CHAPTER 4 – DEVELOPMENT OF THE GENERIC LIVING STANDARD SCALE

Objectives of the Research and Unit of Analysis

The primary objective of the research reported in this chapter is to develop a generic scale[1] suitable for measuring the living standards New Zealanders. The intention was to build on the earlier work by Fergusson et al. (2001), who developed a living standards measure for older New Zealanders called the Material Well-being Scale. As this work is an extension of the previous work, the conceptual model used to develop the measure of Material Well-being for older New Zealanders was used as a guide to develop the generic scale. The model postulates that variations in living standards (as understood in the constrained sense used here) can be validly specified on a single underlying dimension that is the source of the associations between a number of observable variables. This chapter provides an account of the development of a generic scale based on this model.

Confirmatory factor analysis is used to fit a model defined by a specified set of indicators. However, the analysis has an exploratory aspect insofar as one indicator had to be dropped from the initial set in order to achieve an acceptable fit to the data. Multidimensional models were not considered as the goal of the research was to specify a simple unidimensional scale that covered as many aspects of living standards as could be incorporated within one measure.

Unit of Analysis

In order to develop a generic scale of living standards, responses from individuals were used as the unit of analysis. The data was collected from interviews in which each respondent gave information on his or her circumstances in the context of the 'economic family unit' (EFU) of which he or she was a member. Some of the questions that were asked of the respondent (such as those about personal clothing, e.g., possession of a warm winter coat) were particular to the respondent, while others (such as those relating to non-personal, household amenities such as a washing machine) related to the respondent's EFU. In developing the generic living standard scale, questions of both types were regarded as providing information about the respondent. Thus, for example, the above illustrative items might have led to the respondent being characterised as a person who has a warm winter coat and has the advantages of being in a household with a washing machine.

These comments lay the groundwork for explaining how responses have been used to produce the results given in this report. For the purposes of analysis, the assumption has been made that it is sensible to speak of the living standard of the EFU as a whole, and that its living standard is indicated

[1] The term *generic scale* is used here to refer to a scale that is validly applicable to the population as a whole, rather than just a particular sub-group within the population.

by the score of the respondent. In other words, the members of the EFU are considered to have a broadly common standard of living, which is estimated by the respondent's score.

This assumption of a broadly common standard of living within an EFU will not always be true. Some EFUs may arrange their affairs so that some members have a lower living standard than the respondent, and others so that other members have a higher living standard. This will not distort the types of results given in this report if the departures from the assumption occur in both directions. In that case, through a process of 'swings and roundabouts', the effects will tend to average out. However, the departures constitute 'noise' in the data which will weaken the statistical associations between items and cause the structure of the data to be less clearly defined.

The Confirmatory Factor Analysis Approach

The key assumption of the model developed in the present research is that individuals and EFUs can be ranked along a continuum reflecting their living standard. The method used to examine this assumption is called structural equation modelling with latent variables (SEM: Bentler, 1989; Bollen, 1989; Joreskog and Sorbom, 1989). The use of the model requires testing whether the data meets conditions necessary for its valid application and (if the tests are met) providing a specification of the latent variable thus established to exist. The modelling approach begins by estimating the extent to which indicators of a given construct correlate with one another because they share a common source: the latent, underlying construct. This estimation is made using the method of confirmatory factor analysis (CFA). It is assumed that the indicators will not covary perfectly (i.e., the intercorrelations are less than 1.00) because of measurement error and unique aspects of the indicators (Hull, Lehn, and Tedlie, 1991). In other words, it is assumed that the pattern of covariance between the indicators can be explained effectively by their being reflections of the latent, underlying construct with no need for recourse to other constructs when account is taken of measurement error and the unique aspects of the indicators.

The structural equation modelling approach assumes that the latent variable exists independently of the particular set of indicators chosen to reveal it. In principle, the same latent variable should be able to be observed in the properties of a different set of indicators. Confirmatory factor analysis provides a rigorous means of testing the hypothesised pattern of relationships between the indicator variables and the latent variable in order to identify a robust and replicable conceptual model (see Hair, Anderson, Tatham and Black, 1995; Kerlinger and Lee, 2000; Pedhazur, 1997; Reise, Widaman, and Pugh, 1993; or Tabachnik and Fidell, 1996, for further information on SEM and CFA).

It is this procedure that was used in the present research to develop the generic scale which then served as a guide to developing the general use form of the measure. The CFA yields a regression equation that estimates values of the latent variable from the observed indicators. A second stage of the analysis (described in Chapter 5) then derives an easy-to-calculate form of the scale, giving values that closely approximate those obtained by the regression equation. This general use form of the generic scale is referred to as the Economic Living Standard Index or ELSI.

The Confirmatory Factor Model

In a typical CFA model, each indicator variable X_m, where m = 1, …, n, is represented as a linear function of one particular latent variable, ξ_p, where p = 1, …, r, and a stochastic error term δ_m. This relationship may be expressed as:

$$X_m = \lambda_{mp} \xi_p + \delta_m \qquad [1]$$

where λ_{mp} is the regression coefficient representing the regression of X_m on ξ_p and other terms are just defined. Assuming the presence of n measured variables and r latent variables and concatenating the parameters in Equation 1 into matrices leads to the following:

$$X = \Lambda \xi + \delta \qquad [2]$$

where X is a ($n \times 1$) column vector of scores of a person i on n measured variables, Λ is a ($n \times r$) matrix of loadings of the n measured variables on the r latent variables, ξ is a ($r \times 1$) matrix of factor scores of person i on the r latent ξ variables, and δ is a ($n \times 1$) matrix of measurement residuals. It is possible to show that Equation 2 implies the following equation:

$$\Sigma = \Lambda \Phi \Lambda' + \Psi \qquad [3]$$

where Σ is the ($n \times n$) population covariance matrix among the measured variables in Equation 2, Φ is a ($r \times r$) matrix of covariances among the latent variables, Ψ is a ($n \times n$) matrix of covariances among the measurement residuals or unique factors, and Λ is as just defined.

The model in Equation 3 can be fitted to a sample covariance matrix S from a sample size N, leading to:

$$S \cong \hat{\Lambda}\hat{\Phi}\hat{\Lambda}' + \hat{\Psi} = \hat{\Sigma} \qquad [4]$$

where S is the ($n \times n$) observed sample covariance matrix among measured variables and the $\hat{\Lambda}$, $\hat{\Phi}$, $\hat{\Psi}$, and $\hat{\Sigma}$ matrices contain sample estimates of the population parameters in the corresponding matrices in Equation 3. As shown in Equation 4, the observed covariances among the n measured variables in *S* are approximated by the linear CFA solution $\hat{\Lambda}\hat{\Phi}\hat{\Lambda}' + \hat{\Psi}$; this solution, in turn, produces $\hat{\Sigma}$, which contains estimates of the population covariances among the measured variables, $\hat{\Sigma}$, under the assumption that the stated model is a proper representation of the data and therefore holds for the population (Reise et al., 1993).

Equation 4 suggests the principle by which estimated CFA models can be evaluated: latent factor models imply particular covariance matrices. The statistical acceptability of the estimated CFA model depends on how close the estimated covariance matrix $\hat{\Sigma}$ is to the observed covariance matrix *S*.

Assessing the Fit of a CFA Model

The adequacy of CFA models are typically judged in two ways. Using methods of estimation (e.g., maximum likelihood), CFA programs such as LISREL and SAS provide a log likelihood ratio chi square statistic to test whether the covariance matrix reproduced from the estimated parameters, $\hat{\Sigma}$, differs significantly from the observed sample covariance matrix *S*.

The chi square statistic, however, is overly sensitive to small differences between $\hat{\Sigma}$ and *S* if the sample size is large (Bentler and Bonett, 1980; Reise et al., 1993). Hence, various indices of fit are also often used to evaluate CFA models. Although the merits of different indices are debated, it is safe to follow two principles. First, at least two indices should be calculated when developing a model. Second, no CFA model should be accepted or rejected on statistical grounds alone; theory, judgement, and persuasive argument should play a key role in determining the adequacy of any estimated model (Reise et al., 1993).

In the current application, the effective sample size of produced by combining the three surveys is approximately *N* = 3200, which is large for CFA model testing. Cautions about the over-sensitivity of the chi square with large samples and the importance of using multiple indices to assess fit are therefore particularly germane to the present analysis.

