Heydrich et al. 2013). While cluster displacement might induce large misclassification errors when calculating the distance between clusters' centroids and health facilities or other specific locations (Skiles et al. 2013), the random displacement of the centroid of the communities is unlikely to affect the results of this study. First, the region of each community is directly calculated from DHS, so no issue of displacement arises even when the random error is introduced. Second, the distance to the closest municipal capital is the variable that mostly could be affected by the random error, but it is still considered a better approximation of the rural–urban continuum (Woods 2003) than the binary variable provided by DHS, which has

---

[2] The Haversine distance does not reflect real distance, especially in a territory like Bolivia, which is highly mountainous in the South-West areas. It is reasonable to think that Bolivians willing to reach the closest municipal capital might have to cover longer distances than the great-circle line connecting their village to the target. A better estimate of such distance would be the walking (or driving) path from each community to the municipal capital. However, no reliable GPS dataset on minor streets and trails has been found. The only available dataset is related to main roads (GeoBolivia 2013), but this is not specific enough to include all the walking trails that Bolivians might take. Therefore, the Haversine formula has been considered the best available approximation of the real distance to the closest municipal capital.

only the two categories "urban" and "rural". Third, the areas for risk of drought are very large and the risk of displacement of a community is very low.

4.3 Selection of deprivation indicators

The full set of 12 items available in the DHS dataset included *Electricity*, *Water*, *Sanitation*, *Floor*, *Cooking fuels*, *Radio*, *Television*, *Refrigerator*, *Motorbike*, *Bicycle*, *Car* and *Telephone*. These are the same items used for the construction of the DHS wealth index. These items were divided into two sets: the first five items were related to the living environment, while the last seven were assets or possessions.

The aim of the investigation of the correlation matrix was to select the observed items used to construct the latent variable, in order to avoid multicollinearity and to have a coherent set of indicators measuring household deprivation. Tetrachoric correlations estimated the correlation between two theorized normally distributed latent variables from two observed binary variables (Divgi 1979). With the aim of analysing a unique latent variable for household deprivation, the aforementioned observed variables were selected according to their tetrachoric correlations. The items *Bicycle*, *Motorbike*, *Car*, and *Radio* showed a weak tetrachoric correlation with the rest of the items, and were therefore excluded from the measurement model. Although the correlations between *Television*, *Telephone* and the retained items were sufficiently strong, they were excluded from the measurement model on a theoretical basis. These items cannot be considered as basic needs in the context of a low-income country such as Bolivia. On the other hand, the asset *Refrigerator* is the only one that was retained in the measurement model, due to its strong association with health outcomes. By allowing us to keep food fresh, a refrigerator can indeed be related to hygiene and diseases (Lagendijk et al. 2008). Therefore, the six selected items for the measurement model of household deprivation were *Electricity*, *Water*, *Sanitation*, *Floor*, *Cooking fuel* and *Refrigerator*. These items had a tetrachoric correlation higher than 0.5 (Table 2), suggesting that they were manifestations of the same underlying concept.

4.4 Measurement model for household deprivation

The measurement model of Eq. (1) can be interpreted as a single-level model. The total variance of the latent variable σ_η^2 was estimated as 19.15. The Spearman rank correlation with the DHS wealth index was high in the single-level latent variable, with a value of 0.92. This result is consistent with previous attempts to construct a latent variable for wealth (Vandemoortele 2014).

Note that the discrimination parameter related to the item *Electricity* was constrained to 1 for identification. As can be seen in Table 3, *Cooking fuel* and *Electricity* were the items that best discriminated between households with different deprivation scores, while *Water* and *Sanitation* had the least discriminatory power. Therefore, having electricity discerned household deprivation better than, for instance, having clean water. Moreover, *Water* and *Sanitation* were the more likely items to be owned (those with lower values in the difficulty parameters), while *Cooking fuel* was the least likely.

