A Graphical System for Longitudinal Modeling using Dynamic Documents: Application to NLSY97 Religiosity Data

By

Andrey V. Koval

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

August, 2014

Nashville, Tennessee

Approved:

Joseph L. Rodgers, Ph.D.

James H. Steiger, Ph.D.

Kristopher J. Preacher, Ph.D.

Gary T. Henry, Ph.D.

DEDICATION

To

the Hauks,

the Snynders,

the Vaughns, and

the Kemps

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Joe Rodgers for making me work by letting me play. It is doubtful I would have made it any other way.

I would like to thank Will Beasley for introducing me to R and for coping with the aftermath. His friendship and generosity has been a constant source of inspiration to me.

Finally, I'd like to thank my wife Oleksandra Koval for her support and encouragement.

LIST OF TABLES

Chapter II

Chapter IV

LIST OF FIGURES

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

TABLE OF CONTENTS

ABSTRACT

This dissertation proposes a graphical analysis and presentation system for fitting, evaluating, and reporting longitudinal models in social sciences. The graphical innovations demonstrated here address practical issues that arise in evaluating *sequences* of statistical models. A progression of nested or otherwise related models in a sequence creates a context for model comparisons. The proposed graphical methods provide the researcher with visualization tools to facilitate model evaluation, using data mapping and interactive document design. The study applies these methods to examine empirical trends of religious involvement using a nationally representative household sample of American youth, the National Longitudinal Survey of Youth, 1997 (NLSY97). Annual measures in the NLSY97 from 2000 to 2011 provided panel data on church attendance from approximately 9,000 individuals born between 1980 and 1984. These data are examined using latent curve models (LCM) to study the nature of change in religious involvement between ages 13 and 31. Data, code, and reproducibility instructions for this study are published as a GitHub repository and are available to the research community.

CHAPTER I

INTRODUCTION

Overview

Statistical modeling has become an integral part of the scientific methodology in social and behavioral domains through methodological and technological developments of recent decades. Embedded in software technologies, statistical models have become a primary "window" into the world of abstract mathematical structures that are used to operationalize research theories. Social researchers sometimes liken research design and statistical models to such scientific tools as telescopes and microscopes – technologies that help them observe, examine, and ultimately explain phenomena behind human activity (Collins, 2006). Statistical models are not palpable, like microscopes, but certainly not less real or useful. Developing this analogy, the present work offers a "microscope" for statistical models: graphical methods for conducting comparisons of multiple related models, for helping the researcher to interpret the results of fitting the models, and for preparing the results of such analysis for publication in the spirit of reproducible research.

The purpose of these graphical methods is to facilitate evaluation and comparison of statistical models. A common analytic approach in analyzing longitudinal data is to fit a sequence of increasingly complex models (e.g., Singer & Willett, 2003). In practice, statistical modeling involves evaluating a series of model pairs, in which one model is somewhat different from the other. If the models are nested, each model comparison can be conceptualized as a null hypothesis significance test (NHST) that rules on the tradeoff between complexity differences between the models and differences in their performance in fitting empirical data (Rodgers,

2010). Examination of these comparisons informs and gives empirical grounds to the substantive theories developed to explain the phenomena behind the modeled data. Searching for the optimal model in a sequence of *competitors* during this process is associated with making a nontrivial number of model comparisons, each of which is potentially complex and messy.

The proposed graphical methods capitalize on the idea that statistical models "compete" during the estimation phase, but "collaborate" when interpreted. A better statistical fit might guide the researcher to the mathematical structure that reproduces the observed data patterns with the highest fidelity, but implications of such superiority would make sense only in comparison to other mathematical structures. As words help define other words, so models help define other models. Instead of arguing for the superiority of a single model selected as "the winner," the new graphical system proposed here directs the focus to telling a more complex story of the entire *sequence* of related models, thus making them collaborate in contextualizing the meaning of each other.

This dissertation will develop a new mechanism for comparing, interpreting, and reporting a series of latent curve models. Although the implementation of this graphical approach is developed for quantitative methodologists, by application the results can assist the methodologist in communicating modeling results to the wider research community. By providing clear pedagogical value, such graphical reports of model sequences are designed to facilitate understanding, interpretation, and communication of statistical models by *all* members of the research team, narrowing the gap between methodologists and applied researchers. One of the key problems addressed by the proposed methods is the information overload that often accompanies projects involving multiple models. Managing multiple specifications, parameters, indices, conditions, and constraints can frequently hide the forest behind the trees. That is, the limited resource of human attention is wasted on cognitive tasks that could be eliminated through intelligent report design. The reporting mechanism developed in this work gives the researcher the ability to effectively show how alternative models compare to some "winning" model, to demonstrate how "winning" is defined and justified, and to describe how model interpretation might change if more than one "winner" seems appropriate.

Foremost, this dissertation will emphasize the importance of *synthesizing* multiple statistical models into a coherent whole that offers something greater than the sum its parts. I will demonstrate how an interactive system of comparisons and contrasts can create a setting for contextualizing the behavior and interpretation of each model in a sequence. Although it remains a common practice to report only the results of the "winning" model, such proclivity is frequently explained by technical limitations and the cost of reporting the "failed" models, rather than sound methodological considerations. This dissertation develops a tool for synthesizing "winning," "competing," and even "losing" models into a richer understanding of the patterns involved in the competition. The graphical methods proposed here will assist the methodologist in producing more thorough, inclusive and informative model reports by offering a series of guides and templates to reduce the cost of similar report production. The intention of this study is to empower the practitioner to draw broader and more contextualized insights from their models, providing insight that may be difficult to achieve with standard approaches to model reporting.

Model Sequences

To illustrate the principles behind the proposed mechanism for model reporting, consider bottom-up and top-down model-building strategies, especially relevant in exploratory analyses. In the bottom-up (a.k.a. build-up or forward selection) approach, we start with a simplest possible model (1.1) and by adding terms incrementally (1.2, 1.3, and 1.4)[1], we arrive at some model specification that satisfies us with respect to both statistical fit and interpretational utility.

[1] Here I used Snijders & Bosker (2012) notation for multilevel models with *i* and *j* representing the first and second levels respectively. Later I will change it to *t* and *i*, to be consistent with Bollen & Curran (2006). For simplicity of illustration, these models were specified only partially: they represent only level-1 components, but ignore level-2 components and the covariance structure.

$$y_{ij} = \beta_{0j} + \varepsilon_{ij} \tag{1.1}$$

$$y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \varepsilon_{ij} \tag{1.2}$$

$$y_{ij} = \beta_{0j} + \beta_{1} X_{1ij} + \beta_{2j} X_{2ij} + \varepsilon_{ij} \tag{1.3}$$

$$y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \beta_{3j} X_{3ij} + \varepsilon_{ij} \tag{1.4}$$

where *i* and *j* are indices of the first and second level respectively, y_{ij} is the dependent variable, X_{1ij}, X_{2ij}, and X_{3ij} are independent variables, β_{0j}, β_{1j}, β_{2j}, and β_{3j} are estimated weights, and ε_{ij} is the residual. In a more complex model estimation setting, a typical "winning" model in HLM/MLM (Hierarchical Linear Modeling/ Multilevel modeling) might look like (1.5),

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \beta_{3j} X_{3ij} + \varepsilon_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01} W_{1j} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21} Z_{1j} + u_{2j} \\ \beta_{3j} &= \gamma_{30} \end{aligned} \tag{1.5}$$

where *i* and *j* denote levels in this mixed effects model, W_{1j} and Z_{1j} are second level predictors, and u_{0j} and u_{2j} are disturbances of the random effects.

The top-down (a.k.a. teardown or backward elimination) strategy reverses the logic and starts with the most complex model as reasonably possible inside the research agenda. For example, we might start with the following complex structure (1.6), and then seek ways to simplify the structure by removing elements that did not prove useful. One can imagine a series of steps that could reduce (1.6) into (1.5) or into many other possible simpler structures. Each modeling step (within which we remove or add an element or feature) can be formulated as a statistical test, the significance of which would advocate for support of the modification to the model proposed by the step. A carefully constructed sequence of model comparisons guides

researchers in formulating the conclusions of the analysis and informs substantive interpretations of data patterns. Clearly, reporting a sequence of models, as opposed to only reporting the "winning" one, is more informative and thorough.

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \varepsilon_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + \gamma_{03}W_{3j} + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}W_{1j} + \gamma_{12}W_{2j} + \gamma_{13}W_{3j} + u_{1j} \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \gamma_{23}Z_{3j} + u_{2j} \\
\beta_{3j} &= \gamma_{30} + \gamma_{31}Z_{1j} + \gamma_{32}Z_{2j} + \gamma_{33}Z_{3j} + u_{3j}
\end{aligned}
\qquad (1.6)
$$

However, we quickly run into a number of problems when working with sequences of models. When modifications are easy to track, as in (1.3) compared to (1.2), and the number of elements in the sequence is manageable (e.g. 1.1 – 1.4), performing model comparisons may be relatively straightforward. However, this process can get out of hand very fast as sequences become longer and include models that are more complex. For example, a sequence of models that reduces (1.6) into (1.5) might have a model pair (1.7) and (1.8):

(1.7)

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \varepsilon_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + \gamma_{03}W_{3j} + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}W_{1j} + \gamma_{12}W_{2j} + u_{1j} \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \gamma_{23}Z_{3j} + u_{2j} \\
\beta_{3j} &= \gamma_{30} + \gamma_{31}Z_{1j} + u_{3j}
\end{aligned}
$$

(1.8)

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \varepsilon_{ij} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}W_{1j} + \gamma_{12}W_{2j} + u_{1j} \\
\beta_{2j} &= \gamma_{20} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \gamma_{23}Z_{3j} + u_{2j} \\
\beta_{3j} &= \gamma_{30} + \gamma_{31}Z_{1j} + u_{3j}
\end{aligned}
$$

Comparison between these two models tests the usefulness of the term $\gamma_{03}W_{3j}$ in (1.7). It takes some time to study the models and identify the difference, but this inspection gives only the most basic information about the model. Each of these models would generate estimates, fit statistics, residuals, and other various quantitative output that describes an estimated model, to say nothing of the reproduced patterns of data the model recreates. Ideally, *all* of these results would need to be compared to fully understand the influence that the term being tested exerts

on the overall structure. Further, this is but a single comparison in the sequence that may count several or even several dozen competing/collaborating models.

Two interconnected challenges confront the modeler when working with model sequences: how to *represent* each of the models in a comparison and how to *construct* a sequence of models so that those comparisons are most meaningful and relevant to the research agenda. The graphical methods proposed here help the researcher address these challenges, but they cannot be discussed easily independent of data. To see these challenges illustrated with real data from NLSY97, consider the following brief example, which will be elaborated in the Methods and Results chapters.

An Applied Example

Consider a longitudinal multilevel model (1.10) in Snijders and Bosker (2012) notation, with predictors on both levels, located on the top right side of Figure 1.1, in which three time effects reproduce data trajectories over occasions *i* in individuals *j*. Intercept is modeled as random, while other time effects are modeled as fixed. Each time effect is regressed onto the age difference of the individual[2]. Let's say we would like to compare this model to its less restrictive counterpart (1.9). Identifying the difference is trivial: the cubic term $\gamma_{31}\boldsymbol{cohort}_j$ in (1.10) disappears in (1.9). The comparison between this pair of models corresponds to an NHST of the $\gamma_{31}\boldsymbol{cohort}_j$ prediction term in (1.10). When fitted, each of these models generates a collection of numeric descriptors, such as estimates and fit statistics, a partial list of which is given in Figure 1.1. Here, I used the lme4 *R* package for estimation, but one can imagine similar outputs from software like *Mplus*, SAS, SPSS and others.