To assess fit, we used the following statistics:

1. The log likelihood ratio chi square goodness of fit test. This gives a chi squared test of the extent to which the model parameters reproduce the observed variance/covariance matrix. A non-significant chi square value is taken to indicate that the hypothesised model does not significantly differ from the data (Joreskog and Sorbom, 1993a). In practice, the chi square should be regarded as a measure of fit rather than the test statistic (Pedhazur, 1997).

2. The Adjusted Goodness of Fit Index (AGFI: Joreskog and Sorbom, 1993a). This index gives a measure of the improvement of model fit when compared with a null model in which all parameters are zero. The AGFI has a maximum value of 1.00, with this value denoting a perfectly fitting model. Well fitting models have AGFI values in excess of .90.

3. The Root Mean Square of Approximation (RMSEA: Steiger and Lind, 1980). The RMSEA is a measure of the discrepancy between the observed and the fitted data adjusted for the number of degrees of freedom (Joreskog and Sorbom, 1993b). Generally, RMSEA values less than .05 are taken as indicating a close fit of the model to the data; values of .08 reflect reasonable fit of a model.

4. The Root Mean Square Residual (RMR). The RMR is given by the sum of squares of the residual variances and covariances from the fitted model. Well fitting models have RMR values less than .05, if all variables are standardised (i.e., *S* is a correlation matrix).

Suitability of the Material Well-being Model for the Population

The same conceptual model used to develop the Material Well-being Scale (MWS) for older New Zealanders was a natural starting point for constructing the generic scale for the population. Fergusson et al. (2001) proposed a model containing six indicator variables. This was expressed through a latent variable model (see Figure 4.1) in which the six indicators were linked to a single latent dimension. These indicators were: (1) ownership restrictions; (2) social participation restrictions; (3) economising; (4) serious financial problems; and self-ratings of (5) standard of living and (6) adequacy of income[2] [3]. For indicators defined by multiple items, the respondents score on the indicator was specified as the sum of the items that comprise the set.

[2] The indicators used by Fergusson et al. (2001) were sets of items selected from those reported in Chapter 3 of the present report. The selection was made as the result of initial item analysis to weed out items which (for a variety of reasons) were unsuitable. The items included in the indicators were as follows: Ownership restrictions – secure locks; microwave; washing machine; dryer; waste disposal; dishwasher; food processor; heating in all main rooms; warm bedding; best clothes; TV; video; stereo; car; warm coat; good shoes. Social participation restrictions – funeral; family activities; give presents; visit hairdresser; holiday away; overseas holiday; night out; friends for a meal; special meal at home; enough room. Economising behaviour – less/cheaper meat; less fruit and vegetables; second-hand clothing; wore old clothes; put off buying new clothes; relied on gifts; wore worn shoes; put up with feeling cold; stayed in bed for warmth; postponed visits to doctor; postponed visits to dentist; went without glasses; not picked up prescription; cancelled insurance; less visits to family and friends; less trips to shops; less time on hobbies.

[3] It was noted in Chapter 1 that an indicator variable relating to housing was not included in the model because it was thought that it would probably preclude an acceptable fit to the data. Fergusson has indicated (personal communication) that results from subsequent supplementary analysis of the older persons' data tended to support the correctness of this expectation. The conclusion is consistent with the results of the exploratory factor analysis carried out in the 1974 study of older New Zealanders (Fergusson et al., 1974).

Figure 4.1: Conceptual model of material well-being scale for people 65 years and over

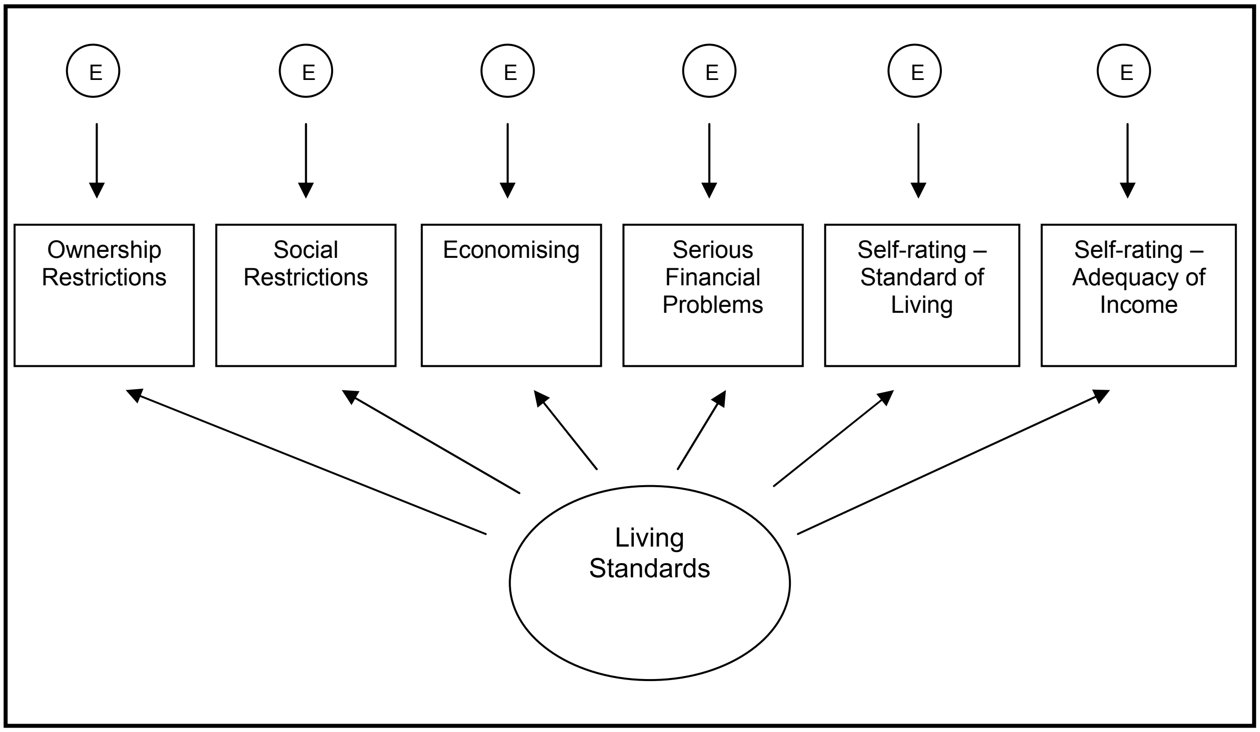
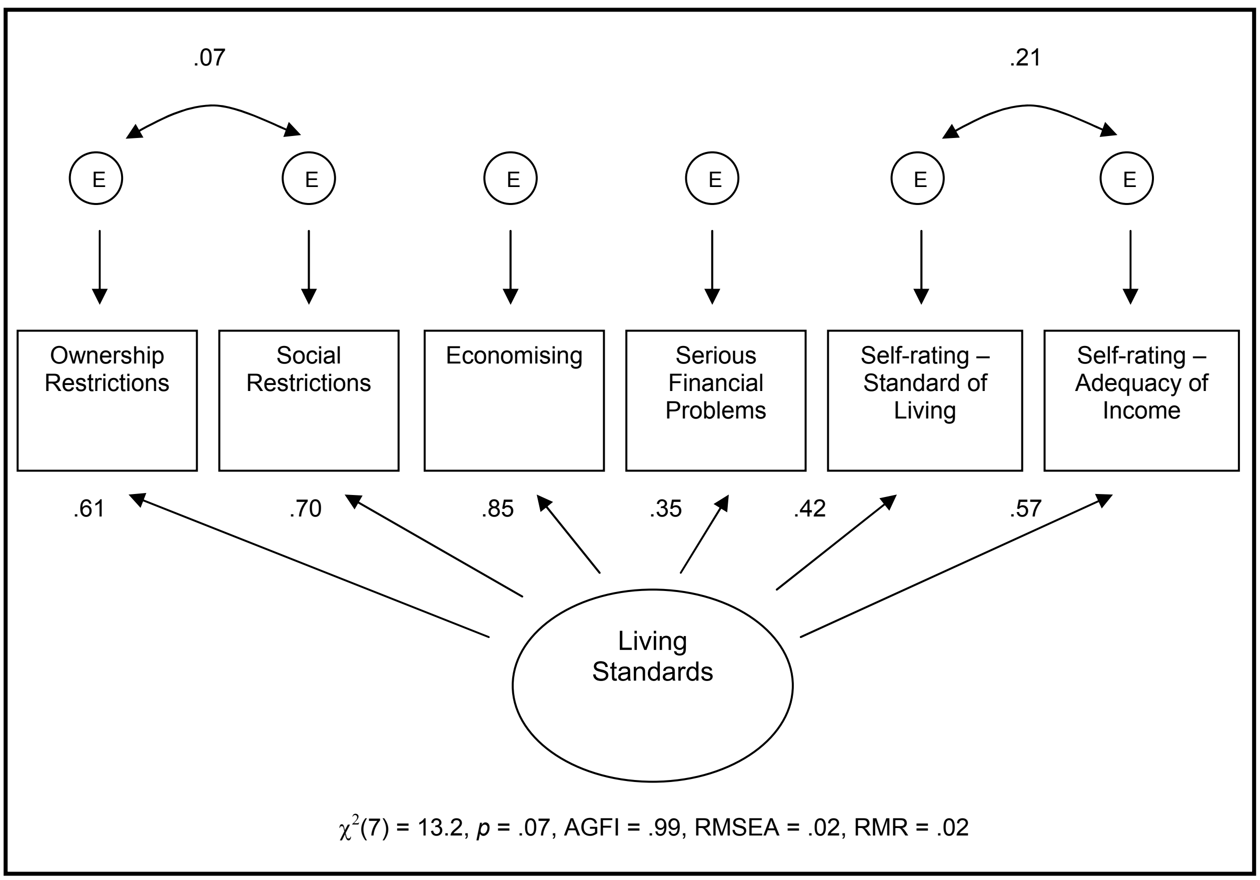