Table 2 Tetrachoric correlation matrix, retained items only

	Electricity	Water	Sanitation	Floor	Cooking fuel	Refrigerator
Electricity	1					
Water	0.66	1				
Sanitation	0.66	0.51	1			
Floor	0.78	0.64	0.59	1		
Cooking fuel	0.86	0.64	0.67	0.84	1	
Refrigerator	0.81	0.52	0.68	0.71	0.72	1

Table 3 Discrimination and difficulty parameters from the measurement model for deprivation

Item	Discr. (α_{r1})	SE (α_{r1})	Diff. (α_{r0})	SE (α_{r0})
Electricity	1.00	(Constrained)	-4.05	0.12
Water	0.41	0.02	-6.39	0.06
Sanitation	0.43	0.02	-3.91	0.04
Floor	0.72	0.04	-3.00	0.07
Cooking fuel	1.02	0.06	-2.32	0.36
Refrigerator	0.57	0.03	1.86	0.04

4.5 Results from the empty multilevel model

The aim of the multilevel structural models of Eqs. (2) and (3) was to analyse the distribution of the latent variable for household deprivation between and within Bolivian communities. In the multilevel model, the between- and within-community variance components were, respectively, 19.51 and 1.77. The intra-community correlation, that is the proportion of variation in the latent variable explained by the grouping of households within communities, allowed an assessment of the level of segregation: a high level of community-level variance reflects substantial differences in household deprivation across communities (Leckie et al. 2012). For this model, a high proportion of variation in the latent variable (around 92%) was due to the grouping of households within communities. Thus, households within the same community had very similar scores on the latent variable of deprivation. This finding is consistent with previous studies: Castellanos (2007) points out the relatively low level of inequality among indigenous households in rural Bolivian communities.

4.6 Results from the models including contextual factors of deprivation segregation

SEM allowed investigation of the factors associated with deprivation segregation, by including community-level covariates in the model. Table 4 shows the results of the univariate and multivariate models of Eqs. (2) and (3). First, the coefficient of *Indigenous* was significantly negative: communities with a majority indigenous population were found to be more likely to have higher mean deprivation. Indigenous origins are associated with poverty in rural Bolivian communities (Albo 1994; Grootaert and Narayan 2004); the Bolivian indigenous population is mainly clustered in the *Altiplano* and *Valle* regions in isolated rural communities, with high vulnerability to natural hazards and a subsequent lack of roads, access to markets, and social infrastructure (Buzaglo and Calzadilla 2009). Therefore, due to their disadvantaged position, the concentration of indigenous households in certain areas leads to the segregation of deprivation.

Second, both coefficients related to male education were significant and positive. The between effect indicated that the higher the mean level of male education within a community, the lower the mean level of deprivation of that community. Education underlies a broad range of socioeconomic factors, including lower economic conditions (Wight et al. 2006), leading to deprivation segregation. While indigenous origins are associated with lower formal education in the literature (Castellanos 2007), the multivariate model in this paper indicated that education is associated with segregation of deprivation while also taking into account ethnicity.

Third, two regions, Potosí and Beni, had a significantly higher level of deprivation than La Paz. The territory of Potosí, located in the South-West of the country, is mainly mountainous, posing issues of accessibility, as well as difficulties in promoting extensive agricultural exploitation. This region presents the highest presence of indigenous population (Castellanos 2007), and has been affected several times by severe drought (Gray-Molina et al. 2002). Beni's case is different: this region is rich in raw materials and represents one of the biggest agricultural centres in Bolivia (Vadez et al. 2004). Despite its richness in natural resources, the level of poverty is still high, being a mainly rural territory, lacking big urban centres and being in a logistically marginal area when compared to the leading Bolivian economic poles (Weisbrot and Sandoval 2008).

As a fourth result, the coefficient of *Distance to municipal capital* was significantly positive: every additional kilometre of distance from the closest municipal capital was associated with an average decrease of 0.18 in the community-level score of the latent variable for household deprivation. Rural populations are strongly dependent on farming productivity, which leads to a high vulnerability to shocks such as drought or flooding (Castellanos 2007). Rural populations are also exposed to endemic diseases that can affect labour productivity and consequently levels of deprivation (Buzaglo and Calzadilla 2009), since 26.7% of rural households retrieve water from a source considered unsafe, and 56.7% lack basic sanitation services (against, respectively, 5.4 and 9.3% in urban areas) (Coa and Ochoa 2009).

