CHAPTER 22

A Comparison of Graphical Models and Structural Equation Models for the Analysis of Longitudinal Survey Data

Peter W. F. Smith, Ann Berrington and Patrick Sturgis
University of Southampton, UK

22.1 INTRODUCTION

Graphical chain modelling (GCM) and structural equation modelling (SEM) are two approaches to modelling longitudinal data. Both approaches have their origins in path analysis and provide pictorial representations of the association between variables, which are usually ordered temporally. Both methods also aim to identify the direct and indirect effects of one variable on another. While the SEM approach specifies a single model for the complete system of variables being studied, the GCM approach permits a model for the complete system to be built up by fitting a sequence of submodels.

In this chapter we briefly discuss the similarities and differences of the GCM and SEM approaches to modelling univariate recursive systems and their application to complex survey data. We identify the strengths and limitations of each approach. By using an example of the relationship between changes in labour force status following entry into parenthood, and changes in gender role attitude, we illustrate their use with survey data. A sample of 632 women, childless and aged between 16 and 39 in 1991, taken from the British Household Panel Survey (BHPS), is used. This survey is a particularly useful resource for this type of analysis since it is nationally representative, continuous since 1991 and collects a wide range of socio-demographic and attitudinal information (Taylor *et al.*, 2007). Information on employment and parenthood status is obtained annually and attitude statements concerning gender roles are collected biennially in 1991, 1993, 1995 and 1997.

22.2   CONCEPTUAL FRAMEWORK

Recent interest has focused on the role that values and attitudes play in life-course decisions. For example, in the UK there has been considerable debate on the extent to which women's attitudes towards work and family care are exogenous (Crompton and Harris, 1998; Hakim, 1998). Whilst numerous studies based on cross-sectional data have demonstrated significant differences in attitudes according to current family formation status, panel data are required to identify the potential causal direction of the relationship. As noted by Lesthaeghe and Moors (2002), in most longitudinal analyses researchers have either focused on selection effects of attitudes (attitudes as predictors of later events) or on adaptation effects of attitudes (events as determinants of change in attitudes). For the analyses in this paper, we focus on the reciprocal relationship between changes in labour force status (following entry into parenthood) and changes in gender role attitude, and how we can best model it using longitudinal survey data.

The conceptual framework, presented in Figure 22.1, conjectures how the variables are related to each other. (See Berrington *et al.* (2008) for a fuller description.) Here the set of variables is partitioned into blocks. The first block contains our background control variables (age, highest level of education and whether or not the respondent's mother worked when the respondent was 14 years old). The rest of the blocks are temporally ordered. Hypothesized associations are represented by arrows. For example, gender role attitude in 1991 is hypothesized to explain variations in the risk of becoming a parent between 1991 and 1993, which is a selection effect. Becoming a parent between 1991 and 1993 is hypothesized to be associated with changes in gender role attitude during the same period, which is an adaptation effect. Women are thought likely to adjust their working hours upon entry into motherhood. Labour force participation may also be predicted directly by gender role attitudes. The background variables may predict any of the later outcomes, including initial attitude, attitude change and the likelihood of becoming a parent or of leaving the labour force. For clarity, these hypothesized relationships are indicated by the large black arrow in Figure 22.1.
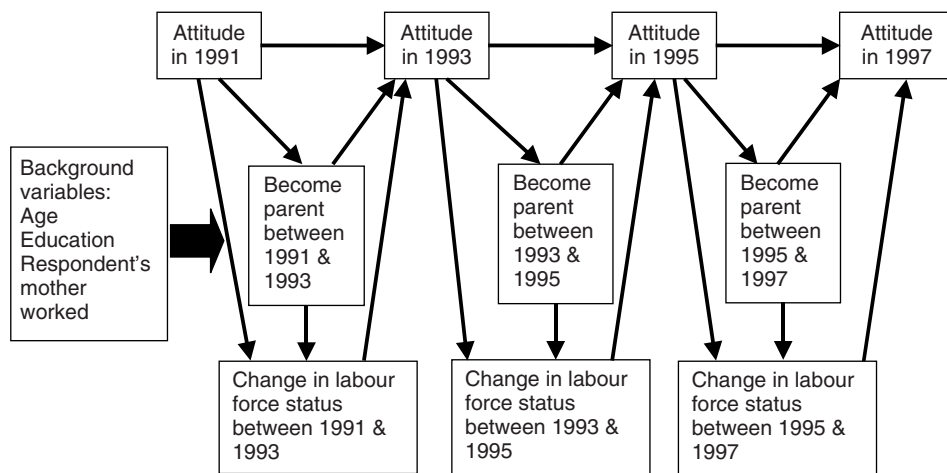

Figure 22.1   Conceptual framework.