Figure 4.2: Fitted model of material well-being scale for people 65 years and over



The fitted model included an elaboration to accommodate the existence of a method effect, in which variables measured in the same way were more strongly correlated with each other than variables measured in different ways. This involved permitting non-zero correlations between the ratings for standard of living and adequacy of income, and between the ownership and social participation indicators. Figure 4.2 gives the final fitted model of material well-being among older people found by Fergusson et al. (2001). The numbers shown are the model estimates of the correlations between the latent variable (material well-being) and the indicator variables.

The approach taken in the present research was to investigate whether the index of living standards developed for older New Zealanders could be legitimately applied to the population. That is, can a generic scale of living standards be constructed for the population using the same indicators that were used for older New Zealanders? Furthermore, if such a scale can be constructed, then to what extent do the model's parameters vary across different groups within the population?

The population means and standard deviations for each of the indicator measures are given in Table 4.1, as well as the correlation matrix. Note that the ownership restriction indicator comprises the same items used in the scale of Material Well-being for older New Zealanders (Fergusson el at., 2001).

Table 4.1: Indicator measures for the population: means, standard deviations, and correlations

	1	2	3	4	5	6
1. Ownership restrictions	*					
2. Social participation restrictions	.48	*				
3. Economising	.51	.59	*			
4. Financial Problems	.41	.35	.54	*		
5. Rating – Standard of living	.33	.36	.41	.26	*	
6. Rating – Adequacy of Income	.37	.43	.50	.37	.50	*
Mean	0.80	1.07	6.53	0.52	2.62	2.34
Standard Deviation	1.55	1.38	6.71	1.16	0.78	0.93

The CFA model was tested using SAS 8.2 with asymptotic distribution (ADF) estimation used to take account of the non-normal distribution of the indicator measures (Browne, 1982; 1984). The model parameters computed by this procedure are given in Table 4.2 and 4.3 for the adult population, and for a number of sub-groups.

Table 4.2: Confirmatory factor analysis – six indicator model

	Factor Loadings						Specific Factor Covariances		Specific Factor Correlations	
Sub-groups	Ownership	Social Participation	Economising	Financial Problems	Rating of Standard of Living	Rating of Adequacy of Income	Social Participation & Ownership	SOL rating* & Ad. of Inc. rating	Social Participation & Ownership	SOL rating* & Ad. of Inc. rating
Working-age	.61	.69	.86	.62	.55	.65	.06	.15	.10	.24
Older People 65+	.59	.69	.85	.33	.42	.56	.07	.22	.11	.29
Partner	.54	.68	.84	.20	.41	.57	.10	.22	.16	.29
Single	.67	.71	.86	.41	.41	.56	.03	.21	.05	.27
Older Māori 65-69	.73	.63	.87	.64	.54	.62	.07	.08	.13	.12
Partner	.75	.65	.87	.56	.47	.64	.01	.13	.01	.19
Single	.73	.62	.87	.67	.59	.59	.11	.01	.21	.01
Total Population	.62	.71	.86	.62	.50	.62	.05	.21	.10	.30

* SOL rating = Rating of standard of living

Table 4.3: Model diagnostics and regression coefficients for representative sample and sub-group

	Model Fit Diagnostics				Regression Coefficients					
	Chi Square (7 df)	AGFI	RMSEA	RMR	Ownership	Social Participation	Economising	Financial Problems	Rating of Standard of Living	Rating of Adequacy of Income
Working-age	91.6***	.99	.06	.03	.13	.19	.48	.15	.09	.15
Older People, 65+	17.8**	.99	.02	.02	.14	.22	.55	.07	.06	.13
Partner	15.4*	.99	.03	.03	.12	.23	.55	.04	.06	.15
Single	13.2	.99	.02	.03	.18	.22	.52	.08	.05	.12
Older Māori, 65-69	12.0	.99	.04	.04	.20	.12	.48	.15	.10	.13
Partner	26.1***	.99	.10	.06	.23	.15	.48	.11	.06	.14
Single	7.1	.99	.01	.03	.19	.10	.48	.16	.12	.12
Total Population	155.1***	.99	.06	.03	.13	.19	.51	.15	.07	.12

* $p < .05$, ** $p < .01$, *** $p < .001$

The log likelihood ratio chi square goodness-of-fit statistic showed strong evidence of poor fit to the data, χ^2 (7) = 155.1, $p < .001$; AGFI = .99, RMSEA = .06, RMR = .03. Investigation of the sources of poor fit suggested the presence of residual correlation structure between the Serious Financial Problems indicator and the other indicators. Although this provides clear evidence of lack of fit, the model parameters were very similar to those obtained for older New Zealanders. The notable exception is the higher loading of the Serious Financial Problems indicator on the latent factor for working-age people than for the older people.

These results indicate that the model of a single latent dimension of material well-being reflected in each of the observed indicators (ownership restrictions, social participation restrictions, economising behaviours, serious financial problems, and self-ratings of standard of living and adequacy of income) was not appropriate when applied to the total population. The matrix of correlations between the six indicator variables could not be adequately described by a single factor model representing variation in the underlying dimension of economic living standards.