[2] Only general familiarity with this model is required for present purposes, for detailed specification and estimation report of this model the reader is directed to methods (III) and results (IV) chapters of this thesis, respectively.

(1.9)

$$y_{ij} = \beta_{0j} + \beta_{1j} timec_{ij} + \beta_{2j} timec_{ij}^{\ 2} + \beta_{3j} timec_{ij}^{\ 3} + \varepsilon_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01} cohort_j + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} cohort_j$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21} cohort_j$$
$$\beta_{3j} = \gamma_{30}$$

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: attend ~ 1 + timec + timec2 + timec3
+ cohort + cohort:timec + cohort:timec2 + (1 | id)

     AIC      BIC   logLik deviance df.resid
103902.3 103977.4 -51942.1 103884.3    31059

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.0437 -0.5070 -0.0690  0.3967  5.6120

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 2.511    1.585
 Residual             1.269    1.127
Number of obs: 31068, groups: id, 2589

Fixed effects:
              Estimate Std. Error t value
(Intercept)    2.8647540  0.0644133   44.47
timec         -0.1148575  0.0189983   -6.05
timec2         0.0200345  0.0035672    5.62
timec3        -0.0010118  0.0002057   -4.92
cohort         0.2383658  0.0252356    9.45
timec:cohort  -0.0613956  0.0049632  -12.37
timec2:cohort  0.0033849  0.0004347    7.79

Correlation of Fixed Effects:
            (Intr) timec  timec2 timec3 cohort tmc:ch
timec       -0.367
timec2       0.246 -0.911
timec3      -0.158  0.753 -0.952
cohort      -0.823  0.209 -0.081  0.000
timec:cohrt  0.314 -0.549  0.247  0.000 -0.382
timec2:chrt -0.260  0.529 -0.256  0.000  0.316 -0.964
```

(1.10)

$$y_{ij} = \beta_{0j} + \beta_{1j} timec_{ij} + \beta_{2j} timec_{ij}^{\ 2} + \beta_{3j} timec_{ij}^{\ 3} + \varepsilon_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01} cohort_j + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} cohort_j$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21} cohort_j$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31} cohort_j$$

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: attend ~ 1 + timec + timec2 + timec3
+ cohort + cohort:timec + cohort:timec2 + cohort:timec3 + (1 | id)

     AIC      BIC   logLik deviance df.resid
103901.3 103984.8 -51940.7 103881.3    31058

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.0440 -0.5025 -0.0732  0.3951  5.5889

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 2.511    1.585
 Residual             1.269    1.127
Number of obs: 31068, groups: id, 2589

Fixed effects:
              Estimate Std. Error t value
(Intercept)    2.8383501  0.0662133   42.87
timec         -0.0777854  0.0287157   -2.71
timec2         0.0112332  0.0062337    1.80
timec3        -0.0004784  0.0003719   -1.29
cohort         0.2509412  0.0262713    9.55
timec:cohort  -0.0790519  0.0113935   -6.94
timec2:cohort  0.0075767  0.0024733    3.06
timec3:cohort -0.0002540  0.0001476   -1.72

Correlation of Fixed Effects:
            (Intr) timec  timec2 timec3 cohort tmc:ch tmc2:c
timec       -0.410
timec2       0.327 -0.960
timec3      -0.278  0.900 -0.984
cohort      -0.833  0.342 -0.272  0.232
timec:cohrt  0.342 -0.833  0.800 -0.750 -0.410
timec2:chrt -0.272  0.800 -0.833  0.820  0.327 -0.960
timec3:chrt  0.232 -0.750  0.820 -0.833 -0.278  0.900 -0.984
```

Figure 1.1 Partial Estimation results for two models: (1.9) - left and (1.10) - right

To determine a "better" model in this or any other pair we may refer to a sequence of formal (and informal) statistical comparisons. For example, nested models (as in this case) could be compared using a variety of criteria such as deviance, AIC, AICC, or BIC, to name a few.

$$D = -2\ell$$
$$AIC = D + 2q$$
$$AICC = D + 2q\frac{N}{N-q-1}$$
$$BIC = D + q\ln N$$

$\ell = \text{logLikelihood}$

$D = \text{Deviance}$

$q = \text{number of estimated parameters}$

$N = \text{total sample size}$

Formally, only deviance is subjected to a direct statistical test, such as the chi-square difference test $(D_0 - D_1) \sim \chi^2(\mathrm{df} = q_1 - q_0)$, which is frequently used as a starting point in a sequence of model comparisons. This test is given as a t-statistic (treated as a *z*-value in interpretations) of cohort in (1.10): parameter estimate = 0.2509412, standard error = .0262713, t-value = 9.55.

Other indices (such as BIC, AIC) have meaning only in comparison with rival models[3]. In the context of model comparison, the model with the lower AIC or BIC is better fitting, and the value of deviance indicates the total unadjusted misfit computed from the likelihood function. Model (1.10) outperforms model (1.9) in terms of absolute fit ($D_{(1.9)}$ =103,884.3> $D_{(1.10)}$ =103,881.3), as would be expected from a more complex model (i.e., a model with more parameters). However, after adjusting for parsimony ($AIC_{(1.9)}$= 103,902.3 > $AIC_{(1.10)}$ =103,901.3) the model (1.9) seems to be a more reasonable choice, but not when model performance is adjusted for sample size ($BIC_{(1.9)}$= 103,977.4 < $BIC_{(1.10)}$ =103,984.8). A significant *t* in the formal test (*t* = 9.55) justifies the increase of model complexity involved in adding $\gamma_{31}cohort_j$ to (1.9). However, it is important to note that the evaluation of differences in relative information criteria can be informed by the performance of other models in the sequence; models help us define and interpret other models. For example, knowing how much AIC/BIC changed when the term $\gamma_{21}cohort_j$ is removed from (1.9) would contextualize the meaning of the difference between AIC/BIC in comparing (1.9) and (1.10).

Fit and information criteria, however, only describe how well a model *does* something (predicts values) per unit of complexity (*df*); for what a model actually *is*, we must refer to the estimated parameters, predicted values, residuals, and other indices. To inspect how adding $\gamma_{31}cohort_j$ to (1.9) disturbs the values of the estimated effects, their precision, and covariations one would have to compare the values for the corresponding estimates:

[3] Many information criteria and fit indices have been developed: GFI, AGFI, non-normed index Delta2 (Bollen, 1989), normed index Rho1(Bollen, 1986), NFI (Bentler & Bonett, 1980) , CFI (Bentler, 1990), RNI (McDonald & Marsh, 1990), RMSEA (Steiger & Lind, 1980), each placing its own emphasis in the definition of the "best" model. Depending on model type, data, and research agenda at hand researchers may need to choose specific indices, however most software systems report at least deviance, AIC, and BIC. The present thesis uses these three quantifications of model performance.

$Intercept_{(1.9)} = 2.8647540 > Intercept_{(1.10)} = 2.8383501$
$timec_{(1.9)} = -0.1148575 < timec_{(1.10)} -0.0777854$
$timec2_{(1.9)} = 0.0200345 > timec2_{(1.10)} = 0.0112332$
...

and so on, until the desired list of value comparisons is exhausted. Evaluating differences in the values of the descriptors in model pairs is an arduous, sequential task. The graphical methods introduced here expand this operation from one involving only minute inspection of the outputs of model estimation, to one including a visual processing exercise.

Graphical Methods for Model Sequences

Figure 1.2 is a linked screenshot of a prototypical model sequencer, where model specification (partial) of (1.10)[4] is given along with the selected estimation output, graphs of predicted individual trajectories (thin red lines) in the bottom left, and a graph of model performance indices in the bottom right. Clicking the link "m6R1" in the guide menu on the left margin of the screen switches the view to the report of the corresponding model, specified by (1.9). This interactive report will be discussed and illustrated in detail by using the NLSY97 data in the Results chapter. For simplicity of the present demonstration, I have selected the point estimates of the standard deviation of the residual, and the fixed effects and the standardized covariance matrix for the random effects. The graph in the bottom right shows raw deviance of all models in the sequence, highlighting the model m10 (known as m7R1 in the model span), currently "under the microscope."

[4] The models used in the study were given descriptive names: (1.10) is referred as m7R1, while (1.9) corresponds to model m6R1 in the model span. The Results chapter will elaborate on the convention for model names.

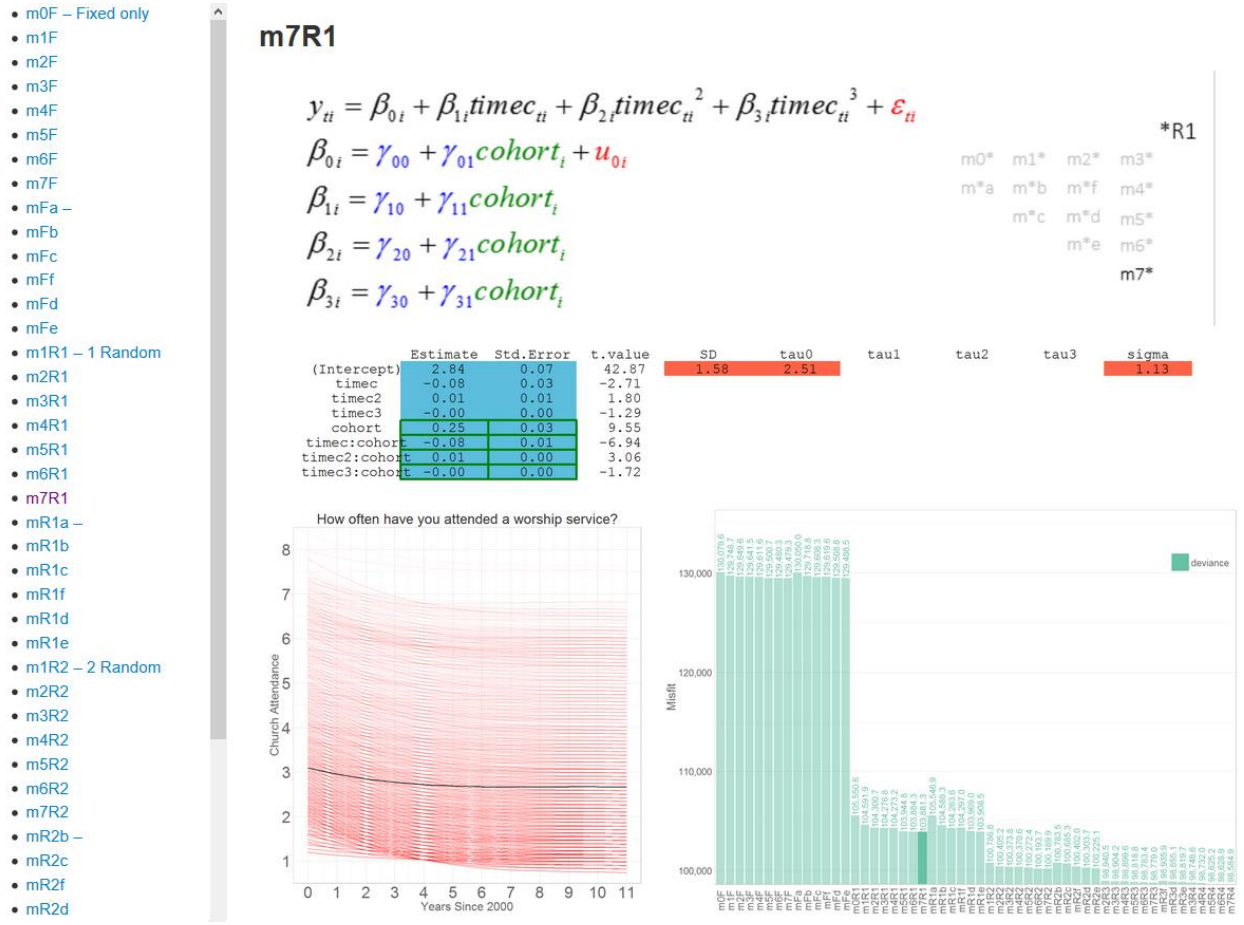


Figure 1.2 Screen shot of a prototypical LCM model sequencer

Clicking between m7R1 and m6R1 (one can use hot-keys “Alt + Left/Right Arrow” for smoother transition) we immediately make several useful observations. First, by switching the views we can immediately identify the distinction between the models, allowing our eye to be drawn to the movement on the screen as added elements appear in the specification. Second, we notice that estimates for the fixed effects do change between the models, though not very much. Third, by studying the graph of the fit indices, the researcher can give relative meaning to the model performance indices. We see that although m7R1 improves on m6R1, this improvement is relatively small compared to changes from m4R1 to m5R1, yet rather substantial compared to other pairs of models. Finally, one can see how predicted individual trajectories (bottom left, thin red lines) change with the introduction of the extra predictor: as expected