Gender role attitude is derived from six items, each measured on a five-point scale from strongly agree to strongly disagree, asked biennially in the BHPS from 1991. The six questions are: (1) a preschool child is likely to suffer if his or her mother works; (2) the whole family suffers when the woman has a full-time job; (3) a woman and her family would all be happier if she goes out to work; (4) both the husband and wife should contribute to the household income; (5) having a full-time job is the best way for a woman to be an independent person; and (6) a husband's job is to earn money, a wife's job is to look after the home and family. Becoming a parent is defined as having a first live birth and hence refers solely to biological parenthood. Change in labour force status is based on changes in the number of hours worked. Movement from being a full-time student into paid work is included as an increase in work hours and movement in the other direction as a decrease. Becoming unemployed or economically inactive due to illness is defined as a reduction in work hours, whilst leaving the paid labour force to take on family responsibilities is referred to as leaving for family care. In summary, change in labour force status has four categories: (1) no change in hours of paid work, (2) increase in hours of paid work, (3) decrease in hours of paid work, and (4) begin family care.

22.3   GRAPHICAL CHAIN MODELLING APPROACH

The graphical chain model (see, for example, Whittaker, 1990; Cox and Wermuth, 1996) is a stochastic model specified via a mathematical graph that is a set of nodes and edges. Nodes represent variables and undirected edges (lines) represent significant associations between pairs of variables. Asymmetric relationships between variables, i.e. one anticipating the other, are represented through directed edges (arrows). Chain graphs allow us to specify directions of causality. Variables are entered into the model in a series of blocks. These reflect the temporal ordering of the variables and the assumed causal ordering of relationships. Hence, directed edges can only flow from variables in a preceding block to variables in subsequent block(s). Fundamental to graphical modelling is the concept of conditional independence. In a graphical chain model a missing edge or arrow corresponds to a conditional independence between the two variables given all the variables in the current and preceding blocks. Graphical chain models are ideally suited to longitudinal data since the temporal ordering of variables can be used to help specify the *a priori* causal ordering of variables. For some applications of graphical chain models to survey data, see Mohamed *et al.* (1998), Magadi *et al.* (2004) and Evans and Anderson (2006).

A graphical chain model can be fitted to our data by fitting a sequence of regression models. First, gender role attitude in 1991 is regressed on the background variables. Second, becoming a parent between 1991 and 1993 is regressed on the background variables and attitude in 1991. Next, change in labour force status between 1991 and 1993 is regressed on the background variables, becoming a parent between 1991 and 1993 and attitude in 1991. The sequence continues by regressing each of the 'attitude', 'becoming a parent' and 'change in labour force status' variables in turn on the background variables and the hypothesized determinants as indicated in the conceptual framework in Figure 22.1.

To permit the use of standard regression techniques when fitting the graphical chain model, we restrict our attention to modelling the observed variables. Hence, we first

created a gender role attitude score by coding the responses to each of the six items, in such a way that a higher score indicates a more liberal gender role attitude, and then summing them. The score ranges from 6 to 30, with a mean in 1991 of 21.2 and a standard deviation of 3.7.

One advantage of graphical chain models is their ability to handle response variables of different types: continuous, binary and multicategory. When the response is continuous, as is the case for the attitude score, we use linear regression models. When the response is binary, as is the case for whether or not the respondent becomes a parent, or multicategory, as is the case for the respondent's change in labour force status, we use binary and multinomial logistic regression models, respectively. Furthermore, sample weights, clustered samples and item nonresponse can be handled using the standard techniques developed for the component models. A disadvantage in the GCM approach considered here is that it does not use the information in the repeated measures or model the reciprocal relationship simultaneously. Hence, it is not easy to assess the overall goodness of fit of the model to the data. Also, it is difficult to explicitly adjust for measurement error without introducing a separate stage of latent variable modelling.