Testing a Modified Model of Living Standards

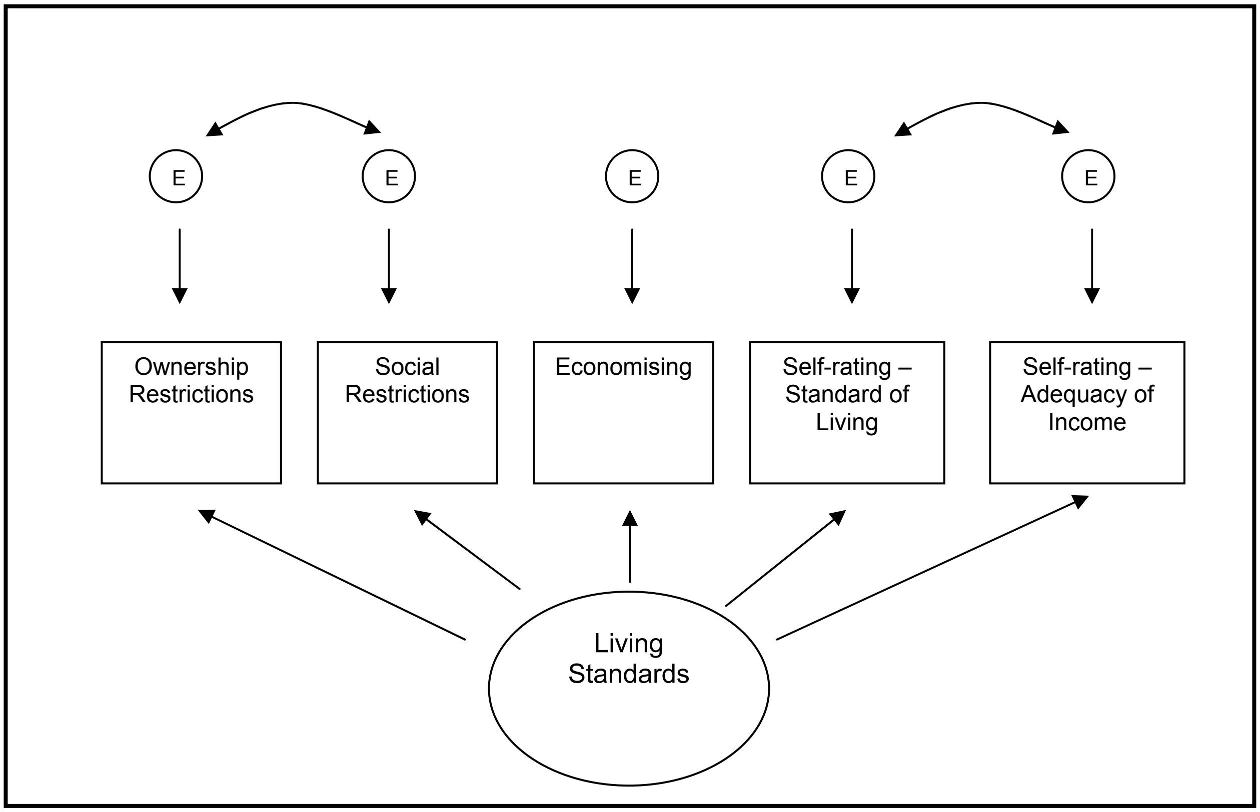
The above results from the unsuccessful attempt to obtain a unidimensional fit of the MWS model to the population pointed to the inclusion of the Serious Financial Problems indicator as the possible reason for lack of fit. Although the Serious Financial problems indicator had a substantial correlation with the estimate of the latent variable, the results suggested that the indicator may also be influenced by another factor and thus introduces an extraneous source of variation into the estimate of the latent variable. Accordingly, it was decided to test the model that resulted when this indicator was removed.

It was decided also to revisit the choice of items that comprise each indicator[4]. For instance, in the Material Well-being Scale, 16 ownership items were used to measure the indicator Ownership Restrictions; however, the survey data provided a total of 29 items that could be utilised to measure this facet of living standards.

[4] For the most part, the indicator items remained largely unchanged, with the exception of the ownership restrictions indicator where the number of items comprising the indicator was increased from 16 to 25. The items comprising the ownership indicator were as follows: secure locks; microwave, washing machine, dryer, waste disposal, dishwasher, food processor, heating in all main rooms, warm bedding, best clothes, TV, video, stereo, car, warm coat, good shoes, boat, bach, PC, internet, pay TV, good bed, contents insurance, pet, phone. The items comprising social participation were: funeral; family activities; give presents; visit hairdresser; holiday away; overseas holiday; night out; friends for a meal; special meal at home; enough room. Economising behaviour included the following: less/cheaper meat; less fruit and vegetables; second-hand clothing; wore old clothes; put off buying new clothes; relied on gifts; wore worn shoes; put up with feeling cold; stayed in bed for warmth; postponed visits to doctor; postponed visits to dentist; went without glasses; not picked up prescription; cancelled insurance; less visits to family and friends; less trips to shops; less time on hobbies.

The revised conceptual model of living standards is presented in Figure 4.3. The suitability of this model was tested using a similar procedure to that used for the Material Well-being scale: the fit between the model and the data was tested for the population, and for a number of subgroups within the population.

Figure 4.3: Conceptual model of generic living standards scale



The indices of fit showed that the model fitted the data for the population well, as evidenced by a non-significant chi square value, χ^2 (3) = 5.5, p = .14, a high level of fit (AGFI = .99), and small error statistics (RMSEA = .01, RMR = .01.) Thus, a generic model of living standards based on a single latent dimension reflected the indicators of ownership restrictions, social participation restrictions, economising, and the ratings of adequacy of income and standard of living was found to be appropriate for the population overall.

The suitability of the model was also investigated for a number of different sub-groups within the population by fitting the model to different EFU types (one-parent families, two-parent families, couple-only families, single people); Māori and non-Māori, and different age groups (18-24, 25-44, 45-64, 65+). The estimates of the model's parameters and goodness of fit statistics for the sub-groups are given in Tables 4.4 and 4.5. A good fit between the model and the data was obtained for the following sub-groups: two-parent families and sole-parent families; working-age Māori and non-Māori; 18 to 24 year olds and 25 to 44 year olds; single and coupled older people; single and coupled older Māori. There was some evidence of lack of fit for single people without children and people aged 44 to 64, as evidenced by significant chi square statistics; however, other model diagnostics showed a good fit to the data. A poor fit to the data was observed for couple-only families in the working-age population, suggesting that for this particular sub-group, the correlation structure between the five indicators cannot be completely represented by a single latent dimension of living standards. Analysis of the relative magnitude of the factor loadings for the five indicator measures was consistent across the different sub-groups. The economising subscale was consistently the best discriminator of living standards across all sub-groups, followed by the social participation subscale. The self-rating of standard of living indicator was the least effective discriminator of living standards.

The regression weights were used to compute score values for each of the sub-groups described previously and for the total population. The correlations between each sub-group specific scale and the generic measure are presented in Table 4.5. All are in excess of 0.99 indicating that the difference between the generic scale and those obtained from using a specific sub-group scale is not important as the generic scale captures much of the variation in living standards.

In summary, the generic model of living standards based on a single latent dimension reflecting the indicators of ownership restrictions, social participation restrictions, economising, and ratings of adequacy of income and standard of living was found to be appropriate for the population[5]. The model defined by those indicators is displayed in Figure 4.4.

[5] It was noted in Chapter 1 that a housing indicator variable had not been included in the MWS model because it was expected, on the basis of previous work, that it would preclude a satisfactory fit from being obtained. In Chapter 7 it is suggested that further research might usefully be undertaken on the relationship between housing measures and the living standard scale defined by the latent variable.

Table 4.4: Factor loadings and specific factor covariances for representative sample sub-groups

Population	Sub-group	Factor Loadings				
		Ownership	Social participation	Economising	Adequacy of income	Ratings of standard of living
Working-age		.64	.69	.83	.65	.56
EFU composition	Two-parent family	.64	.69	.80	.65	.56
	One-parent family	.54	.61	.90	.63	.44
	Couple-only	.64	.67	.88	.59	.51
	Single person	.63	.68	.78	.62	.61
Ethnicity	Māori	.57	.60	.80	.65	.51
	Non-Māori	.64	.71	.83	.64	.56
Age	18-24	.61	.64	.72	.54	.58
	25-44	.62	.70	.83	.68	.60
	45-64	.70	.71	.87	.67	.55
Older People, aged 65+		.67	.68	.83	.58	.44
EFU composition	Partner	.64	.67	.81	.59	.45
	Single	.71	.71	.84	.57	.43
Older Māori, aged 65-69		.72	.63	.85	.62	.52
EFU composition	Partner	.73	.64	.79	.63	.49
	Single	.72	.61	.90	.60	.52
Total Population		.66	.71	.84	.60	.50