Moreover, the coefficients indicated that the communities located in the medium- and low-risk areas of drought had a lower mean level of deprivation than the communities in areas of high risk. Climate change has triggered rural–urban migrations; a rapid process of urbanization has been observed in Bolivia between the 1980s and the 2000s (World Bank 2015). Punch (2004) observes that in a rural Bolivian village in Tarija (located in the area at medium risk of drought) migration rather than education is considered the best way to

Table 4 Results for the structural models, all models

Model	Variable	Univariate models		Multivariate model	
		Coeff.	95% CI	Coeff.	95% CI
Indigenous village [Ref. non indigenous]	Indigenous	−2.67	[−3.25; −2.09]	−1.47	[−1.84; −1.10]
	(LR test—versus empty model)	$X^2=86.23$	$d.f.=1$		
Male education	Community-level mean years of male education	1.08	[0.99; 1.17]	0.92	[0.84; 1.01]
	Group mean centred years of male education	0.19	[0.17; 0.20]	0.19	[0.17; 0.20]
	(LR test—versus empty model)	$X^2=1720.81$	$d.f.=2$		
Administrative region [Ref. La Paz]	Chuquisaca	−0.66	[−1.76; −0.43]		
	Cochabamba	0.48	[−0.49; 1.45]		
	Oruro	−0.70	[−1.77; 0.37]		
	Potosí	−1.44	[−2.49; −0.40]		
	Tarija	0.53	[−0.56; 1.62]		
	Santa Cruz	0.76	[−0.15; 1.68]		
	Beni	−2.40	[−3.77; −1.13]		
	Pando	0.94	[−0.72; 2.60]		
	(LR test—versus empty model)	$X^2=40.31$	$d.f.=8$		
Distance to the closest municipal capital	Distance	−0.19	[−0.21; −0.16]	−0.07	[−0.08; −0.05]
	(LR test—versus empty model)	$X^2=269.15$	$d.f.=1$		
Risk of drought [Ref. High]	Very low	2.14	[−0.46; 4.73]	0.30	[−1.16; 1.70]
	Low	4.92	[2.33; 7.52]	2.03	[0.60; 3.47]
	Medium	3.89	[1.36; 6.43]	1.53	[0.14; 2.92]
	(LR test—versus empty model)	$X^2=43.50$	$d.f.=3$		

improve living standards, since migrant work offers more security and immediate benefits. Rural–urban migration was associated with the uneven residential sorting of the migrants within the urban environment, increasing the level of urban residential segregation.

Little difference was found in the multivariate model simultaneously including these variables: rural, indigenous communities with a lower mean level of male education and at higher risk of drought were significantly more likely to have higher mean deprivation. *Region* was not included in the model, since it was highly correlated with *Risk of drought*: the areas of risk overlapped with many of the Bolivian regions. *Risk of drought* was preferred because of its higher theoretical value as a potential explanation for segregation of deprivation within communities, being a cause of selective rural–urban migration (Balderrama 2011).

5 Discussion

This paper proposes a general SEM approach to the study of geographical segregation, by extending the multilevel modelling approach proposed by Goldstein and Noden (2003) to handle constructs measured by multiple indicators. This approach enables us to not just quantify the extent of segregation but to model patterns of segregation as functions of contextual factors.

The proposed multilevel SEM approach is applied in a study of deprivation segregation in Bolivia, a country that presented among the highest indicators of poverty and deprivation in Latin America (Coa and Ochoa 2009). By analysing 2008 DHS data, a latent variable for household deprivation was created from a set of six observed items, and simultaneously included in the SEM models, overcoming issues related to measurement error (Muthén and Muthén 2010). Bolivia was found to have a high level of segregation of deprivation, since a high proportion of variation in the latent variable was due to the grouping of households within communities. Ethnicity, education, administrative region, distance to urban centres and drought-induced migration significantly explained differences in the mean level of deprivation across Bolivian villages. This analysis highlighted the differences in the use of the latent variable in comparison to the DHS wealth index; the inclusion of this latter measure leaded to an underestimation of the magnitude of the segregation of deprivation in Bolivia, since the DHS wealth index did not take into account measurement error and the items used in the construction of the two indices were slightly different.