22.4   STRUCTURAL EQUATION MODELLING APPROACH

SEM is a framework for statistical analysis that brings together confirmatory factor analysis, path analysis and regression (see, for example, Bollen, 1989; Finkel, 1995). SEM can usefully be conceived of as having two components, or stages. The first is the measurement model in which the relationships between a set of observed (manifest) variables and a set of unobserved (latent) variables are specified. The second is the structural model, which specifies causal and associational relationships between measured constructs, whether latent or observed.

We specify a measurement model for our data in which gender role attitude is a latent variable measured by the six indicator items. The latent variable represents a common factor that influences the observed items. The square of the standardized factor loading of an item reflects the proportion of the item variance that is explained by the common factor. By definition, one minus the square of the factor loading is the proportion of variance in the item that is measurement error. Relative to summing scores over the individual items to form a composite scale, as is common practice, the latent variable approach enables us to make a correction for differential measurement error across items, which is reflected by variation in the magnitude of the factor loadings. Measurement error in independent variables is known to result in attenuated effect sizes, so this error correction should reduce the probability of Type II errors in our anlayses (Bollen, 1989). Because we have repeated measures of the same items on the same individuals over time, we estimate covariances between the error terms for the same indicator at each time point, as the unmeasured causes of responses to each item are likely to remain stable over time. Failing to account for covariances between error terms in panel studies can bias estimates of both stability and cross-lag parameters (Williams and Podsakoff, 1989).

In summary, SEM includes the measurement models for attitude in the four waves, all of the structural paths tested for in GCM and correlations between the errors in the repeated attitude items over time. All of the regression parameters and error correlations

are estimated simultaneously. As well as reducing measurement error, another advantage of the SEM approach is that it uses the information in the repeated measures and models the reciprocal relationship simultaneously. Hence, it is possible to measure the overall goodness of fit of the model to the data and it is straightforward to test for equality of effects across time. A disadvantage in the SEM approach is that it cannot easily handle unit nonresponse, attrition, sample weights and clustering. Also, categorical response variables, particularly those with more than two categories, can cause problems for maximum likelihood estimators. However, recent developments in these areas have substantially improved the flexibility of SEM models for dealing with complex survey data, nonresponse and categorical endogenous variables (Muthén, 2001). Mplus (Muthén and Muthén, 1998–2007) and LISREL (Scientific Software International, 2006) are two statistical package that are more advanced in these areas (Stapleton, 2006).

22.5   MODEL FITTING

Initial modelling revealed a very strong association between becoming a parent and change in labour force status. The vast majority of women taking on family care between one time period and the next were new mothers. Preliminary modelling revealed that becoming a parent was not, of itself, significantly associated with attitude change but that it was the change in labour force status, particularly the change to family care, that was significant in the graphical chain model. Therefore, only a binary variable contrasting 'begin family care' with 'did not begin family care' was included in the final graphical chain and structural equation models. Note that the binary variable, whether or not the respondent begins family care, is estimated using a probit regression in the SEM framework and a logistic regression in the GCM framework.

The BHPS is a stratified and clustered sample with weights to adjust for unequal selection probabilities and panel attrition (see Taylor *et al*., 2007, for details). In the first instance, to permit comparison of the GCM and SEM approaches, these complexities are ignored. We refer to this as the 'simple' approach. To identify whether taking into account the complex survey design and unit response would make a difference to our substantive conclusions we repeat the GCM analyses taking account of the stratification (using region as the stratifying variable), clustering and weighting. We refer to this as the 'complex' approach. The sample is restricted to women who gave a full interview in every year from 1991 to 1997. Since our sample considers only childless women of childbearing age, only a small percentage belong to a household containing at least one other sample member. Hence our cluster variable refers to the primary sampling unit – i.e. postcode sector. The wave 7 longitudinal respondent weights take account of initial unequal selection probabilities into the BHPS sample and subsequent differential nonresponse at waves 1–7. Linearization-based variance estimators (Binder, 1983; Skinner, 1989) are used to produce robust estimates that are reported alongside the naïve estimates in Tables 22.1 and 22.2. (Note that we also repeated the SEM analyses taking account of complex survey design and weights, and found that the impact on the point estimates and standard errors is very similar to that with GCM.)

The GCM analyses were undertaken using the Stata software with the svy commands (Stata Corp, 2007) when taking account of the complex survey design and weights. The SEM analyses, including those incorporating adjustments for stratification, clustering

and weights, were undertaken using the Mplus software (Muthén and Muthén, 1998–2007). The full SEM results are available from the authors on request.