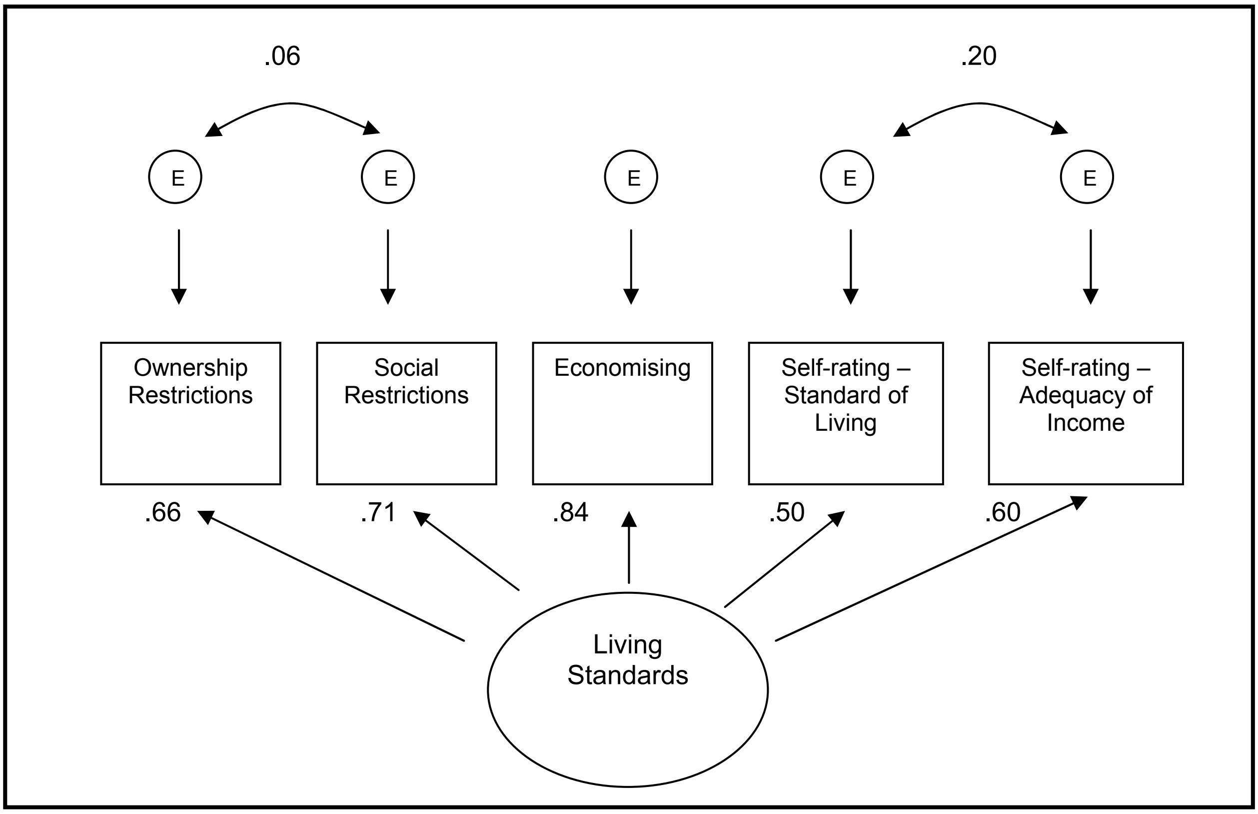
Table 4.4 continued
Factor loadings, specific factor covariances for representative sample sub-groups

Population	Sub-group	Regression coefficients					Specific factor covariances	
		Adequacy of income	Ratings of standard of living	Economising	Social participation	Ownership	Social & ownership	SOL & adequacy
Working-age		.17	.11	.47	.22	.17	.07	.14
EFU composition	Two-parent family	.19	.12	.44	.23	.18	.07	.13
	One-parent family	.14	.06	.69	.12	.09	.08	.10
	Couple-only	.12	.08	.60	.17	.15	.07	.18
	Single person	.18	.17	.40	.23	.18	.10	.11
Ethnicity	Māori	.23	.12	.49	.16	.14	.20	.10
	Non-Māori	.16	.11	.47	.24	.17	.04	.15
Age	18-24	.16	.18	.39	.23	.20	.15	.14
	25-44	.19	.13	.46	.21	.15	.07	.10
	45-64	.16	.08	.52	.19	.18	.06	.17
Older People, aged 65+		.14	.07	.50	.22	.21	.07	.20
EFU composition	Partner	.16	.08	.49	.21	.19	.12	.20
	Single	.13	.06	.48	.23	.23	.02	.19
Older Māori, aged 65-69		.16	.08	.49	.23	.18	.10	.06
EFU composition	Partner	.19	.10	.40	.18	.28	.06	.11
	Single	.13	.10	.62	.09	.18	.15	.01
Total Population		.14	.08	.49	.23	.18	.06	.20

Table 4.5: Total sample and sub-groups: measures of fit and correlation coefficients

Population	Sub-group	Chi Square (3 df)	Measures of fit AGFI	RMSEA	RMR	Correlation with generic scale
Working-age		3.2	1.00	.01	.01	.99
EFU composition	Two-parent family	3.6	1.00	.01	.01	.99
	One-parent family	2.7	.99	.01	.01	.99
	Couple-only	18.7***	.99	.07	.02	.99
	Single person	9.1*	.99	.05	.01	.99
Ethnicity	Māori	3.8	.99	.03	.01	.99
	Non-Māori	3.1	1.00	.01	.01	.99
Age	18-24	1.5	1.00	.01	.01	.99
	25-44	7.5	1.00	.03	.01	.99
	45-64	9.6*	.99	.04	.01	.99
Older People, aged 65+		4.5	1.00	.01	.01	.99
EFU composition	Partner	3.6	1.00	.01	.01	.99
	Single	3.4	1.00	.01	.01	.99
Older Māori, aged 65-69		2.0	1.00	.01	.01	.99
EFU composition	Partner	3.8	.99	.03	.01	.99
	Single	0.6	1.00	.01	.01	.99
Total Population		5.5	.99	.01	.01	

Figure 4.4: Fitted model of generic living standards scale



Multiple Group Analysis

Confirmatory factor analysis provides a framework for examining the effects of sample heterogeneity on scale parameters through the use of multiple-group modelling methods (Joreskog and Sorbom, 1993a; Muthen, 1989). This provides a basis for statistically testing whether two (or more) sub-populations can be described by the same general factor model; and whether the model parameters vary between sub-populations. This procedure provides a rigorous means of testing the stability of the scale's factor structure.

The following multiple-group models were tested:

- the effects of ethnicity and age: a comparison between working-age Māori, working-age non-Māori, older Māori, older non-Māori

- the effects of dependent children: a comparison between families with dependent children and families without dependent children for the working-age population

The multiple group analysis was conducted as follows:

1. The sample was stratified according to the different respondent characteristics described above (e.g., working-age Māori, working-age non-Māori, older Māori, older non-Māori, etc.).

2. For each of the strata, the weighted covariance matrices and means and standard deviations of the indicator measures were computed. These were used for the input data for the multiple group comparisons.

3. Alternative models were fitted to the data. The first model assumed that the model shown in Figure 4.3 described all strata and that the model parameters are the same across strata. The alternative model assumed that all strata are described by the model shown in Figure 4.3, but that model parameters vary across strata. The difference between the log likelihood statistics for the two models (and the corresponding difference in the degree of freedom) was used as a test of equality of the model parameters across strata. Where a p value less than .05 was obtained, this provided evidence that the fit of the constrained model (whereby all the factor loadings are constrained to be the same across strata) is significantly worse than the unconstrained model (where the factor loadings differ across strata), therefore indicating that the model parameters (or the factor loadings) do differ between the strata.

4. Where the procedure described above suggested differences in the model parameters, alternative models were fitted to each strata to identify which parameters were the same and which differed across the strata. The alternative models were again compared to unconstrained model, and this provided a basis of a log likelihood ratio chi square test of equality of the model parameters across the strata. Finally, the overall fit of the resulting model is examined.

Model fitting was conducted using the computer package LISREL 8 (Joreskog and Sorbom, 1993a) with Maximum Likelihood Estimation. Maximum Likelihood Estimation was used instead of ADF estimation due to the complexity of correctly accounting for sample weights in LISREL under ADF estimation. While the data was clearly not normally distributed, both estimation methods produced models with very similar parameter estimates and standard errors for the single group models obtained in the previous analyses. This suggests that for the data set being used, Maximum Likelihood Estimation would yield similar to ADF estimation.

The Effects of Ethnicity and Age

Differences in the structure of the scale between Māori and non-Māori in both the working-age and older population were investigated, and a multiple group model was fitted to the following four groups: working-age Māori (n = 457); working-age non-Māori (n = 3225); older Māori (n = 542); and older non-Māori (n = 2988). The correlation matrices and the means and standard deviations are presented below (Table 4.6).