from the minuscule point estimate, the change in the shape of trajectories is barely noticeable, compared to other model pairs.

This demonstrated graphical method of model comparison conveyed a lot of relevant information about latent curve models very quickly. Traversing the constructed sequence gives us the vocabulary to describe the latent construct of the study: in this case, the shape of change of religious attendance. Naturally, the present example can be extended to include other results of model estimation interesting to the researcher, for example, correlations and standard errors of the fixed effects (see the output in figure 1.1). Other statistical models (mixtures, hazard, etc.) would call for different ensembles of model manifestations to optimally represent complexity and different types of modeling steps from which to construct meaningful sequences. In general, the graphical methods for sequencing longitudinal models will look different from one statistical method to another, but will be united by three design principles:

1) Ensembles of model manifestations must fit onto the same surface area across models
2) Only differences between models should be noticeable during alternation of two views
3) Choosing models for viewing must occur via interactions with page elements

As mentioned earlier, two main challenges arise from working with model sequences: how to *represent* each of the models in a comparison and how to *construct* a sequence. The present dissertation offers possible solutions to these challenges for LCM, leaving mixture growth, Markov, and other models for future work. The design principles listed above address these challenges, making a falsifiable statement: "Implementation of these principles in reports of statistical models offers analytical opportunities superior to those of traditional methods of analyzing sequences of longitudinal models". The rest of this dissertation provides an evaluation of this claim.

Organization and Chapter Summary

Recent reviews of quantitative methodologies for longitudinal data (e.g. Collins, 2006) point to an emerging challenge associated with the increasing complexity of statistical models. As models become more powerful and nuanced, they naturally grow more difficult to

understand, operate, interpret, and discuss. The challenge intensifies when entire sequences of models are estimated and compared. The present dissertation offers graphically-oriented methods to structure and analyze Latent Curve Models (LCM) .

Chapter II, Literature review, helps us recognize the general trend of increasing complexity in modern modeling methods. There, I expose the problem that my graphical methods address. After reviewing trends in statistical modeling in general, I focus on *longitudinal* models in the social and behavioral sciences, giving a brief overview of the methodological field for the last few decades, from which statistical methods were selected for the present study. Chapter II concludes with a brief overview of past published research articles and statistical analyses of religiosity, to provide context and rationale for the empirical research reported here.

Chapter III, Methods, gives a detailed description of the sample, data, and methodology used in the analyses. All analyses and visualizations in this dissertation can be reproduced with publically available code and templates; therefore, special attention was given to preparing the reader for reproducing these methods with their own data. First, I describe the NLSY97 sample, its data structure, and temporal design. Then, variables selected for analysis from the NLSY97 database are discussed with the help of a computerized databox, following Cattell (1952). Transformations of the clean data in preparation for modeling the focal variable (church attendance) are described. Finally, I specify LCM in its general form.

Chapter IV, Results, reports a sequence of latent curve models fit to the data of church attendance in the NLSY97 sample. There, I describe and illustrate the report mechanism for presenting a sequence of statistical models compactly and efficiently. I demonstrate how models compete in determining the "best" model and how they collaborate to arrive at meaningful substantive interpretations of data structures and predicted values.

Chapter V, Discussion, reviews the effectiveness of the proposed graphical method, discusses the way the graphical results informed the analysis of the NLSY97 religiosity data, and draws some general conclusions for substantive research on religiosity. Next, the value of reproducible research and dynamic reporting of the model sequences is discussed. I conclude with discussion of limitations and ideas for future research.

CHAPTER II

LITERATURE REVIEW

The Challenge of Model Complexity

Complexity on the rise

Quantitative methodology offers a certain kind of active and interesting challenge. In the last 50 years, the variety and amount of data being collected on human-related activity have been accelerating. Technology has made it easier to collect, process, and share data. A remarkable variety of methodological approaches has been developed to accommodate new types and amounts of data. Some of those approaches are models, which have become numerous, specialized, and complicated. Consequently, human limitations in attention, perception, and information processing have become relevant in working with models.

Contributing to the challenge is the particular difficulty associated with the measurement of the primary subject matter of social science, behavior. A number of years ago, in a discussion concerning the role of methodology in the future of psychology, Raymond Cattell (1988, p. 5) noted that "[in order] to overcome the difficulties due to unusually complex subject matter, it is now necessary for psychologists to become unusually explicit and sophisticated about methodology." Indeed, since then psychologists have become substantially more sophisticated with their methodology than they were in 1988, to the degree that in some (perhaps many) cases it is now prohibitive, if not impossible, for non-experts to appreciate, much less to employ, the methodological fruits of their labors.

Longitudinal models are at the forefront of this challenge because they operate on a very general (and important) data structure in social research. Any statistical model in psychology can be thought of a special case of its longitudinal extension. Considering the pivotal role

longitudinal designs play in establishing causality (Pearl, 2000; Rubin, 1974; Shadish, Cook, & Campbell, 2002), it is understandable why addressing methodological issues in developmental models subsumes a great variety of other analytical instruments. This section will discuss the challenge of methodological complexity in the social sciences. In later chapters, I will demonstrate some new graphical methods for model comparison that show promise in working with some specific longitudinal models. I apply recent technological advances to produce a visual integration of various model manifestations. This allows for quicker and easier evaluation and management of statistical models. To guide the development of such technological innovation I define three directions in which statistical models in general, and longitudinal models specifically, have been evolving.

Types of complexity

First, models have become more *numerous* (though in some cases only by appearance). Many authors point out the wide range of options in analytical strategies available to developmental researchers (Collins, 2006; Cudeck & Harring, 2007; McArdle, 2005). Understandably, a wide variety of tools to answer a broad spectrum of questions can create either clarity or confusion. Card and Little (2007, p. 207) noted in the introduction to the special issue of *International Journal of Behavioral Development* on longitudinal modeling: *"*Given [this] tremendous amount of literature on longitudinal data analysis, the problem for developmental researchers is not a lack of information but rather an over-abundance of information." Increasing specialization of models has produced an expansive and elaborate vocabulary, often laden with historical and/or disciplinary baggage. The disparate terminology that has been emerging, instead of clarifying the distinctions among the models, arguably contributes to the confusion instead of clarifying it. For example, Raudenbush (2001b) lists model names that can be used to discuss the same approximate concept from different perspectives : "covariance components models," "hierarchical models," "latent curve analysis," "latent growth models," "mixed models," "mixed linear models," "multilevel models," "multilevel linear models," "random effects models," "random coefficient models," and "structural equation modeling." Gibbons, Hedeker, and DuToit (2010) give additional approximate synonyms, adding "two-stage models," "empirical Bayes models," and "random regression". Frequently, models with different names,

forms, and notations reveal themselves under scrutiny to be mathematically equivalent. Raudenbush (2001a) pointed out that disparate terminology can frequently be traced to software, rather than conceptual differences.

Second, models have become more *complex.* Here, a literal meaning of "complexity" is invoked: the number of elements of which the whole is comprised. Encouraged by the ease and affordability of estimation, even cross-sectional models in econometrics, for example, may employ hundreds of predictors in a single regression equation. A typical model in psychometrics, to take a less severe example, might incorporate a few dozen variables, although psychometric models defined at the item level may also include hundreds of elements. Of course, once placed in a longitudinal setting the number of elements is multiplied by at least the number of time points. Larger numbers of variables are not only a chore to handle during estimation, but also a challenge for interpretation. Interpretations are supposed to simplify the precision of the mathematical structure into patterns understandable by human language. Curran, Obeidat, and Losardo (2010) articulated: *"*And as any developmental researcher can attest, statistical models for longitudinal data can become exceedingly complex exceedingly quickly, both in terms of fitting models to data and properly interpreting results with respect to theory*"* (p. 122).

Third, models have become exceedingly *sophisticated* (in a manner that is distinct from complexity). Here, model sophistication implies some system of constraints by which components are united into mechanisms for generating predictions. Multiple components of a very complex model can nevertheless be united in a structurally straightforward fashion, as, for example, in a multiple regression with a large number of predictors. Others, like growth mixture models and dynamic item response theory, for example, offer more intricate ways of combining equation elements.

The concern about increasing inaccessibility of modern modeling methods by the broader community of social scientists is rising in both methodological and applied areas. The mathematical and programming expertise required to build, estimate, and interpret statistical models frequently becomes an obstacle for applied researchers, whose data analytical skills understandably lag behind those of professional methodologists. In addition to technical

expertise, researchers must possess experience and be willing to invest a considerable mental effort to make sense of models and exploit their inferential potential fully. The more sophisticated a model is, the more difficult it is to specify, estimate, and communicate its findings. Thus, model complexity and sophistication are at constant odds with the ease and transparency of inference: *"*The challenge we face is that we must carefully balance the complexity of our theoretical models with the requisite complexity demanded by the empirical evaluation of our theory*"* (Curran & Willoughby, 2003, p. 581).

Marketing offers a useful analogy to modern quantitative methodologists: the success of a "product" (i.e. methods, tests, software) depends not only on its utility, but also to a significant degree on the ease with which it can be used by the wide public. A microwave that requires a Ph.D. in mechanical engineering to operate will not sell well despite its out-of-this-world performance. Within our social/behavioral science research domain, we want the model to be complex enough to accommodate the research agenda, but simple enough to be attractive to and usable by practicing researchers. Theoretical developments in methodology, changing data culture, and evolving demands of substantive social research push for a larger number of models that are more complex and more sophisticated. These trends expose an important vulnerability of such confluence: models become exceedingly difficult to operate.