22.6 RESULTS

Figure 22.2 depicts the selected graphical chain model using the estimates from the 'simple' approach. Here the full lines indicate significant positive effects (i.e. attitude becomes more liberal) and the dashed lines indicate significant negative effects (attitude becomes less liberal), both at the 5 % level. The corresponding diagram for the structural equation model is the same except that attitude in 1995 is no longer a significant predictor of beginning family care between 1995 and 1997. A test for the equality of the effects of lagged attitude on attitude is significant (meaning that the effect of lagged attitude changes over time), whereas there is no evidence against equality of the effects of beginning family care on attitude, that is, after controlling for the other variables the adaptation effect is of the same magnitude over time.

Table 22.1 presents the parameter estimates for the four linear regressions used to construct the graphical chain model. For each year the first column present the estimates from the 'simple' approach and the second column contains those from the 'complex' approach. Note that the standard error appears in parentheses beneath each estimate. Here we find that all three background variables are associated with gender role attitude in 1991, with younger, more educated women whose mothers worked when they were aged 14 years tending to have more liberal attitudes. For 1993, 1995 and 1997 we are modelling change in attitude, since attitude two years previous is included in the model. Here we find that, in general, background variables are not associated with changes in attitude, whereas women who begin family care tend to become more traditional. Hence, there is a significant adaptation effect. Table 22.2 presents the parameter estimates for the logistic regression of beginning family care. The evidence for a selection effect is weaker. If there is one, then, as expected, more traditional women are more likely to give up work to undertake family care.
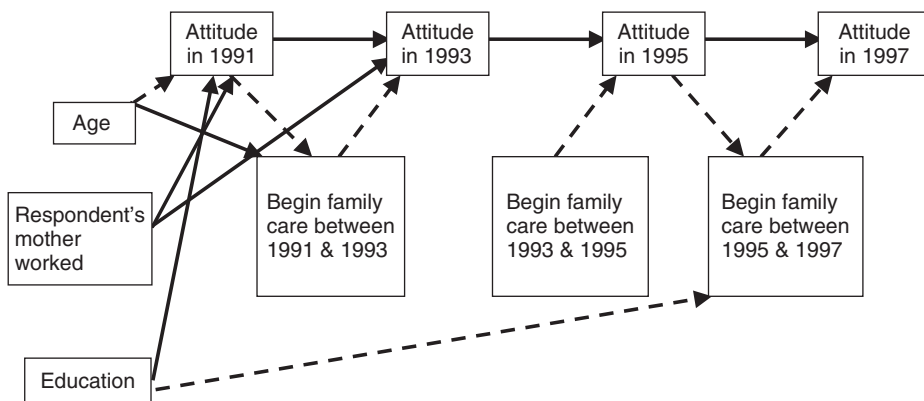

Figure 22.2 Significant associations found in the GCM using the 'simple' approach. Full lines indicate significant positive effects and dashed lines indicate significant negative effects, both at the 5 % level.

Table 22.1 Parameter estimates for the linear regression models for attitude: 'simple' and 'complex' approaches compared.

	Attitude in 1991		Attitude in 1993		Attitude in 1995		Attitude in 1997	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Constant	20.72**	20.87**	9.70**	10.01**	6.47**	6.33**	6.76**	6.43**
	(0.33)[a]	(0.38)	(0.69)	(0.93)	(0.69)	(0.69)	(0.64)	(0.75)
Age (Ref: 16–21 years)								
22–29	−0.98**	−1.14**	−0.06	−0.13	−0.49*[†]	−0.56*[†]	−0.25	−0.18
	(0.33)	(0.33)	(0.26)	(0.26)	(0.25)	(0.27)	(0.24)	(0.28)
30–39	**−1.12****	**−1.11***	−0.43	−0.43	−0.03	−0.12	−0.44	−0.42
	(0.39)	**(0.43)**	(0.30)	(0.32)	(0.29)	(0.32)	(0.28)	(0.30)
Education in 1991 (Ref: Less than A level)								
A level or above	1.03**	1.01**	0.06	0.32	0.27	0.18	0.16	0.26
	(0.29)	(0.30)	(0.22)	(0.22)	(0.22)	(0.23)	(0.21)	(0.23)
Mum worked (Ref: No)								
Yes	1.03**	1.00**	0.68**	0.67**	0.24	0.30	0.33	0.37
	(0.27)	(0.30)	(0.23)	(0.25)	(0.23)	(0.24)	(0.22)	(0.21)
Attitude two years previous			0.52**	0.50**	0.68**	0.68**	0.65**	0.66**
			(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Began family care in the preceding two years (Ref: No)								
Yes			−1.58***	−1.54**	−1.91**	−1.97**	**−1.34****	**−1.15***
			(0.41)	(0.46)	(0.41)	(0.55)	**(0.42)**	**(0.49)**