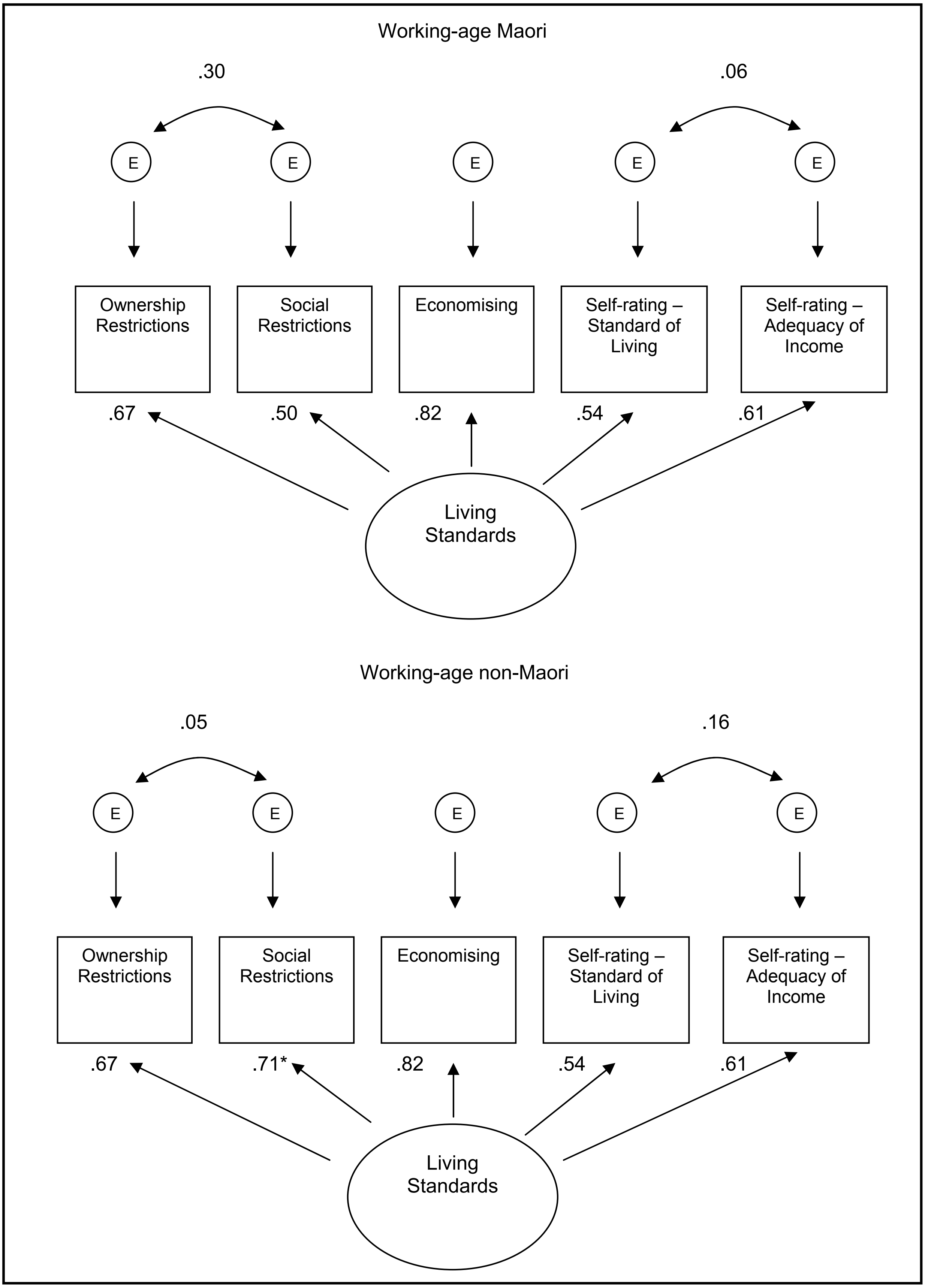
Table 4.6: Correlations, means and standard deviations – working-age Māori, working-age non-Māori, older Māori, older non-Māori

	M	SD	1	2	3	4	5
Working-age Māori							
1. Ownership restrictions	3.56	3.30	*				
2. Social participation restrictions	1.40	1.38	.52	*			
3. Economising	10.61	7.60	.42	.49	*		
4. Rating – Standard of living	2.85	0.74	.32	.32	.40	*	
5. Rating – Adequacy of Income	2.64	0.95	.38	.36	.52	.44	*
Working-age non-Māori							
1. Ownership restrictions	2.19	2.69	*				
2. Social participation restrictions	1.13	1.41	.50	*			
3. Economising	6.77	6.54	.53	.58	*		
4. Rating – Standard of Living	2.55	0.79	.37	.39	.46	*	
5. Rating – Adequacy of Income	2.28	0.94	.40	.46	.53	.50	*
Older Māori							
1. Ownership restrictions	2.47	3.19	*				
2. Social participation restrictions	1.00	1.39	.54	*			
3. Economising	7.36	8.13	.61	.51	*		
4. Rating – Standard of Living	3.02	0.72	.37	.30	.45	*	
5. Rating – Adequacy of Income	2.91	0.86	.45	.40	.52	.39	*
Older non-Māori							
1. Ownership restrictions	0.76	1.52	*				
2. Social participation restrictions	0.54	1.07	.52	*			
3. Economising	2.62	4.34	.52	.56	*		
4. Rating – Standard of Living	2.78	0.68	.30	.29	.37	*	
5. Rating – Adequacy of Income	2.43	0.85	.40	.39	.47	.46	*

Two models were fitted to the data. The first model (the constrained model) assumed that the model shown in Figure 4.3 described all four strata and that the factor loadings were the same across strata. There was strong evidence of lack of fit, χ^2 (24) = 189.1, p < .0001, AGFI = .99, RMSEA = .06, RMR = .18. The second model (the unconstrained model) assumed that the four strata were described by the model shown in Figure 4.3, but that the factor loadings varied across strata. This model showed good fit to the data, χ^2 (12) = 18.8, p = .094, AGFI = .99, RMSEA = .02, RMR=. 01. To test whether the parameters were the same across the four strata, the difference between the chi square statistics for the two models was computed. The model parameters were found to be significantly different, χ^2 (12) = 170.3, p<.0001. Therefore, alternative models were fitted to each stratum to identify which parameters were the same and which were different. These are presented in Figure 4.5. The difference between the chi square statistics for this model and the model where all factor loadings can vary across the four strata was not significant, χ^2 (6) = 9.9, p =.13. Many of the factor loadings were the same across the four sub-groups with only six of the 30 factor loadings not equivalent. The factor loadings for the ownership indicator were the same for working-age Māori, working-age non-Māori and older Māori, but different to that of older non-Māori. The factor loadings for social participation were the same for working-age Māori, and older Māori, but different to working-age non-Māori and older non-Māori. The factor loading for economising indicator was the same across the four strata. The factor loadings for the self-rating of standard of living were the same for working-age Māori, working-age non-Māori and older non-Māori, but different to that of older Māori. The factor loadings for self-rating of adequacy of income were the same for working-age Māori and non-Māori, but different to those of working-age older Māori and non-Māori[6]. The overall fit of this model to the data was very good, χ^2 (18) = 28.9, p = .052, AGFI = 1.00, RMSEA = .02, RMR = .03.