The results of the analysis have implications for social and health policies. By identifying the contextual factors associated with the segregation of deprivation, this paper provides evidence on the mechanisms leading to economic and social segregation. This analysis helps in identifying segregation of deprivation within Bolivia, and highlights crucial sectors to be developed in order to fight spatial unevenness in the distribution of wealth, linked to social exclusion, diminished opportunities for human capital development and lower access to public services. Finally, reducing inequality across Bolivian communities could also positively affect health indicators, since contexts of concentrated deprivation are associated with higher mortality and higher exposure to infectious diseases (Fiscella and Franks 1997; Szwarcwald et al. 2002).

References

Albo, X.: Ethnic Violence: the Case Of Bolivia; in the Culture of Violence, pp. 119–143. The United Nations University, Shibuya-ku, Tokio (1994)

Allen, R., Burgess, S., Davidson, R., Windmeijer, F.: More reliable inference for the dissimilarity index of segregation. Econom. J. **18**(1), 40–66 (2015)

Allison, P.D.: Measures of inequality. Am. Sociol. Rev. **43**, 865–880 (1978)

Amara, M., El Lahga, A.: Tunisian constituent assembly elections: how does spatial proximity matter? Qual. Quant. **50**(1), 65–88 (2016)

Amarasinghe, U., Samad, M., Anputhas, M.: Spatial clustering of rural poverty and food insecurity in Sri Lanka. Food Policy **30**(5), 493–509 (2005)

Andersen, L.E.: Rural-urban migration in Bolivia: advantages and disadvantages. Documento de Trabajo, Instituto de Investigaciones Socio-Económicas, Universidad Católica Boliviana (No. 05/02) (2002)

Balderrama, C.: Rural migration in Bolivia: the impact of climate change, economic crisis and state policy. IIED; No. 31 (2011)

Bartholomew, D.J., Knott, M., Moustaki, I.: Latent variable models and factor analysis: a unified approach, vol. 904. Wiley, New York (2011)

Bell, W.: A probability model for the measurement of ecological segregation. Soc. Forces **32**, 357–364 (1954)

Benton, J.: The role of womens organisations and groups in community development. A case study of Bolivia. USAid working papers; Document No. 127039; Retrieved April 17, 2016, from https://www.popline.org/node/339234 (1993)

Bibby, J.: Methods of measuring mobility. Qual. Quant. **9**(2), 107–136 (1975)

Borjas, G.J., Tienda, M.: The economic consequences of immigration. Science **235**, 645–652 (1987)

Bruch, E., Atwell, J.: Agent-based models in empirical social research. Soc. Methods Res. **44**(2), 186–221 (2015)

Bruch, E., Mare, R.D.: Neighborhood choice and neighborhood change. Am. J. Sociol. **112**(3), 667–709 (2006)

Bulle, N.: A Method of measuring inequality within a selection process. Sociolog. Methods Res. **45**(1), 69–108 (2016)

Buzaglo, J., Calzadilla, A.: Towards a new consensus: poverty reduction strategies for Bolivia. Int. J. Dev. Issues **8**(1), 18–39 (2009)

Castellanos, I.O.V.: Extreme Poverty: Vulnerability and Coping Strategies Among Indigenous People in Rural Areas of Bolivia. Cuvillier Verlag, Gottingen (2007)

Chakravorty, S.: A measurement of spatial disparity: the case of income inequality. Urban Stud. **33**(9), 1671–1686 (1996)

Cliff, A., Ord, J.K.: Spatial Autocorrelation. Taylor & Francis; No. 04; QA278. 2, C5 (1973)

Coa, R., Ochoa, L.H.: Bolivia Encuesta Nacional de Demografía y Salud 2008. Ministerio de Salud y Deportes and Macro International, Calverton, Maryland. Retrieved October 15, 2015, from http://dhsprogram.com/pubs/pdf/FR228/FR228.pdf (2009)

Cornia, G.A.: Falling Inequality in Latin America: Policy Changes and Lessons. OUP Oxford, Oxford (2014)

Dawkins, C.J.: Space and the measurement of income segregation. J. Reg. Sci. **47**(2), 255–272 (2007)

Divgi, D.: Calculation of the tetrachoric correlation coefficient. Psychometrika **44**(2), 169–172 (1979)