*Significant at 5 % level; **significant at 1 % level; [†]overall age is not significant at the 5 % level. Figures in bold denote a change in statistical significance level between the 'simple' and 'complex' approach.

[a] Standard errors are naïve for the unweighted parameter estimates and robust for the weighted estimates.

Table 22.2   Parameter estimates for the logistic regression models for beginning family care: 'simple' and 'complex' approaches compared.

	Begin family care between 1991 and 1993		Begin family care between 1993 and 1995		Begin family care between 1995 and 1997	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Constant	− 0.32	0.27	− 1.53	− 1.37	− 0.30	0.23
	(0.89)	(1.05)	(0.97)	(0.99)	(0.95)	(0.81)
Age (Ref: 16–21 years)						
22–29	**1.10****	**0.94***	0.27	0.17	0.02	− 0.20
	(0.40)	**(0.44)**	(0.37)	(0.39)	(0.36)	(0.36)
30–39	0.63	0.58	0.20	0.04	**− 0.81**	**− 1.09***†
	(0.47)	(0.55)	(0.43)	(0.43)	**(0.53)**	**(0.52)**
Education in 1991 (Ref: Less than A level)						
A level or above	**− 0.37**	**− 0.58***	0.19	0.07	− 0.78*	− 0.82*
	(0.31)	**(0.29)**	(0.32)	(0.30)	(0.35)	(0.33)
Mum worked (Ref: No)						
Yes	0.29	0.25	− 0.24	− 0.04	0.09	− 0.23
	(0.32)	(0.35)	(0.32)	(0.29)	(0.34)	(0.39)
Attitude at	− 0.14**	− 0.16**	− 0.06	− 0.06	**− 0.10***	**− 0.11****
beginning of period	(0.04)	(0.04)	(0.05)	(0.05)	**(0.05)**	**(0.04)**

*Significant at 5 % level; **significant at 1 % level; †overall age is not significant at the 5 % level. Figures in bold denote a change in statistical significance between the 'simple' and 'complex' approach.

Adjustments for stratification and clustering generally impact on the standard errors but not on the point estimates, whereas including weights affects the magnitude of both the point estimates and the standard errors (see Skinner *et al.*, 1989). We find that the weighted parameter estimates are generally similar to the unweighted estimates and that the robust standard errors are generally larger than those assuming a simple random sample, but this is not always the case. In the linear regression models for attitude there are two parameters where the estimates become less significant, moving from the 1 % to 5 % level (indicated by bold figures in 22.1). In the logistic regression models for beginning family care, there is one parameter where the estimate becomes less significant and three parameters where the estimates become more significant. However, only in one case do we find contradictory evidence for the presence of an edge after accounting for the complex sample design and weights: the association between education and beginning family care between 1991 and 1993 becomes significant at the 5 % level (with a slightly more negative parameter estimate and slightly smaller standard error). Note that adjusting for the complex sample design and weights within the SEM framework had a similar effect, that is, changes were of a similar magnitude and in the same direction.

Standardized parameter estimates from the SEM are presented in Tables 22.3 and 22.4. They have been standardized only with respect to the latent attitude variables. Hence, in Table 22.3, they are the estimated average amount (in standard deviations) by which the latent attitude variables change for a one-unit change in the predictor variable. For attitude two years previous, this corresponds to a change of one standard deviation,

since all the latent attitude variables have been standardized, whilst for the categorical variables this refers to the difference between the category of interest and the reference category. For example, for whether or not the respondent began family care, it is the difference between "yes" and "no". For comparison, standardized parameter estimates for the (naïve) linear and logistic regression models used to construct the GCM are also presented in Tables 22.3 and 22.4. Here the attitude variables constructed by summing

Table 22.3 Standardized parameter estimates for the regressions of attitude (using the 'simple' approach).