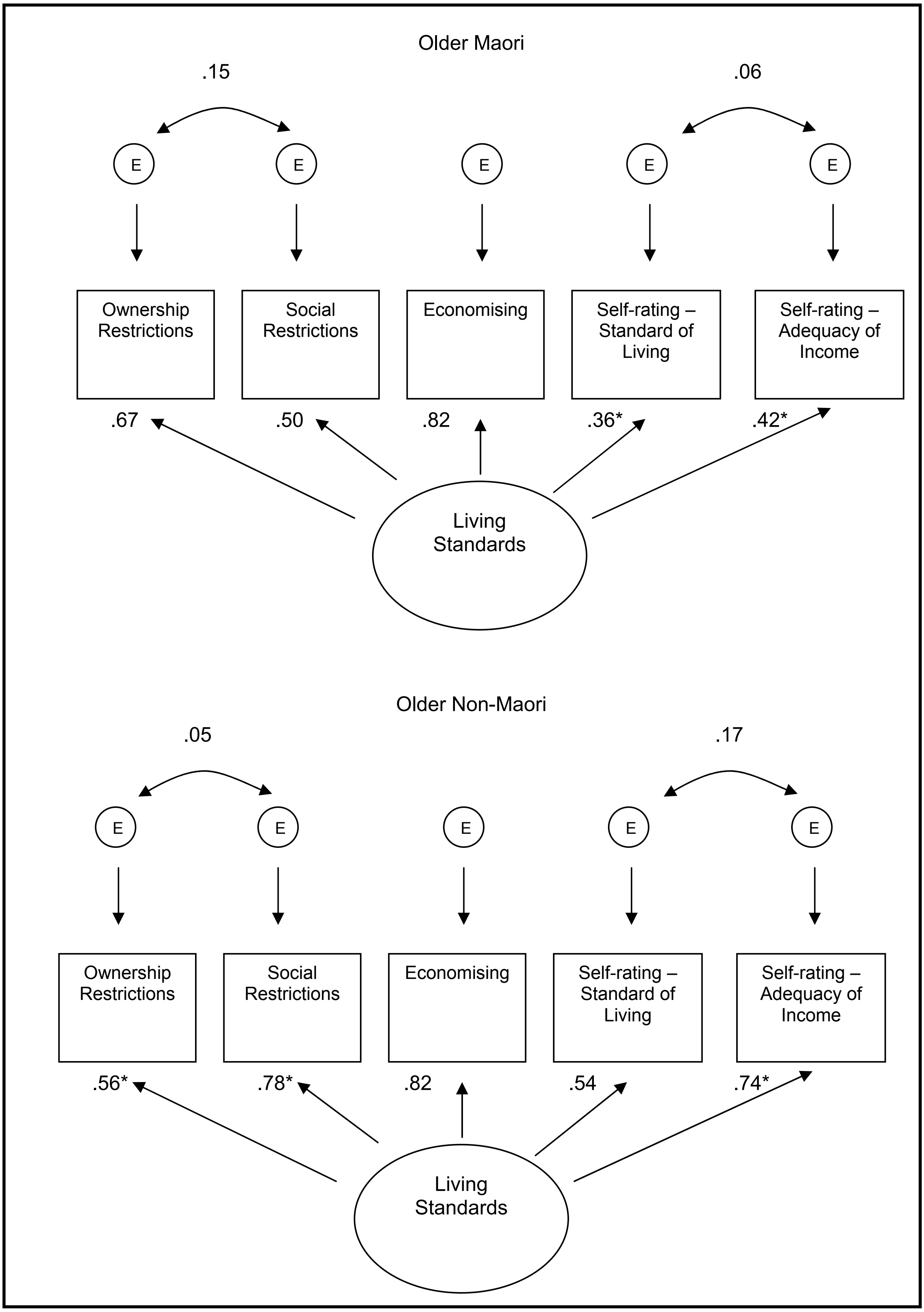
[6] When alternative scales of living standards were constructed using the model parameters that were specific to each group, it was found that these group specific scales were all correlated in excess of .99 with the generic scale obtained earlier.

Figure 4.5: Fitted model of living standards for working-age Māori, working-age non-Māori, older Māori, older non-Māori



* indicates parameters differed significantly across strata

Figure 4.5 continued: Fitted model of living standards for working-age Māori, working-age non-Māori, older Māori, older non-Māori



* indicates parameters differed significantly across strata

The Effects of Dependent Children

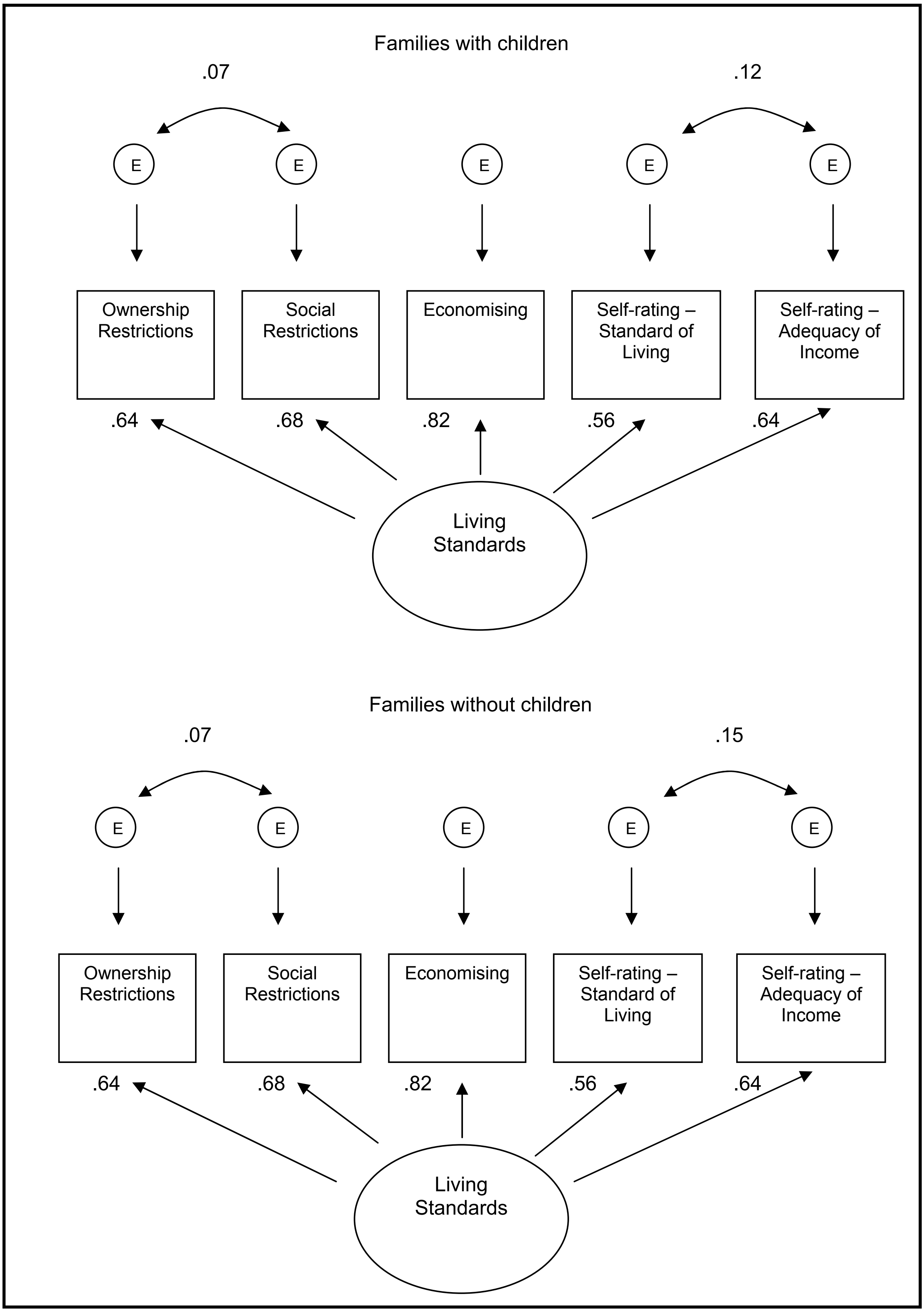
The working-age sample was stratified into families with children (*n* = 1658) and families without children (*n* = 2024). The correlation matrices and the means and standard deviations for each of these groups are present in Table 4.7.