How can this challenge of methodological complexity be resolved, managed, or at least addressed? Some solutions are emergent: *taxonomic* devices and conceptual *frameworks* organize the methodological field into convenient clusters; I give a brief overview of them in the next section. Then, I develop and demonstrate a visual system of information management, designed to address the challenges in modern modeling methods for longitudinal data, using LCM as an example. Collins (2006, p. 508) drew an analogy: *"*In the natural sciences, the investigator may choose an instrument, such as a microscope, to provide a view of the phenomenon of interest. In the social and behavioral sciences, research design is a similarly important instrument that provides a view of the change phenomenon of interest.*"* In certain ways, graphical methods are even better analogies to microscopes than are research design principles and statistical models, because they are directly visual. The present work develops

and presents a graphical microscope for examination of statistical models, the data they reconstruct, and the relationship between the two.

The microscope analogy is a rough one; after all, modern microscopes not only magnify, but also equip researchers with additional capabilities, including scales to evaluate size of the examined objects, controllers of zoom and spectrum of light, and even cameras to take stills and videos. More elaborate microscopes (e.g. fMRI, NIRS, astronomic spectroscopes) demonstrate that it is not sufficient to merely perceive the patterns in order to be well equipped to interpret them. Enhancing the tools that enhance the senses is frequently necessary to uncover the patterns and the meaning behind them. In working with complex abstractions, we need the ability to enhance our perception of them.

Unlike previous (and overlapping) developments of interactive data visualization tools (DataDesk (Velleman, 1989), ViSta (Young & Bann, 1996), and Mondrian (Theus, 2003)) the present study does not offer a tool for *conducting* data analysis, which is left for specialized software. Instead, the focus is on expanding interpretation by exploring the *sequences* into which they can be organized. The methodological scope of the current work is limited to latent curve models, offering a proof of concept, which can be extended to other statistical methods in the future. This choice was informed by a review of relevant methodological literature in social sciences for the past several decades. The next section reviews several ways to organize the rapidly evolving field of longitudinal modeling.

Review of Longitudinal Methods

A brief overview of the main literature on longitudinal modeling in social sciences revealed a variety of ways to organize statistical tools that model change. For historical accounts of longitudinal models, the reader is directed to Bollen (2007), Bollen & Curran (2006), and Fitzmaurice & Molenberghs (2009). Representative examples of structuring the field of longitudinal modeling can be found in recent review articles (Card & Little, 2007; Gibbons et al., 2010; Hertzog & Nesselroade, 2003; McArdle, 2009). Celebrating the abundance of recently developed statistical models of change, Collins (2006, p. 509) remarked: “With an unprecedented array of statistical models from which to choose, today’s behavioral scientist has

an excellent chance of identifying and applying a statistical model well suited to the theoretical model of interest" (p 509). However, as was discussed in the previous section, handling this "unprecedented array" is becoming a challenge.

Despite the variety of nuances, several ways to distinguish longitudinal models were especially prevalent. Certain features, such as the scale of latent variables, the scale of observed variables, and scale of change itself help think through the selection of possible models for operationalizing theoretical models of change. This section reviews several taxonomies that help motivate the use of latent curve methods I present in this paper and sets up the stage for extending the graphical methods to other family of models.

The first taxonomy that will be reviewed comes from Little, Preacher, Selig, and Card (2007). They placed a ubiquitous taxonomic device, type of data, at the pivot of organizing general SEM models (Figure 2.1), distinguishing between the scale of *latent* and *observed* variables in the model. Their table illustrates how the options for analytic strategy changes with reconceptualization of the latent trait or with the transformation of the data that enters the statistical model. Such decision can be made during both the design and/or the analysis phases of a research project. The bottom-right quadrant hosts probably the most populous category: considered as reformulations of each other (Curran, 2003), latent curve and random coefficient models are considered to be *"*currently the most widely used longitudinal data analysis technique in psychology*"* (Kuljanin, Braun, & DeShon, 2011, p. 1).

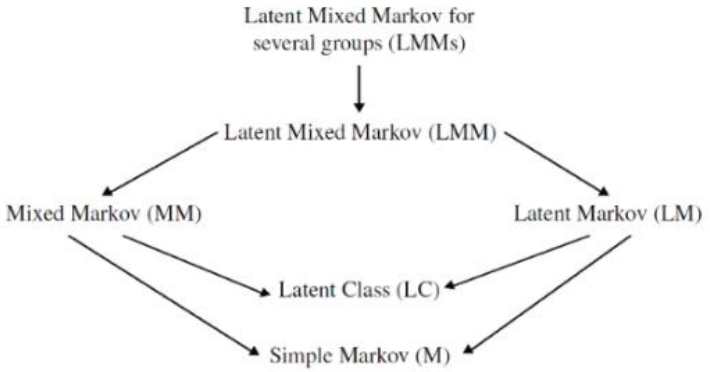
The taxonomy offered by Kaplan (2008), focused on the case when both latent and manifest variable are categorical, portrayed in the top-left quadrant of Little et al. (2007). Elaborating on works of Langeheine (1994; Langeheine & Van de Pol, 2002) the taxonomy in Figure 2.2 shows how various statistical models of stage-sequential change can be represented with a diagram, in which arrows indicate that the model at its end is a special case of the models at its origin. The most general model here, mixture latent Markov, when applied to studying continuous growth is known as the general mixture model (Muthén , 2004). This observation helps with describing the relatedness of GMM and Markov/EMOSA models when the future research is discussed in Chapter V.

A Simple Taxonomy of the Nature of the Measured Variables Crossed with the Nature of the Latent Variables Found in General SEM Models

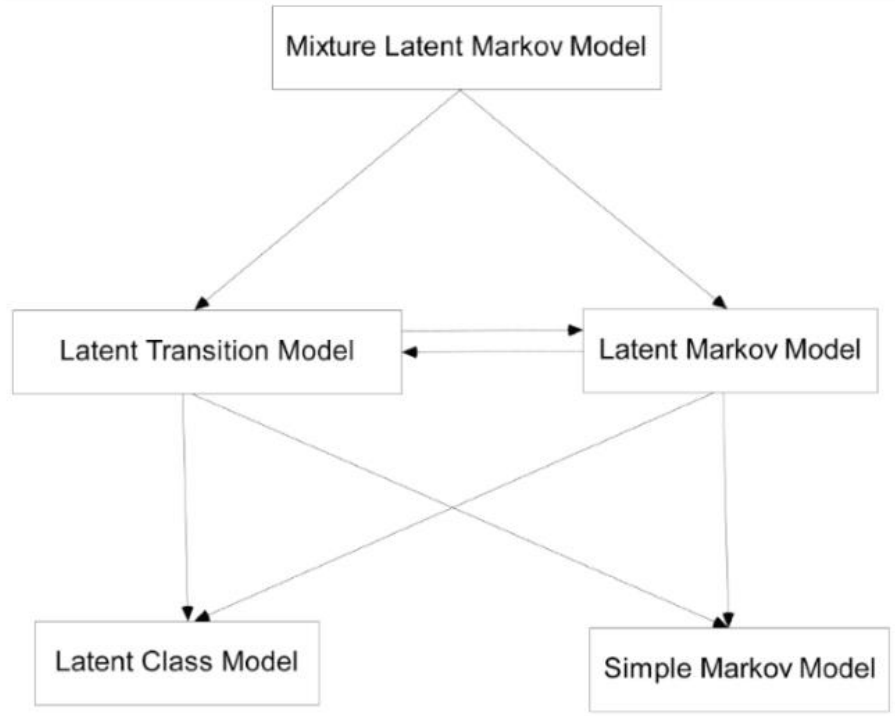
		Nature of the Unobserved, Latent Variables	
		Categorical	**Metrical**
Nature of the Observed, Manifest Variables	**Categorical**	Latent Transition Analysis	Latent Trait Analysis; IRT
	Metrical	Latent Class; Mixture Modeling	Latent Variable Analysis; CFA

Note: *Categorical* refers to variables that reflect nominal or polytomous categories. *Metrical* refers to variables that reflect ordinal, interval, or ratio-level properties. Many models for longitudinal data can contain more than one of these kinds of variables.

Figure 2.1 Screen shot from Little et al. (2007)



A hierarchy of Markov chain models (adapted from Langeheine, 1994).



A modified hierarchy of Markov models (adapted from Langeheine & Van de Pol, 2002). Arrows represent special cases.

Figure 2.2 Screen shot from Langeheine & van del Pol (2005) - top, and Kaplan (2008) - bottom. Arrows represent special cases

Muthén and Curran (1997) demonstrated how the history of academic disciplines shaped the development and classification of longitudinal models. They distinguished three academic traditions of particular importance: biostatistics, education, and psychometrics. Each of the disciplines framed the questions in the language relevant to its own agenda. Not surprisingly, the models that provided answers to these questions had to reflect the idiosyncrasies of the respective discipline. Key terms, references, and software are organized in Table 2.1.

Table 2.1 Created from classification of longitudinal models by Muthén & Curran (1997)

	Biostatistics	Education	Psychometrics
Terms	Repeated measurement Random-effects ANOVA Mixed model Random coefficient modeling.	Slopes-as-outcomes Multilevel modeling Hierarchical linear modeling	Latent curve analysis Latent variable structural equation modeling.
References	Rao (1958) Laird and Ware (1982) Diggle, Liang, and Zeger (1994)	Cronbach (1976) Burstein (1980) Goldstein (1987) Bock (1989) Bryk and Raudenbush (1992) Longford (1993)	Tucker (1958) Meredith and Tisak (1990) McArdle and Epstein (1987)
Software	BMDP5V SAS PROC MIXED, MIXED, and MIXOR.	MLn HLM VARCL	Amos, CALIS EQS, LISCOMP L1SREL, MECOSA MX

The last taxonomy in this overview comes from Collins (2006), who surveyed representative longitudinal models, organizing them with respect to the *scale of the outcome* and the *temporal design* of the study. While the Little et al. (2007) taxonomy used the scales of manifest and latent variables, Collins considered how the scale of the time itself influences the choice and/or development of the statistical model. In particular, she distinguished between two general types of longitudinal data: panel (4-8 time points) and intensive (20 and more time points). She also articulated the principle that is becoming popular in modern methodology: a good longitudinal research design seamlessly integrates a theoretical model of change, temporal design, and a statistical model of change. The theoretical model describes the nature of change in the modeled phenomenon, discussing such aspects as shape, periodicity, and the scale of change, as well as the nature and role of covariates. The temporal design structures observations in time, describing timing, frequency, and spacing of measurement points. The statistical model tests a specific mathematical operationalization of the theoretical model against

the observed structures of data. The models that Collins (2006) chose to exemplify her taxonomical categories are organized in Table 2.2.

A simple longitudinal model may have different "maps" of how it can be extended, depending on what assumptions we are willing to make or what questions are driven to answer. For example, "slope-as-outcome" model, random effects ANOVA, or unconditional growth model may offer different potential for extensions, despite being very similar. With this in mind, the reviewed taxonomies should not be approached as ontological statements of "what is" in the field of quantitative methodology, but rather as roadmaps to remind the researchers what their model *can become* under specific conditions. The statistical model that I chose to illustrate my sequence reporting technique maps well into the reviewed taxonomies: One can easily locate and contextualize LCM in each of them. Such versatility of relatedness offers a hope that my graphical methods can be extended to other related models as well. This subsection has reviewed taxonomies of statistical modeling methods. The next subsection offers a brief review of statistical models of religiosity.