	Attitude in 1991		Attitude in 1993		Attitude in 1995		Attitude in 1997	
	Linear regression	SEM	Linear regression	SEM	Linear regression	SEM	Linear regression	SEM
Age 22–29	− 0.269**	− 0.272**	− 0.017	0.104	− 0.137*†	− 0.048	− 0.071	− 0.015
Age 30–39	− 0.305**	− 0.319**	− 0.124	− 0.106	− 0.009	0.099	− 0.123	− 0.143
Education in 1991								
A level or above	0.282**	0.396**	0.017	− 0.009	0.074	0.050	0.046	− 0.042
Mum worked	0.282**	0.340**	0.198**	0.229**	0.066	− 0.022	0.093	0.056
Attitude two years previous			0.552**	0.717**	0.642**	0.839**	0.664**	0.829**
Began family care			− 0.458**	− 0.169**	− 0.530**	− 0.308**	− 0.377**	− 0.187*

*Significant at 5 % level; **significant at 1 % level;
†overall age is not significant at the 5 % level.

Table 22.4 Standardized parameter estimates for the regression of beginning family care (using the 'simple' approach).

	Begin family care between 1991 and 1993		Begin family care between 1993 and 1995		Begin family care between 1995 and 1997	
	Logistic regression	SEM	Logistic regression	SEM	Logistic regression	SEM
Age 22–29	0.609**	0.517**	0.149	0.123	0.012	0.007
Age 30–39	0.346	0.280	0.108	0.079	− 0.445	− 0.378
Education in 1991						
A level or above	− 0.202	− 0.131	0.103	0.098	− 0.431*	− 0.342*
Mum worked	0.158	0.177	− 0.130	− 0.098	0.049	0.068
Attitude at beginning of period	− 0.279**	− 0.292**	− 0.105	− 0.093	− 0.193*	− 0.176

*Significant at 5 % level; **significant at 1 % level.

the responses to the six items have been divided by their sample standard deviations. Since the SEM used probit regressions for the binary responses, the parameter estimates for the logistic regression modes are standardized by dividing them by the standard deviation of the standard logistic distribution: $\pi/\sqrt{3}$ (see Agresti, 2002, p. 125 for more details).

The conclusions from the significant parameter estimates for the SEM are very similar to those from the linear and logistic regression models. As noted when discussing Figure 22.2, the main difference is the lack of a significant selection effect of attitude in 1995 on beginning family care between 1995 and 1997. When the magnitudes of the significant standardized parameter estimates are compared for GCM and SEM, they are surprisingly similar given the difference in the approaches. The largest discrepancies are for the effects of beginning family care on family role attitude, where the estimates from the SEM are smaller in magnitude. To the extent that these differences are systematic, this could be a result of controlling for correlated measurement error in the responses to the six items used to measure the attitude.

22.7   CONCLUSIONS

Individual attitudes of women in our sample change over time. They appear to be adjusting their attitudes in response to life-course events, such as beginning family care, rather than the 'decisions' being influenced by their attitudes. In other words, in this particular sample, the adaptation effects are stronger than the selection effects. Our graphical chain model and our structural equation model gave very similar conclusions. However, we would not wish to overemphasize the generality of this point; with different data and different models, agreement between SEM and GCM should not always be anticipated. We have already noted a number of advantages and disadvantages with each approach. Graphical chain models allow the researcher to fit relatively simple models, often already familiar to the social science researcher, for various types of response (continuous, binary, ordinal, etc.). By using the Markov properties of chain graphs (e.g. Whittaker, 1990; Cox and Wermuth, 1996), we can draw conclusions about conditional independence and dependence structures of subsets of variables, and hence identify direct and indirect pathways through which explanatory variables are related to the outcome of interest. The SEM approach has the advantage of being able to decompose associations between variables into direct and indirect effects and also permits the inclusion of multiple indicator latent variables, which are useful for identifying and adjusting for measurement error. However, not all SEM software packages allow for complex survey designs and the inclusion of weights, though this is an area currently experiencing rapid development.