Dong, G., Harris, R.: Spatial autoregressive models for geographically hierarchical data structures. Geogr. Anal. **47**(2), 173–191 (2015)

Dorfman, R.: A formula for the Gini coefficient. Rev. Econ. Stat. **61**, 146–149 (1979)

Duncan, O.D., Duncan, B.: A methodological analysis of segregation indexes. Am. Sociol. Rev. **20**(2), 210–217 (1955)

Filmer, D., Kinnon, S.: Assessing asset indices. Demography **49**(1), 359–392 (2012)

Fiscella, K., Franks, P.: Poverty or income inequality as predictor of mortality: longitudinal cohort study. Bmj **314**(7096), 1724 (1997)

Gastwirth, J.L.: The estimation of the Lorenz curve and Gini index. Rev. Econ. Stat. **54**, 306–316 (1972)

GeoBolivia: Red vial fundamental de Bolivia 2012. Retrieved May 2, 2017, from https://geo.gob.bo/geone twork/srv/esp/catalog.search#/metadata/9c141353-e99b-44c9-b121-6f8fe75ac240 (2013)

GeoBolivia: Capitales municipales de Bolivia (339 municipios), 2013. Retrieved May 2, 2017, from https://geo.gob.bo/download/?w=fondos&l=CapitalesMunicipales (2017a)

GeoBolivia: Mapa de amenaza de sequía meteorológica. Retrieved May 2, 2017, from https://geo.gob.bo/geonetwork/srv/esp/catalog.search#/metadata/e304b259-8dcf-4886-b0f8-d0bcf99c02d0 (2017b)

Goldstein, H., Noden, P.: Modelling social segregation. Oxf. Rev. Educ. **29**(2), 225–237 (2003)

Gray-Molina, G., Jimenez, W.L., de Rada, E.P.: Social exclusion: residential segregation in Bolivian cities. Inter-American Development Bank; Research Network working papers; R-440 (2002)

Greene, R.: Poverty concentration measures and the urban underclass. Econ. Geogr. **67**(3), 240–252 (1991)

Grootaert, C., Narayan, D.: Local institutions, poverty and household welfare in Bolivia. World Dev. **32**(7), 1179–1198 (2004)

Harris, R.: Measuring the scales of segregation: looking at the residential separation of White British and other schoolchildren in England using a multilevel index of dissimilarity. Trans. Inst. Br. Geogr. **42**(3), 432–444 (2017)

Harris, R., Owen, D.: Implementing a multilevel index of dissimilarity in R with a case study of the changing scales of residential ethnic segregation in England and Wales. Environ. Plan. B Urban Anal. City Sci. **45**(6), 2399808317748328 (2017)

Howe, L.D., Hargreaves, J.R., Huttly, S.R.: Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. Emerg. Themes Epidemiol. **5**(1), 1 (2008)

Hutchens, R.: Numerical measures of segregation: desirable properties and their implications. Math. Soc. Sci. **42**(1), 13–29 (2001)

James, F.J.: A new generalized "exposure-based" segregation index: demonstration in Denver and Houston. Sociol. Methods Res. **14**(3), 301–316 (1986)

Jencks, C., Mayer, S.E.: The social consequences of growing up in a poor neighborhood. Inner City Poverty U. S. **111**, 186 (1990)

Johnston, R., Forrest, J., Jones, K., Manley, D.: The scale of segregation: ancestral groups in Sydney, 2011. Urban Geogr. **37**(7), 985–1008 (2016)

Jolliffe, I.T.: Principal component analysis and factor analysis. In: Jolliffe, I.T. (ed.) Principal Component Analysis, pp. 115–128. Springer, Berlin (1986)

Jones, K., Subramanian, S.V.: Developing multilevel models for analysing contextuality, heterogeneity and change using MLwiN Volume 2. Bristol: Centre for Multilevel Modelling, University of Bristol. Retrieved October 20, 2018, from https://www.researchgate.net/publication/260772180_Developing _multilevel_models_for_analysing_contextuality_heterogeneity_and_change_using_MLwiN_Volum e_2 (2013)

Jones, K., Johnston, R., Manley, D., Owen, D., Charlton, C.: Ethnic residential segregation: a multilevel, multigroup, multiscale approach exemplified by London in 2011. Demography **52**(6), 1995–2019 (2015)