Table 2.2 Created from the survey of longitudinal models by Collins (2006)

Theoretical Model	Scale of Outcome and Time Calendar effect Periods/cycles Shape of change Time variant covariates Time invariant covariates	
Temporal Design	Timing Frequency Spacing	
Statistical Model		
	Change in Continuous Variable	Movement between states
Panel Design	**MLM/HLM & SEM/LCM** Raudenbush (2001a) McArdle and Epstein (1987) Meredith and Tisak (1990) **PIECEWISE & MULTIPHASE** Cumsille, Sayer, and Graham (2000) Cudeck and Klebe (2002) **AUTOREGRESSIVE & HYBRID** (Bollen & Curran, 2004) McArdle and Hamagami (2001) **Growth Mixture Models** (D. Nagin and Nagin (2005); Daniel S Nagin (1999)) D.S. Nagin and Tremblay (2001) Muthén and Muthén (2000) Muthén (2001)	**DISCRETE-TIME SURVIVAL ANALYSIS** D. R. Cox (1972) Singer and Willett (2003) Singer and Willett (2003) **LATENT TRANSITION ANALYSIS** Langeheine (1994) Lanza, Flaherty, and Collins (2003) Lanza, Collins, Schafer, and Flaherty (2005) Lanza and Collins (2002)
Accelerated Panel Design	Bell (1953) McArdle and Hamagami (2001) Duncan, Duncan, and Hops (1996) Miyazaki and Raudenbush (2000)	
Intensive Longitudinal Design	**FUNCTIONAL DATA ANALYSIS** Fan and Gijbels (1996) Ching, Fok, and Ramsay (2006) Li, Root, and Shiffman (2006) **DYNAMICAL SYSTEMS** Boker and Graham (1998) Boker and Nesselroade (2002) Ramsay (2006)	**POINT-PROCESS MODELS** D. Cox and Lewis (1966) Cressie P. J. Diggle and Diggle (1983) Lewis (1972) Rathbun, Shiffman, and Gwaltney (2006)

Modeling Religiosity

The literature in psychology and sociology links adult and adolescent religiosity to positive and negative behaviors and outcomes. Studies abound exploring the association of religiosity with substance use (Mason & Spoth, 2011; Sanchez, Opaleye, Chaves, Noto, & Nappo, 2011; Vaughan, de Dios, Steinfeldt, & Kratz, 2011) , sexual behavior (Rostosky, Wilcox, Wright, & Randall, 2004), gambling (Casey et al., 2011), delinquency (Desmond, Soper, & Kraus, 2011), depression treatment (Schettino et al., 2011), community service (Smith, 2003), identity formation (Puffer et al., 2008), educational outcomes (Hakin Orman, North, & Gwin, 2009), coping (Desrosiers & Miller, 2007), and marital satisfaction (MacArthur, 2008; Orathinkal & Vansteenwegen, 2006), to name just a few of the most recent works. For meta-analysis on the role of religiosity and positive and negative behavioral outcomes see Cheung and Yeung (2011). However, in most cases such studies focus on religiosity as a predictor or explanatory factor for other behaviors of interest, rather than developing models of religiosity itself.

In particular, the change in religiosity during the transition from adolescence into adulthood has only occasionally been treated within developmental psychology until recently (King & Boyatzis, 2004). The stage of life between 18 and 25 years of age, identified by Arnett (2000) as "emerging adulthood," is associated with substantial dynamics in identity formation (Nelson & Barry, 2005), neurological and cognitive development (Steinberg, 2005), as well as transformation of the social environment. The amount and multidimensionality of change experienced by individuals in this period clearly calls for longitudinal modeling, with only a few examples in the literature (Desmond et al., 2011; Petts, 2009; Uecker, Regnerus, & Vaaler, 2007). Otherwise, most studies addressing religiosity of adolescents and emerging adults were either purely cross-sectional, or contained but a few waves of observations, or used small, nonrepresentative samples. In addition to these methodological shortcomings, as Desmond et al. (2011) noted, there exists *"*the lack of strong developmental studies that examine how adolescents' religious attitudes and behaviors grow or decline over time.*"* The empirical portion of the present study helps to fill this gap by analyzing religious attendance of a nationally representative sample of American households (the NLSY97) in longitudinal detail, modeling twelve rounds of panel data.

Most research on acquisition, transmission, and change of religious beliefs and practices operates in an ecological framework, identifying relevant *socializing agents.* The influence of socialization in transmission and development of religiosity among adolescents and emerging adults is well recognized (Hill, 2011; M. D. Regnerus, Smith, & Smith, 2004; Vaidyanathan, 2011). Among the agents of socialization, two classes are most apparent: *familial* and *extra-familial.* The role of *parents* (Day et al., 2009; Milevsky, Szuchman, & Milevsky, 2008), mothers (Hood Jr, Hill, & Spilka, 2009), and fathers (Wilcox, 2002) in transmission of religious beliefs are linked to both formation of religious identities in childhood and religious practices in young adulthood. The role of *siblings* as socialization agents, however, is yet to be explored (McNamara Barry, Nelson, Davarya, & Urry, 2010). The models that look at transition of religiosity between generations include Myers' interactive model of religious inheritance (Myers, 1996), the intergenerational transmission model (Bengtson, Copen, Putney, & Silverstein, 2009), and a broader model of religious socialization (Martin, White, & Perlman, 2003). Inheritance models have been enriched by studies adopting an evolutionary perspective (Weeden, Cohen, & Kenrick, 2008), elaborating on the role of gene-environment interaction in the family context on formation of religious behavior and mate-selection mechanisms that increase the prominence of religious practices in the population (Rowthorn, 2011). Among socializing agents outside of the family, researchers have studied other adults: *mentors* in college (Cannister, 1999), *peers* (Gunnoe & Moore, 2002; Schwartz, 2006), and *media* (Clark, 2002; Pardun & McKee, 1995), an influence that Arnett (1995) suggested is a type of "self-socializing" influence. For a broad discussion of themes in adolescent religiosity, the reader is referred to a special issue of *Applied Developmental Science* (Volume 8, 2004), and the latest journal-article review of the field (McNamara Barry et al., 2010).

It is well documented that religious involvement declines during transition from adolescence into young adulthood (M. Regnerus, Smith, & Fritsch, 2003; Smith & Snell, 2009). Stoppa and Lefkowitz (2010), for example, found that during the first three semesters of college, religious attendance declines across demographic conditions and religious affiliations, heavily mediated by the latter. With approximately 62% of American high school graduates entering institutions of higher education, college experiences play an important role in forming religious

beliefs and practices emerging during early adulthood (Braskamp, 2008; Milevsky et al., 2008; Uecker et al., 2007). However, as the religious participation undergoes substantial change in these years, religious beliefs themselves do not (Desmond et al., 2011). In fact, many researchers have found that the importance of one's religion become greater during this time (Astin & Astin, 2003). Although it was evidenced that the importance of one's religion declined since 1990 among youth from most industrialized nations, American adolescents and emerging adults stand as exceptions to the global secularization trend (Inglehart, 2004). The current study models *attendance* of religious services; all interpretations of the trends presented within this study must be limited to the behavioral component of the religiosity construct. For a thorough discussion of current trends in conceptualizing and measuring religiosity, see DeHaan, Younker, and Affholter (2011).

The abundant cross-sectional evidence for decline in religious involvement during emerging adulthood, however, does not result in strong developmental theories explaining the nature of this change. Cross-sectional data simply do not provide the support for theoretical developments that come from representative samples that combine behavioral and psychological measures of religiosity over multiple time points. Religiosity as a construct offers researchers unique challenges in data collection that the field began to address only recently, most notably with the National Study of Youth and Religion (Smith, Denton, Faris, & Regnerus, 2002). The present study contributes to this effort by offering an in-depth look at the changes in religious attendance using a large number of time points and a nationally representative sample. With rare exceptions (e.g. Day et al., 2009) the utility of the NLSY97 sample has been untapped by the field of religiosity research.

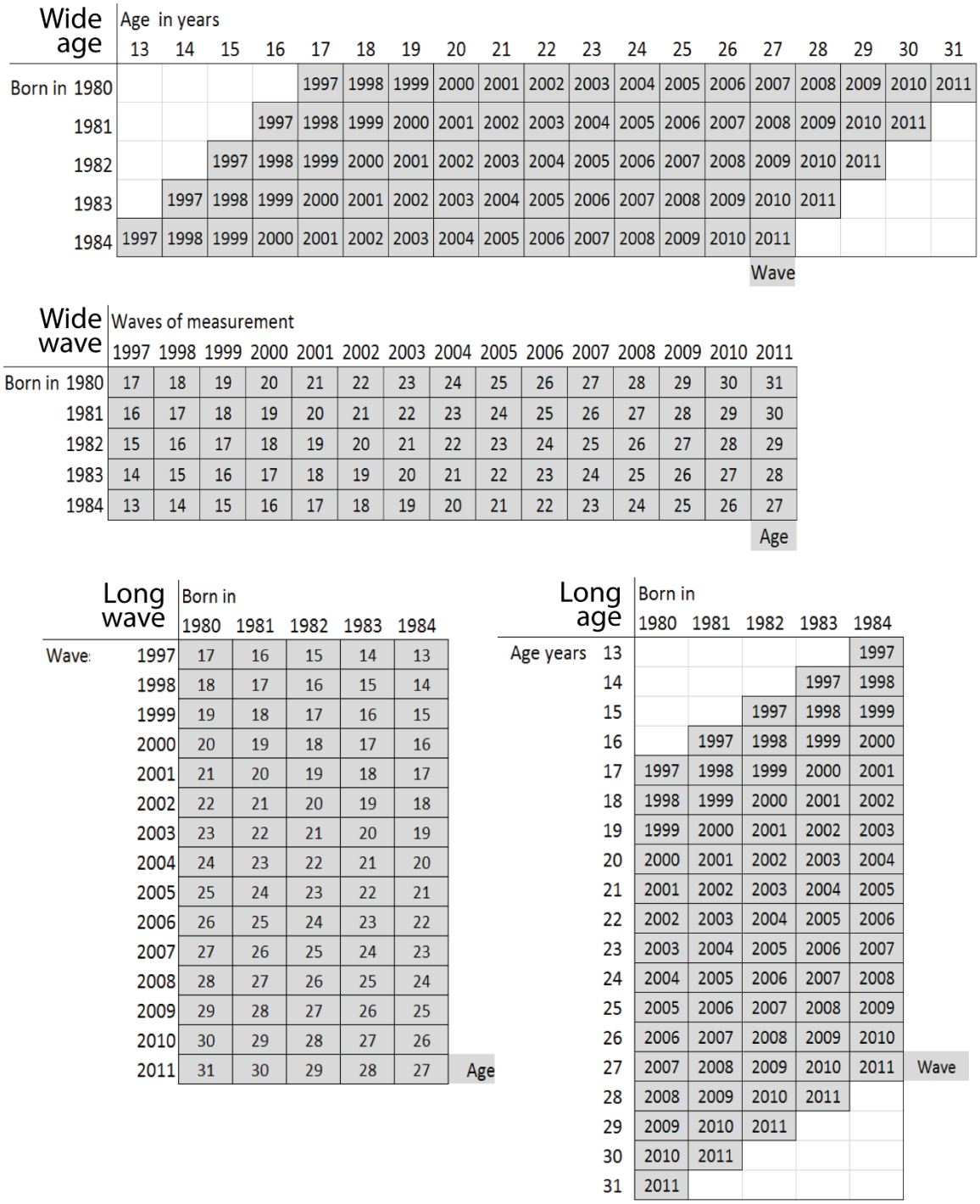
CHAPTER III

METHODS

NLSY97 Sample

The current study uses the data from the NLSY97 study, which is a part of a larger effort of the National Longitudinal Surveys NLS. NLSY97 is a nationally representative sample of households including approximately 9,000 participants. The NLSY97 was based on a household probability sample in which all adolescents between certain ages were surveyed within sampled households. Selected individuals, born between 1980 and 1984, were 12 to 16 years old as of December 31, 1996. They were interviewed annually, starting in 1997 and continuing until the present.