Of particular importance in the analysis of survey data are the ability to correct for measurement error (more readily achieved within the SEM approach) and the ability to correct for complex sample designs and to incorporate weights (more readily achieved within the GCM approach). Our comparison of parameter estimates from the GCM and SEM approaches suggests that failure to correct for measurement error in the gender role attitude may result in an overstatement of the selection effect of attitudes on labour force behaviour. At the same time our comparison of the estimates from the 'simple' approach and 'complex' approach, which takes account of stratification, clustering and weights, suggests that failure to take account of clustering will lead to an overestimation of precision, whilst the inclusion of weights affects both the magnitude of the point estimates

and their estimated variance. In this example the impact of weighting on our substantive conclusions is fairly minimal, perhaps because of the relative size of the weights. For this subgroup, the weights had a median of 0.85, a lower quartile of 0.68 and an upper quartile of 1.17. They range from 0.18 to 1.78 and hence do not reach the trimming level of 2.5 used in BHPS (Taylor *et al.*, 2007). More consideration of the need to adjust for weights, and their likely impact on the results, is required when they have a larger variance.

ACKNOWLEDGEMENTS

REFERENCES

Agresti, A. (2002). *Categorical Data Analysis* (2nd edn.). Hoboken, NJ: John Wiley & Sons, Inc.

Berrington, A., Hu, Y., Smith, P. W. F. and Sturgis, P. (2008). A graphical chain model for reciprocal relationships between women's gender role attitudes and labour force participation. *Journal of the Royal Statistical Society, Series A*, 171, 89–108.

Binder, D. A. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review*, 51, 279–292.

Bollen, K. (1989). *Structural Equations with Latent Variables*. New York: John Wiley & Sons, Inc.

Cox, D. R. and Wermuth, N. (1996). *Multivariate Dependencies*. London: Chapman & Hall.

Crompton, R. and Harris, F. (1998). Explaining women's employment patterns: 'orientations to work' revisited. *British Journal of Sociology*, 49, 118–136.

Evans, G. and Anderson, R. (2006). The political conditioning of economic perceptions. *Journal of Politics*, 68, 194–207.

Finkel, S. (1995). *Causal Analysis with Panel Data*. London: Sage.

Hakim, C. (1998). Developing a sociology for the twenty-first century: preference theory. *British Journal of Sociology*, 49, 137–143.

Lesthaeghe, R. and Moors, G. (2002). Life course transitions and value orientations: selection and adaptation. In R. Lesthaeghe (Ed.), *Meaning and Choice: Value orientations and life course transitions* (pp. 1–44). The Hague: NIDI-CBGS.

Magadi, M. Diamond, I., Madise, N. and Smith, P. W. F. (2004). Pathways of the determinants of unfavourable birth outcomes in Kenya. *Journal of Biosocial Science*, 36, 153–176.

Mohamed, W. N., Diamond, I. and Smith, P. W. F. (1998). The determinants of infant mortality in Malaysia: a graphical chain modeling approach. *Journal of the Royal Statistical Society, Series A*, 161, 349–366.

Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L. M. Collins and A. Sayer (Eds.), *New Methods for the Analysis of Change* (pp. 289–322) Washington, DC: American Psychological Association.

Muthén, L. K. and Muthén, B. O. (1998–2007). *Mplus User's Guide* (4th edn.). Los Angeles, CA: Muthén & Muthén.

Scientific Software International (2006). *LISREL, Version 8.8*. http://www.ssicentral.com/lisrel/index.html.

Skinner, C. J. (1989). Domain means, regression and multivariate analysis. In C. J. Skinner, D. Holt, and T. M. F. Smith, (Eds), *Analysis of Complex Surveys* (pp. 59–87). Chichester: John Wiley & Sons, Ltd.

Skinner, C. J., Holt, D. and Smith, T. M. F. (1989). *Analysis of Complex Surveys*. Chichester: John Wiley & Sons, Ltd.

Stapleton, L. M. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, 13, 28–58.

Stata Corp (2007). *Survey Data Reference Manual, Version 10*. College Station, TX: Stata Corporation.

Taylor, M. F., Brice, J., Buck, N. and Prentice-Lane, E. (2007). *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Chichester: John Wiley & Sons, Ltd.

Williams, L. J. and Podsakoff, P. M. (1989). Longitudinal field methods for studying reciprocal relationships in organizational behavior research: toward improved causal analysis. In B. Staw, and L. Cummings (Eds.), *Research in Organizational Behavior* (pp. 247–292). Greenwich: JAI Press Inc.