Jones, K., Johnston, R., Forrest, J., Charlton, C., Manley, D.: Ethnic and class residential segregation: exploring their intersection: a multilevel analysis of ancestry and occupational class in Sydney. Urban Stud. **55**(6), 1163–1184 (2018a)

Jones, K., Manley, D., Johnston, R., Owen, D.: Modelling residential segregation as unevenness and clustering: a multilevel modelling approach incorporating spatial dependence and tackling the MAUP. Environ. Plan. B Urban Anal. City Sci. **45**(6), 1122–1141 (2018b)

Kawachi, I., Berkman, L.F.: Neighborhoods and Health. Oxford University Press, Oxford (2003)

Kessler, C., Stroosnijder, L.: Land degradation assessment by farmers in Bolivian mountain valleys. Land Degrad. Dev. **17**(3), 235–248 (2006)

Kim, J., Jargowsky, P.A.: The Gini coefficient and segregation on a continuous variable. In: Flückiger, Y., Reardon, S.F., Silber, J. (eds.) Occupational and Residential Segregation, vol. 17, pp. 57–70. Emerald Group Publishing Limited, Bingley (2009)

King, E.M., Hill, M.A.: Women's Education in Developing Countries: Barriers, Benefits, and Policies. Published for the World Bank by the John Hopkins University Press, Baltimore (1993)

Kish, L.: Differentiation in metropolitan areas. Am. Sociol. Rev. **19**(4), 388–398 (1954)

Kolenikov, S., Angeles, G.: The use of discrete data in PCA: theory, simulations, and applications to socio-economic indices, pp. 1–59. Carolina Population Center, University of North Carolina, Chapel Hill (2004)

Lagendijk, E., Assere, A., Derens, E., Carpentier, B.: Domestic refrigeration practices with emphasis on hygiene: analysis of a survey and consumer recommendations. J. Food Prot. **71**(9), 1898–1904 (2008)

Leckie, G., Goldstein, H.: A multilevel modelling approach to measuring changing patterns of ethnic composition and segregation among London secondary schools, 2001–2010. J. R. Stat. Soc. Ser. A (Stat. Soc.) **178**(2), 405–424 (2015)

Leckie, G., Pillinger, R., Jones, K., Goldstein, H.: Multilevel modeling of social segregation. J. Educ. Behav. Stat. **37**(1), 3–30 (2012)

Lieberson, S., Peach, C., Robinson, V., Smith, S.: An asymmetrical approach to segregation. In: Peach, C., Robinson, V., Smith, S. (eds.) Ethnic Segregation in Cities, pp. 61–82. Taylor & Francis (1981)

Manley, D., Johnston, R., Jones, K., Owen, D.: Macro-, meso-and microscale segregation: modeling changing ethnic residential patterns in Auckland, New Zealand, 2001–2013. Ann. Assoc. Am. Geogr. **105**(5), 951–967 (2015)

Massey, D.S., White, M.J., Phua, V.C.: The dimensions of segregation revisited. Sociol. Methods Res. **25**(2), 172–206 (1996)

Matthews, S.A., Parker, D.M.: Progress in spatial demography. Demogr. Res. **28**, 271–312 (2013)

Merschrod, K.: The index of dissimilarity as a measure of inequality. Qual. Quant. **15**(4), 403–411 (1981)

Michels, A.: Migration and inheritance practices in the Bolivian Altiplano. Working paper/World Institute for Development Economics Research (2011)

Morrill, R.L.: On the measure of spatial segregation. Geogr. Res. Forum **11**, 25–36 (1991)

Muthén, L., Muthén, B.: Mplus Software (Version 6). Muthén & Muthén, Los Angeles, CA (2010)

Napierala, J., Denton, N.: Measuring residential segregation with the ACS: how the margin of error affects the dissimilarity index. Demography **54**(1), 285–309 (2017)

Owen, D.: Measuring residential segregation in England and Wales: a model-based approach. Doctoral dissertation, University of Bristol (2015)

Paulson, S., Gisbert, M.E., Quitón, M.: Case studies of two womens health projects in Bolivia. USAid working papers; Document No. 127864; Retrieved April 25, 2016, from https://www.popline.org/node/306664 (1996)