Table 4.7: Correlations, means and standard deviations – working-age families with dependent children and working-age families without dependent children

	M	SD	1	2	3	4	5
	Families with children						
1. Ownership restrictions	2.58	2.88	*				
2. Social participation restrictions	1.45	1.48	.51	*			
3. Economising	8.59	7.08	.52	.57	*		
4. Rating – Standard of living	2.64	0.80	.38	.38	.47	*	
5. Rating – Adequacy of Income	2.43	0.94	.45	.46	.56	.51	*
	Families without children						
1. Ownership restrictions	2.18	2.74	*				
2. Social participation restrictions	0.92	1.29	.50	*			
3. Economising	6.12	6.35	.53	.55	*		
4. Rating – Standard of Living	2.54	0.78	.36	.39	.44	*	
5. Rating – Adequacy of Income	2.23	0.94	.36	.41	.51	.49	*

As with previous analysis, two models were fitted to the data. The first model (the constrained model) assumed that the model shown in Figure 4.3 described both strata and that the model parameters were the same across strata. There was some evidence of lack of fit of this model to the data, χ^2 (10) = 19.3, *p* = .037, AGFI = .99, RMSEA = .02, RMR = .10. The second model (the unconstrained model) assumed that all strata were described by the model shown in Figure 4.3 but that the factor loadings varied across strata. Overall, this model showed an improved fit to the data and there was little evidence of lack of fit, χ^2 (6) = 15.1, *p* = .020; AGFI = .99, RMSEA = .03; RMR = .04. The difference between the log likelihood statistics for the two models tests whether the factor loadings were the same across strata, and was not significant, χ^2 (4) = 4.2, *p* = .20. This result provided no evidence of the factor loadings varying across strata; rather, respondents with children and those without children can be described by the same factor model (see Figure 4.6).

Figure 4.6: Fitted model of living standards for working-age families with dependent children and without dependent children



* indicates parameters differed significantly across strata

In sum, the multiple group analysis suggested the following similarities and differences between models obtained for the various strata.

1. Model structure: The same model structure was generally appropriate across the various sub-groups. All indicator measures were related to the underlying dimension of the scale, and the model permitted correlation between the error terms of: a) ownership restrictions and social participation restrictions; and b) ratings of standard of living and adequacy of income.

2. Model parameters: There were some small between-strata differences in the factor loadings; however, these differences do not lead to appreciable differences in the scale score values. For instance, when alternative scales were constructed using the factor loadings and corresponding factor score coefficients that were specific to each sub-group it was found that the sub-group specific scales were all correlated above .99 with the generic scale.[7]

The results of multiple group comparisons support the conclusion that the generic scale defined by the five indicator model is appropriate for the population.

Generic Scale Reliability and Validity

Reliability refers to consistency of measurement – either across time (test-retest reliability), across different forms of a scale (parallel form reliability), or across items (internal consistency). When performing confirmatory factor analysis, it is possible to compute the composite reliability for each latent variable included in the model. This index is analogous to coefficient alpha, and reflects the internal consistency of the indicators measuring a given factor (Hatcher, 1998). The reliability of a test will be high if the various items or indicators that constitute the scale are strongly correlated with each other. The reliability of the generic scale was estimated at .79, suggesting an acceptable level of internal consistency.

According to Cronbach (1990) validation is an inquiry into the soundness of interpretations proposed for scores from a test. Kline (1998) proposes a more specific definition of validity,

[7] One unanswered question from this analysis concerns whether the living standards of families with children can be adequately described by a general scale measure, or would a measure which includes items specific to children have greater utility? A scale that excludes restrictions specific to children would need to adequately reflect the relative well-being of families with children. If it did not do so, then a separate scale would need to be developed which included the children specific restrictions. In the survey of working-age people, respondents with children were asked about restrictions specific to children. Of these, 12 were economising restrictions, four were ownership restrictions and four related to social participation restrictions (see Tables 3.2 to 3.10 in Chapter 3 for a complete list of common and children sub-population specific restrictions collected). To test whether a scale that does not include children specific restrictions adequately reflects the relative well-being of economic family units with children, an alternative scale measure was constructed whereby the indicator measures comprised both the common and children specific restrictions. This was done separately for respondents who were single and partnered. The correlation between two scales was used to assess whether the general scale measure adequately captured variation in living standards among families with children. The alternative scale measures were found to be highly correlated for both single respondents with children (r = .96) and partnered respondents with children (r = .98) These result suggests that for practical intent and purposes, families with children can be scored on the same general scale measure defined in Figure 4.3.

arguing that a test is valid if it can be used for all purposes to which the test legitimately might be put. Encapsulated within these definitions is the idea that a test must be evaluated against some criteria to determine whether it measures what it was intended to measure.

The validity of the generic scale was tested by examining its association with a number of variables that can be expected to be associated with living standards. Table 4.8 shows the correlation of the generic scale and seven validation measures that were gathered in the surveys. These measures are

1. housing-adjusted equivalised disposable income
2. equivalised disposable income[8]
3. for older people (aged 65+): whether the respondent reported being unable to save on most months
4. whether the respondent reported being unable to find $5000 in an emergency
5. for older people (aged 65+): whether the respondent reported health-related financial stress in the past 12 months
6. for older people (aged 65+): whether the respondent reported being in possession of a Community Services Card
7. for older people: Material Well-being Scale

[8] Equivalent income is obtained as an income transformation using income equivalence ratios, which are taken as specifying the relative amounts of income required by households of different sizes to achieve the same standard of living. There are various ‘income equivalence scales’ that have been used for that purpose. If a single adult household type is taken as the reference type, and assigned the value of 1, then the equivalence ratios of for other household types can be interpreted as measures of ‘effective size’. For example, if the ‘effective size’ of a couple + child household type is 1.9 (say), this indicates that such a household requires 1.9 times the income of the single adult to achieve the same standard of living. Equivalised income can be calculated as (income/‘effective household size’). Conceptually it is similar to per capita income but differs in recognising that households achieve economies of scale, so that a household of four people (say) requires less that double the income of a household of two people to achieve the same standard of living.

The calculation of equivalised income was made using the Revised Jensen Scale (RJS: see Mowbray, 2001). In the present application, the unit of equivalised income has been calculated by first subtracting the EFU’s net accommodation outgoings from its disposable income and then equivalising the amount obtained. There are reasons for expecting that housing adjusted disposable income would correlate more highly with the generic scale than would equivalised disposable income (see Jensen and Krishnan, 2001).

Table 4.8: Correlation coefficients between confirmatory factor model and validation measures

Measure	Correlation
1. Housing adjusted equivalised disposable income	-.52
2. Equivalised disposable income	-.43
3. For older people: Unable to save	.40
4. Unable to get $5000	.52
5. For older people: Health-related stress	.43
6. For older people: Possession of Community Services Card	.33
7. For older people: MWS	-.96

Note: Confirmatory factor model scored so that an increasing score implies decreasing standard of living. All correlations are significant, $p < .0001$.

As expected, the table shows substantial negative correlations between the confirmatory factor model and housing adjusted equivalised disposable income and equivalised disposable income. Thus, higher living standard scores are associated with higher equivalised disposable income, and its derivative housing adjusted equivalised disposable income. Similarly, indicators of hardship were positively correlated with the confirmatory factor model. Inability to save money most months, being unable to find $5000 in an emergency, having health-related financial stress, and possession of a Community Services Card were all related to lower living standard scores. These results provide tentative support for the validity of the generic scale. However, as with all examinations of test validation, it is important to replicate findings to more firmly establish a scale's validity. It is also desirable, as opportunities become available through future research, to make further tests based on other additional validational variables.

Concluding Comment

This chapter has documented the development of a generic living standards scale using the methods of confirmatory factor analysis. The analysis has centred on testing the premise that living standards can be expressed as an unobservable latent variable with properties that are reflected in a series of observable indicator variables.

The analysis began by restating the CFA theoretical model that had been found to give an acceptable fit for older New Zealanders (Fergusson et al., 2001). That model, which comprised six indicator variables, failed to give an acceptable fit for the population. The lack of fit was attributed the presence of one particular indicator (serious financial problems) which showed a residual correlation with other variables after the removal of the common variance accounted for by the latent variable. Accordingly, the serious financial problems indicator was removed. Figure 4.3 shows the resulting revised model, which had five indicators reflecting a single latent variable. These indicators were ownership restrictions, social participation restrictions, economising behaviour, and ratings of adequacy of income and standard of living. Correlated errors were permitted between the ownership restrictions indicator and social participation indicator, and between the adequacy of income indicator and standard of living indicator. This model had a good fit to the data for the adult population and for most sub-groups within the population. Multi-group analysis suggested small differences in the factor loadings of some indicators across different strata but these differences did not lead to appreciable differences in score values. Tests of reliability and validity supported the generic scale's utility as a measure.