As of the current date (April 2014), there are 15 publically available rounds of NLSY97 data (1997-2011), the reports for the other rounds are still to be released. The present study focuses on the span of 12 time points (2000 – 2011) for which an uninterrupted measure of church attendance was taken. We follow American youth starting in their teens (13-17 years of age) until early adulthood (27-31 years of age). Figure 3.1 shows the structure of NLSY97 measurements using two metrics of time (wave and age) and two data formats (wide and long).



Wide age	Age in years																		
	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Born in 1980					1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1981				1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	
1982			1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011		
1983		1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011			
1984	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011				
															Wave				

Wide wave	Waves of measurement														
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Born in 1980	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1981	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1982	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1983	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1984	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
															Age

Long wave	Born in				
	1980	1981	1982	1983	1984
Wave: 1997	17	16	15	14	13
1998	18	17	16	15	14
1999	19	18	17	16	15
2000	20	19	18	17	16
2001	21	20	19	18	17
2002	22	21	20	19	18
2003	23	22	21	20	19
2004	24	23	22	21	20
2005	25	24	23	22	21
2006	26	25	24	23	22
2007	27	26	25	24	23
2008	28	27	26	25	24
2009	29	28	27	26	25
2010	30	29	28	27	26
2011	31	30	29	28	27
					Age

Long age	Born in				
	1980	1981	1982	1983	1984
Age years 13					1997
14				1997	1998
15			1997	1998	1999
16		1997	1998	1999	2000
17	1997	1998	1999	2000	2001
18	1998	1999	2000	2001	2002
19	1999	2000	2001	2002	2003
20	2000	2001	2002	2003	2004
21	2001	2002	2003	2004	2005
22	2002	2003	2004	2005	2006
23	2003	2004	2005	2006	2007
24	2004	2005	2006	2007	2008
25	2005	2006	2007	2008	2009
26	2006	2007	2008	2009	2010
27	2007	2008	2009	2010	2011
28	2008	2009	2010	2011	
29	2009	2010	2011		
30	2010	2011			
31	2011				

Wave

Figure 3.1 Overview of the age-period-cohort structure of the NLSY97

Data and Measures

Selected variables

Religiosity is a multifaceted construct and frequently calls for a psychometric scale to be measured properly (Rohrbaugh & Jessor, 1975). Psychometric scales of religiosity consist of many (sometimes dozens) of questions that span the multidimensional surface of the construct. Although psychometrically sound, such measures can be prohibitively expensive to administer in longitudinal studies. The NLSY97 contains a few items mapping into the domain of religiosity; a description of them follows.

The items of the NLSY97 that were available to operationalize religious involvement for this study can be conceptualized in relation to two dimensions from Cattell's (1966; 1988) databox, shown in Figure 3.2

VARIABLE TITLE	Units	Codename														
CV_SAMPLE_TYPE	1/0	*sample*	1997													
PUBID, YOUTH CASE IDENTIFICATION CODE	integers	*id*	1997													
KEY!**SEX**, RS GENDER	m/f	*sex*	1997													
KEY!**RACE**_ETHNICITY, COMBINED RACE AND ETHNICITY	b/h/m/o	*race*	1997													
KEY!BDATE, RS BIRTHDATE **MONTH**/YEAR	01-12	*bmonth*	1997													
KEY!BDATE, RS BIRTHDATE MONTH/**YEAR**	years	*byear*	1997													
HOW OFTEN PR **ATTEND** CHURCH IN LAST YEAR?	1-8	*attendPR*	1997													
WHAT IS **PR**S CURRENT **RELI**GIOUS **PREF**ERENCE?	1-8	*relprefPR*	1997													
WHAT **RELI**GION WAS **PR RAISED** IN?	1-8	*elraisedPR*	1997													
RS **AGE** IN **MON**THS AS OF INTERVIEW DATE	months	*agemon*	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
RS **AGE** AT INTERVIEW DATE	years	*ageyear*	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
# DAYS PER WEEK TYPICALLY **FAM**ILY DOES SOMETHING **RELI**GIOUS	# days	*famrel*	1997	1998	1999	2000											
HOW OFTEN R **ATTENDED** WORSHIP SERVICE IN PAST 12 MONTHS	1-8	*attend*				2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
R DOES NOT NEED RELIGION FOR GOOD **VALUES**	y/n	*values*						2002			2005			2008			2011
GOD NOTHING **TO DO** HAPPENS TO R	y/n	*todo*						2002			2005			2008			2011
R BELIEVES RELIGIOUS TEACHINGS ARE TO BE **OBEYED** EXACTLY AS WRITTEN	y/n	*obeyed*						2002			2005			2008			2011
R **PRAY**S MORE THAN ONCE A DAY	y/n	*pray*						2002			2005			2008			2011
R ASKS GOD HELP MAKE **DECISIONS**	y/n	*decisions*						2002			2005			2008			2011
WHAT IS R'S CURRENT **RELIGIOUS** PREFERENCE?	cats:35	*relpref*									2005			2008			2011
R A **BORN-AGAIN** EVANGELICAL CHRISTIAN?	y/n	*bornagain*												2008			2011
IMPORTANCE OF RELIGIOUS **FAITH** IN DAILY LIFE	1-5	*faith*												2008			2011
HOW OFTEN R FELT CALM AND **PEACEFUL** IN PAST MONTH	1-4	*calm*				2000		2002		2004		2006		2008		2010	
HOW OFTEN R FELT DOWN OR **BLUE** IN PAST MONTH	1-4	*blue*				2000		2002		2004		2006		2008		2010	
HOW OFTEN R HAS BEEN A **HAPPY** PERSON IN PAST MONTH	1-4	*happy*				2000		2002		2004		2006		2008		2010	
HOW OFTEN R **DEPRESSED** IN LAST MONTH	1-4	*depressed*				2000		2002		2004		2006		2008		2010	
HOW OFTEN R HAS BEEN A **NERVOUS** PERSON IN PAST MONTH	1-4	*nervous*				2000		2002		2004		2006		2008		2010	
HOW MANY HOURS PER WEEK DOES R WATCH **TELEVISION**	cats:6	*tv*						2002					2007	2008	2009	2010	2011
HOW MANY HOURS PER WEEK DOES R USE A **COMPUTER**	cats:6	*computer*						2002					2007	2008	2009	2010	2011
CURRENTLY HAVE ACCESS TO INTER**NET**?	y/n	*internet*						2002	2003	2004	2005	2006	2007	2008	2009	2010	2011

Figure 3.2 Databox slice of variables selected from the NLSY97 for analyses

Variables on vertical dimension and occasions on horizontal intersect over grey-filled boxes displaying the year of the wave for which data are available. Empty cells indicate that the item was not on the NLSY97 questionnaire in that round. The variable "attendance" is marked by red in Figure 3.2 to indicate that this will be the primary quantification of religiosity in the statistical

models used in this study. This figure can provide guidance to future studies using the NLSY97 to study expanded operationalizations of religiosity.

The variable dimension of the databox slice is annotated by three identifiers adjacent to the left of the grid. First is the "Variable title", the verbatim item label from NLS Investigator. The column titled "Codename" gives the short name of the variable used in the R code that accompanies the statistical analyses. "Units" describes the scales used to measure the variable.

The light grey background highlights the variables related to religion and spirituality. The first section of items (attendPR, relprefPR, relraisedPR**)** gives data on the religiosity of the parents of the respondents, whose households were sampled into NLSY97. One of the considered perspectives on religiosity, the channeling hypothesis, suggests that parents pass the meme of religiosity concepts onto the children. These three items help evaluate this hypothesis and explore the generational association in religious behavior. The largest grey section lists the items related to the religiosity of the youth, describing their religious behaviors (relpref, attend, pray, decisions) and attitudes (values, todo, obeyed, bornagain, faith).

Context variables and covariates are on white background. The top section gives basic demographics: the month (bmonth) and year (byear) of birth, sex (sex), race (race), as well as the indicator whether the individual is a member of the cross-sectional sampling or a special oversample of minorities (sample). Two variables measuring age are located between the religiosity sections: age at the time of the interview in months (agemon) and age in years (ageyear). Those are not derivatives of each other, but, understandably, are closely related (details on the measures of age in NLSY97 are given in the Results chapter). At the bottom are self-reports on emotional wellbeing (calm, blue, happy, depressed, nervous) and media activities (internet, computer, tv) of respondents. To review the original questionnaire cards for the NLSY97 survey, as well as descriptive statistics for the selected variables, see the Descriptives section in the Results chapter. Although not all variables described here are actually used in the models of this study, I give context to show what NLSY97 has to offer in testing substantive theories about change in religiosity, perhaps for the future studies. I explore these directions in the Discussion chapter.

Data structures

All models in the study are applied to the same data – records of self-reported church (worship) attendance from 2000 to 2011 (indicated in red in Figure 3.2). The graphical and syntactical expression of the models and their properties used in present work relies on good understanding of the data structures. This section describes the focal dataset and prepares the way for discussing the research methodology to follow. A report in the Appendix narrates the steps in data preparation starting with accessing the gateway to NLS data online (NLS Web Investigator) and ending with the production of a groomed dataset, used as the starting point for each modeling method.

The dataset produced by the report in the Appendix directly relates to the databox slice in Figure 3.2. However, to match the data structures required by the estimation routine, the databox slice was transposed, distributing variables on the horizontal axis. A new column variable *year* placed the wave values, displayed in the grey boxes of the databox slice, onto the vertical dimension. As displayed in Figure 3.3, it separated two kinds of variables: those whose values do not change with time and those measured at multiple occasions. This distinction will be of convenience in later discussion of statistical models.

The dataset in figure 3.3 is referred to as **dsL** throughout this text and the accompanying R code. It defines the scope of the NLSY97 data used in the current study and has a direct correspondence to the databox slice from Figure 3.2.

Time Invariant | Time Variant

sample	id	sex	race	bmonth	byear	attendPR	relprefPR	relraisedPR	year	agemon	ageyear	famrel	attend	values	todo	obeyed	pray	decisions	relpref	bornagain	faith	calm	blue	happy	depressed	nervous	tv	computer	internet
1	1	2	4	9	1981	7	21	21	1997	190	15	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	1	2	4	9	1981	7	21	21	1998	206	17	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	1	2	4	9	1981	7	21	21	1999	219	18	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	1	2	4	9	1981	7	21	21	2000	231	19	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	3	3	3	3	3	NA	NA	NA
1	1	2	4	9	1981	7	21	21	2001	243	20	NA	6	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	1	2	4	9	1981	7	21	21	2002	256	21	NA	2	1	1	1	0	1	NA	NA	NA	4	2	3	2	1	2	5	NA
1	1	2	4	9	1981	7	21	21	2003	266	22	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1
1	1	2	4	9	1981	7	21	21	2004	279	23	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	4	1	4	1	1	NA	NA	0
1	1	2	4	9	1981	7	21	21	2005	290	24	NA	1	0	1	0	0	1	21	NA	NA	NA	NA	NA	NA	NA	NA	NA	1
1	1	2	4	9	1981	7	21	21	2006	302	25	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	4	1	4	1	1	NA	NA	1
1	1	2	4	9	1981	7	21	21	2007	313	26	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2	6	1
1	1	2	4	9	1981	7	21	21	2008	325	27	NA	1	0	1	0	0	1	21	NA	3	3	3	3	3	3	NA	NA	1
1	1	2	4	9	1981	7	21	21	2009	337	28	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2	6	1
1	1	2	4	9	1981	7	21	21	2010	350	29	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	3	3	3	3	3	1	6	1
1	1	2	4	9	1981	7	21	21	2011	360	29	NA	1	0	1	0	0	1	21	NA	1	NA	NA	NA	NA	NA	NA	6	1
1	2	1	2	7	1982	NA	NA	NA	1997	178	14	3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	2	1	2	7	1982	NA	NA	NA	1998	196	16	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Figure 3.3 Generic dataset used in the current study, view for one respondent.