Perez-Heydrich, C., Warren, J.L., Burgert, C.R., Emch, M.: Guidelines on the use of DHS GPS data. In: ICF International (2013)

Population Reference Bureau: 2013 World Population Data Sheet. Retrieved March 30, 2016, from www.prb.org/pdf13/2013-population-data-sheet_eng.pdf (2013)

Punch, S.: The impact of primary education on school-to-work transitions for young people in rural Bolivia. Youth Soc. **36**(2), 163–182 (2004)

Robson, M.G., Stephenson, R., Elfstrom, K.M.: Community influences on antenatal and delivery care in Bangladesh, Egypt, and Rwanda. Public Health Rep. **127**(1), 96–106 (2012)

Robusto, C.C.: The cosine-haversine formula. Am. Math. Mon. **64**(1), 38–40 (1957)

Rutstein, S., Johnson, K.: The DHS Wealth Index. ORC Macro, Measure DHS (2004)

StataCorp: Structural Equation Modelling Reference Manual, Release 13. Stata Journal, StataCorp LP, College Station, TX (2013)

Schroeder, K.: Economic globalization and Bolivia's regional divide. J. Latin Am. Geogr. **6**(2), 99–120 (2007)

Silver, H.: Social exclusion and social solidarity: three paradigms. Int. Lab. Rev. **133**, 531 (1994)

Skiles, M.P., Burgert, C.R., Curtis, S.L., Spencer, J.: Geographically linking population and facility surveys: methodological considerations. Popul. Health Metr. **11**(1), 14 (2013)

Snijders, T.A.B., Bosker, R.J.: Multilevel analysis: an introduction to basic and advanced multilevel modeling. In: Lovric, M. (ed.) International Encyclopedia of Statistical Science, pp. 879–882. Springer, Berlin (2012)

Szwarcwald, C.L., Andrade, C.L.T.D., Bastos, F.I.: Income inequality, residential poverty clustering and infant mortality: a study in Rio de Janeiro, Brazil. Soc. Sci. Med. **55**(12), 2083–2092 (2002)

Theil, H., Theil, H.: Statistical decomposition analysis; with applications in the social and administrative sciences. North-Holland, London, HA33 No. 4 (1972)

Thomas, D.: Intra-household resource allocation: an inferential approach. J. Hum. Resour. **25**(4), 635–664 (1990)

US Aid: Sampling and household listing manual. DHS Toolkit of methodology for the MEASURE DHS Phase III project (2012)

Uthman, O.A., Moradi, T., Lawoko, S.: Are individual and community acceptance and witnessing of intimate partner violence related to its occurrence? Multilevel structural equation model. PLoS ONE **6**(12), e27738 (2011)

Vadez, V., Reyes-García, V., Godoy, R.A., Apaza, V.L., Byron, E., Huanca, T., Wilkie, D.: Does integration to the market threaten agricultural diversity? Hum. Ecol. **34**, 635–646 (2004)

Vandemoortele, M.: Measuring household wealth with latent trait modelling: an application to Malawian DHS data. Soc. Indic. Res. **118**(2), 877–891 (2014)

Weisbrot, M., Sandoval, L.: The distribution of Bolivia's most important natural resources and the auton-
omy conflicts. Center for Economic and Policy Research, 2 (2008)

Wight, R.G., Aneshensel, C.S., Miller-Martinez, D., Botticello, A.L., Cummings, J.R., Karlamangla, A.S.,
Seeman, T.E.: Urban neighborhood context, educational attainment, and cognitive function among
older adults. Am. J. Epidemiol. **163**(12), 1071–1078 (2006)

Wong, D.W.: Spatial indices of segregation. Urban Stud. **30**(3), 559–572 (1993)

Woods, R.: Urban–rural mortality differentials: an unresolved debate. Popul. Dev. Rev. **29**(1), 29–46 (2003)

World Bank: Bolivian data. Retrieved April 2, 2016, from http://data.worldbank.org/country/bolivia (2014)

World Bank: Urbanization trends in Bolivia, opportunities and challenges. Global Program Unit, May 2015
(2015)

Zoloth, B.S.: Alternative measures of school segregation. Land Econ. **52**(3), 278–298 (1976)