All models work with the same primary outcome (church attendance) and use time and the age of respondents to predict its change. These data are contained in four columns of **dsL**,

which are subset in Figure 3.4: id, byear (birth year of respondents), year (survey year) and attend (church attendance, the outcome measure). The latter item first appeared in NLSY97 only in 2000, so years 1997-1999 are omitted. Extending this structure, I express statistical models, connecting them to the code that estimates them, in the spirit of reproducible research.

id	byear	year	attend
1	1981	2000	1
1	1981	2001	6
1	1981	2002	2
1	1981	2003	1
1	1981	2004	1
1	1981	2005	1
1	1981	2006	1
1	1981	2007	1
1	1981	2008	1
1	1981	2009	1
1	1981	2010	1
1	1981	2011	1

id	byear	year	attend	cohort	timec	timec2	timec3
1	1981	2000	1	1	0	0	0
1	1981	2001	6	1	1	1	1
1	1981	2002	2	1	2	4	8
1	1981	2003	1	1	3	9	27
1	1981	2004	1	1	4	16	64
1	1981	2005	1	1	5	25	125
1	1981	2006	1	1	6	36	216
1	1981	2007	1	1	7	49	343
1	1981	2008	1	1	8	64	512
1	1981	2009	1	1	9	81	729
1	1981	2010	1	1	10	100	1000
1	1981	2011	1	1	11	121	1331

Figure 3.4 (left) Basic modeling data view
Figure 3.5 (right) Basic data structure extended for LCM

For example, consider Figure 3.5, in which the basic dataset was modified and augmented with several additional variables to match the structure of latent curve models. Timec is a centered variable (*timec* = *year* - 2000), and represents years since 2000. Another derived variable is cohort (*cohort* = *byear-1980*), which gives the age difference of the respondent with respect to the oldest cohort. Additionally, three shapes are added to quantify time effects: linear, quadratic, and cubic, represented by variables *timec*, *timec2*, and *timec3* respectively. The values for these effects are stored in the lambda matrix, to which the next section refers in LCM specification. Using the names of these variables in the estimation syntax of lme4, one can fit a variety of multilevel growth curve models. For example, the following code

```
lmer (attend ~  1  + timec + timec2 + timec3 + cohort
            + cohort:timec + cohort:timec2
            + (1 + timec + timec2 + timec3 | id))
```

specifies a multilevel model with occasions nested within individuals, three predictors on the first level, all modeled as random effects, with linear and quadratic effects regressed on age

difference at the second level. This model can be specified either in the LCM tradition (Bollen & Curran, 2006) or multilevel tradition (Snijders & Bosker, 2012)[5] as follows:

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$

$$\boldsymbol{\eta}_i = \boldsymbol{\mu}_\eta + \mathbf{\Gamma}\mathbf{w}_i + \boldsymbol{\zeta}_i$$

$$\begin{bmatrix} y_{i0} \\ y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \\ y_{i6} \\ y_{i7} \\ y_{i8} \\ y_{i9} \\ y_{i10} \\ y_{i11} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \\ 1 & 5 & 25 & 125 \\ 1 & 6 & 36 & 216 \\ 1 & 7 & 49 & 343 \\ 1 & 8 & 64 & 512 \\ 1 & 9 & 81 & 729 \\ 1 & 10 & 100 & 1000 \\ 1 & 11 & 121 & 1331 \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_{1i} \\ \beta_{2i} \\ \beta_{2i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i0} \\ \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \\ \varepsilon_{i5} \\ \varepsilon_{i6} \\ \varepsilon_{i7} \\ \varepsilon_{i8} \\ \varepsilon_{i9} \\ \varepsilon_{i10} \\ \varepsilon_{i11} \end{bmatrix}$$

$$\begin{bmatrix} \alpha_i \\ \beta_{1i} \\ \beta_{2i} \\ \beta_{3i} \end{bmatrix} = \begin{bmatrix} \mu_\alpha \\ \mu_{\beta 1} \\ \mu_{\beta 2} \\ \mu_{\beta 3} \end{bmatrix} + \begin{bmatrix} \gamma_{01} \\ \gamma_{11} \\ \gamma_{21} \end{bmatrix} cohort_i + \begin{bmatrix} \zeta_{\alpha i} \\ \zeta_{\beta 1 i} \\ \zeta_{\beta 2 i} \\ \zeta_{\beta 3 i} \end{bmatrix} \qquad \begin{bmatrix} \zeta_{\alpha i} \\ \zeta_{\beta 1 i} \\ \zeta_{\beta 2 i} \\ \zeta_{\beta 3 i} \end{bmatrix} \sim N \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{\alpha\alpha} & & & \\ \psi_{\alpha\beta 1} & \psi_{\beta 1\beta 1} & & \\ \psi_{\alpha\beta 2} & \psi_{\beta 1\beta 2} & \psi_{\beta 2\beta 2} & \\ \psi_{\alpha\beta 3} & \psi_{\beta 2\beta 3} & \psi_{\beta 2\beta 3} & \psi_{\beta 3\beta 3} \end{bmatrix}$$

$$y_{ti} = \beta_{0i} + \beta_{1i} timec_{ti} + \beta_{2i} timec_{ti}^{\,2} + \beta_{3i} timec_{ti}^{\,3} + \varepsilon_{ti} \qquad \varepsilon_{ti} \sim N\left([0], [\sigma^2]\right)$$

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01} cohort_i + u_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11} cohort_i + u_{1i} \\ \beta_{2i} &= \gamma_{20} + \gamma_{21} cohort_i + u_{2i} \\ \beta_{3i} &= \gamma_{30} + u_{3i} \end{aligned} \qquad \begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \\ u_{3i} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & & & \\ \tau_{10} & \tau_{11} & & \\ \tau_{20} & \tau_{21} & \tau_{22} & \\ \tau_{30} & \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \right)$$

[5] Color is used to help match the elements in the specification to the elements of the graph produced by the model sequencer. Thus, blue is used to identify gamma estimates, green points to the predictors at the second level, and red refers to the estimates of the random effects and the residual variance. Also note the change of subscripts from *i,j* in Snijders & Bosker (2011) notation to *t,i* in Bollen & Curran (2006) notation.

Focus Outcome Variable: Church Attendance

The focal variable of interest is attend, an item measuring church attendance in the current year. Although it was recorded on an approximately ordinal scale, its precision allows us to treat it as quantitative for the purpose of fitting statistical models. We have data on attendance for 12 years, from 2000 to 2011. Figure 3.6 gives a cross-sectional frequency distribution of the data across the years, assuming attrition was not related to the outcome.

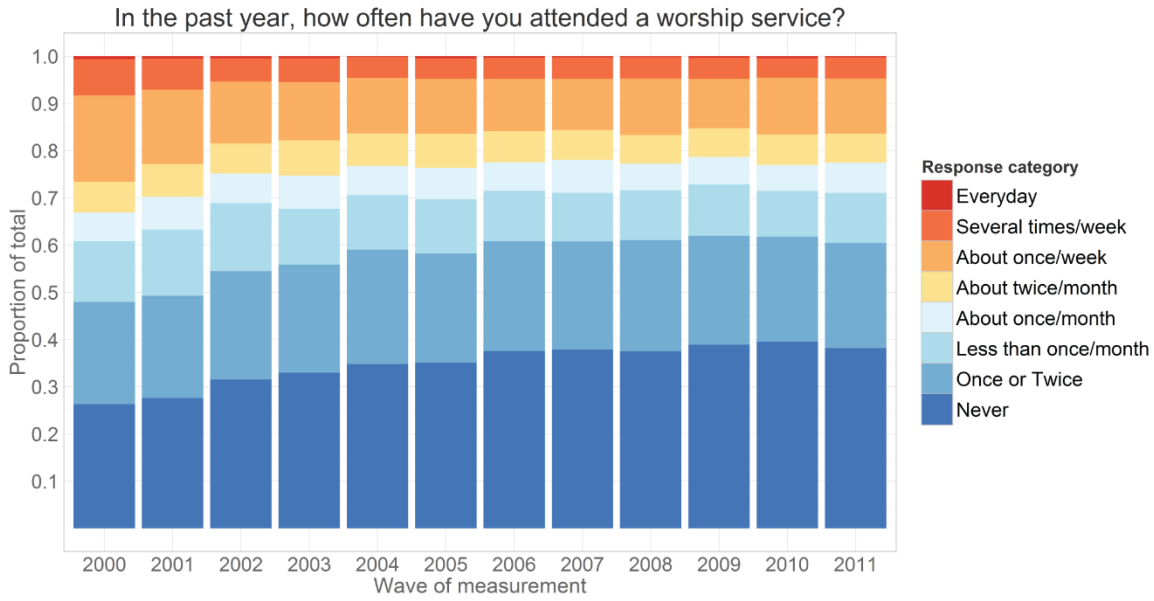


Figure 3 6 Relative frequency of responses for each round of observation

Modeling transitions between the frequencies of endorsing particular response items across time will be the focus of using a Markov model, which treats a set of cross-sectional representations. However, LCM and GMM work with longitudinal data, modeling the trajectory of each individual. To illustrate, the trajectories of subjects with id 4, 25, 35, and 47 are plotted in Figure 3.7

Figure 3.7 Trajectories of church attendance: four respondents over waves

The respondent id = 35 reported attending no worship services in any of the years, while respondent id = 25 attended quite often (indicating weekly attendance in 8 out of the 12 years). Individual id = 47 started as a regular attendee of religious services in 2000 (5 = "about twice a month"), then gradually declined his involvement to nil in 2009 and on. Respondent id = 4, on the other hand started off with a rather passive involvement, reporting attended church only "Once or twice" in 2000, maintained a low level of participation throughout the years, and then increased his attendance in 2011. Each of these trajectories implies a story, a life scenario related to each person's religious involvement. Why one person grows in his religious involvement, whereas another declines, or never develops an interest in the first place, is the empirical subject of the current investigation.

Previous research in religiosity indicated that age might be one of the primary factors explaining interindividual differences in church attendance. To examine the role of age, we change the metric of time from waves of measurement, as in the Figure 3.7, to biological age, calculated as age in months at the time of the interview and converted to years. This re-alignment is represented graphically in Figure 3.8.

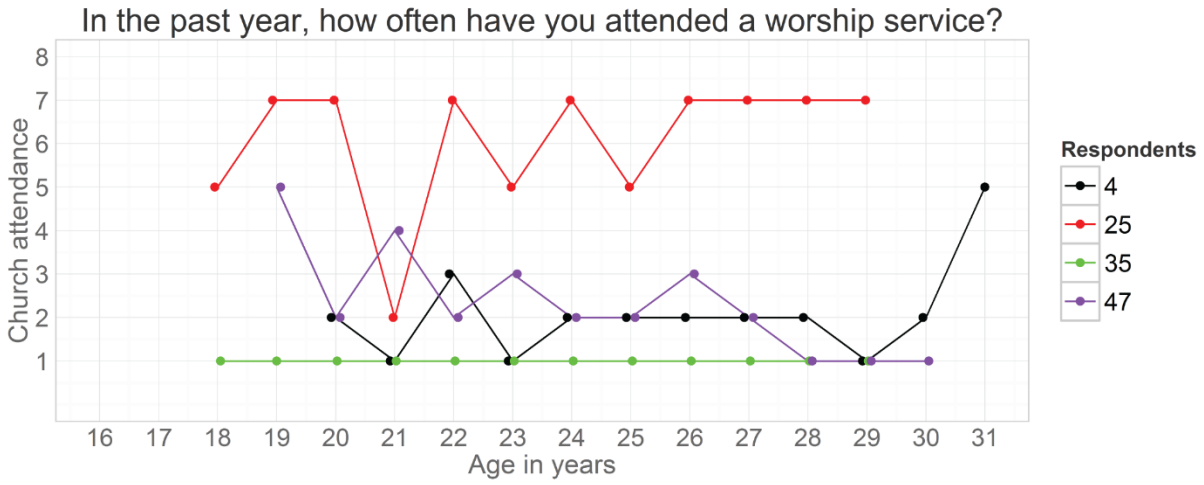


Figure 3.8 Trajectories of church attendance: four respondents over age

Research Methodology

The current study analyses how religiosity changes during adolescence and young adulthood, across ages. Latent curve models (LCM) test certain shapes of the time effect (linear, quadratic, and cubic) in a search for the best-fitting common trajectory that describes church attendance between 2000 and 2011, regressing random terms on age indicator.

Model Specification

The latent curve models (LCMs) considered in the analysis can be expressed in latent curve (Bollen & Curran, 2006), or in the multilevel regression tradition (Snijders & Bosker, 2012)[6] along with some color conventions, explained in the footnote on page 33.

[6] S&B use *i* and *j* for first and second level respectively, however I changed it to *t* and *i* to be consistent with Bollen & Curran (2006) notation and also because it offers a mnemonic "*t* for time and *i* for individual."

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$$
$$\boldsymbol{\eta}_i = \boldsymbol{\mu}_\eta + \mathbf{\Gamma}\mathbf{w}_i + \boldsymbol{\zeta}_i$$

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\mu}_\eta + \mathbf{\Lambda}\mathbf{\Gamma}\mathbf{w}_i + \mathbf{\Lambda}\boldsymbol{\zeta}_i + \boldsymbol{\varepsilon}_i$$

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} \quad \mathbf{\Lambda} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (T-1)^1 & \cdots & (T-1)^n \end{bmatrix} \quad \boldsymbol{\eta}_i = \begin{bmatrix} \alpha_i \\ \beta_{1i} \\ \vdots \\ \beta_{Pi} \end{bmatrix} \quad \boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

$$\boldsymbol{\eta}_i = \begin{bmatrix} \alpha_i \\ \beta_{1i} \\ \vdots \\ \beta_{Pi} \end{bmatrix} \quad \boldsymbol{\mu}_\eta = \begin{bmatrix} \mu_\alpha \\ \mu_{\beta 1} \\ \vdots \\ \mu_{\beta P} \end{bmatrix} \quad \mathbf{\Gamma} = \begin{bmatrix} \gamma_{\alpha 1} & \gamma_{\alpha 2} & \cdots & \gamma_{\alpha K} \\ \gamma_{\beta 11} & \gamma_{\beta 12} & \cdots & \gamma_{\beta 1K} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{\beta P1} & \gamma_{\beta P1} & \cdots & \gamma_{\beta PK} \end{bmatrix} \quad \mathbf{w}_i = \begin{bmatrix} w_{1i} \\ w_{2i} \\ \vdots \\ w_{Ki} \end{bmatrix} \quad \boldsymbol{\zeta}_i = \begin{bmatrix} \zeta_\alpha \\ \zeta_{\beta 1} \\ \vdots \\ \zeta_{\beta P} \end{bmatrix}$$

$\mathbf{y}_i$ - A vector of responses of individual *i* for times T

$\mathbf{\Lambda}$ - Matrix of weights for P functions of time

$\boldsymbol{\eta}_i$ - Vector of person-specific weights for P time effects

$\boldsymbol{\mu}_\eta$ - Vector of fixed effect estimates (mean/intercept)

$\mathbf{\Gamma}$ - Matrix of fixed effect estimates for $\mathbf{w}_i$ with K predictors

$\mathbf{w}_i$ - Time invariant, fixed predictors of $\boldsymbol{\eta}_i$

$\boldsymbol{\zeta}_i$ - Random effect estimates

$\boldsymbol{\varepsilon}_i$ - Residual variance

T - Total number of time points in the data

P - Total number of time effects estimated in addition to mean/intercept

K - Total number of predictors $\mathbf{w}_i$ on time effects $\boldsymbol{\eta}_i$

$$\begin{bmatrix} \zeta_{\alpha i} \\ \zeta_{\beta 1 i} \\ \vdots \\ \zeta_{\beta P i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{\alpha\alpha} & & & \\ \psi_{\alpha\beta 1} & \psi_{\beta 1\beta 1} & & \\ \vdots & \vdots & \ddots & \\ \psi_{\alpha\beta P} & \psi_{\beta 1\beta P} & \cdots & \psi_{\beta P\beta P} \end{bmatrix} \right)$$

$$y_{ti} = \beta_{0i} + \beta_{1i} time_{1ti} + \ldots + \beta_{Pi} time_{Pti} + \varepsilon_{ti}$$

$$\varepsilon_{ti} \sim N\left([0], [\sigma^2]\right)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01} w_{1i} + \gamma_{01} w_{2i} + \ldots + \gamma_{0K} w_{Ki} + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11} w_{1i} + \gamma_{12} w_{2i} + \ldots + \gamma_{0K} w_{Ki} + u_{1i}$$

$$\vdots \quad \vdots \quad \vdots$$

$$\beta_{Pi} = \gamma_{P0} + \gamma_{P1} w_{1i} + \gamma_{P2} w_{2i} + \ldots + \gamma_{PK} w_{Ki} + u_{Ki}$$

$$\begin{bmatrix} u_{0i} \\ u_{1i} \\ \vdots \\ u_{Pi} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & & & \\ \tau_{10} & \tau_{11} & & \\ \vdots & \vdots & \ddots & \\ \tau_{P0} & \tau_{P1} & \cdots & \tau_{PP} \end{bmatrix} \right)$$

where time occasions *t* are nested within individuals *i* and each of the time effects P is regressed on *K* time invariant predictors $\mathbf{w}_i$. Such notation reflects some of the logic of lme4 syntax, however matrix algebra notation is more useful for other purposes. The present work relies on both notations to provide a broad perspective on the mathematical structure of these longitudinal models. The graphical methods illustrated in the first chapter will annotate the analysis fitting the model above to the NLSY97, providing additional nuance and interpretation.

CHAPTER IV

RESULTS

Descriptives

This section pursues two distinct goals. The first is to prepare the reader for the modeling exercise that is to follow. The second is to familiarize the reader with the structure and the potential of the NLSY97.

Reproducible research, the aspiration of the current work, ideally presents the reader not only with the distilled statements about the nature of the world and the means of replicating the analysis, but also with the room to take the study into the directions unforeseen by the initial author. Providing the reader with understanding of the structure of the NLSY97, and particularly its longitudinal aspects, may (I hope) incline the reader to further exploit the utility of this sample and minimize starting costs of initiating a research project.

Because of these dual goals, the graphical presentations may in a few cases provide more expansive information than what is required to link the graph to the modeling exercise presented later in the Results chapter. In those cases, the additional information is presented to give the reader further depth of understanding of the NLSY97 data.

Age and basic demographic

The NLSY97 includes 8,983 respondents, of which 6,748 were selected from randomly sampled households, and 2,236 came from the oversample of racial minorities. The demographic composition of the sample is given in Figure 4.1.

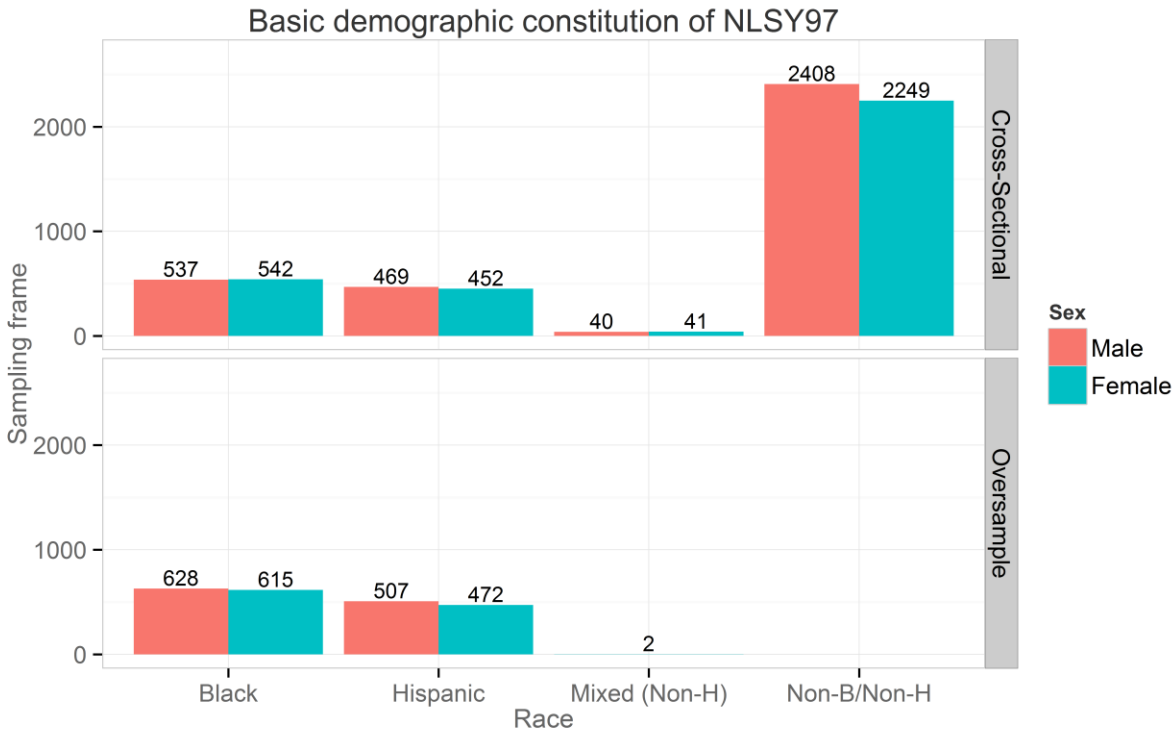


Figure 4.1 Race demographics in NLSY97: counts of respondents

Respondents' age was of particular interest in explaining church attendance. The NLSY97 contains static and dynamic indicators of age. Variables **byear** and **bmonth** (static) were recorded once in 1997 and contained the birth year and birth month respectively. Two age variables were recorded continuously at each interview: age at the time of the interview in months **agemon** and in years **ageyear** (dynamic). Figure 4.2 shows how births in the NLSY97 sample were distributed over calendar months from 1980 to 1984.

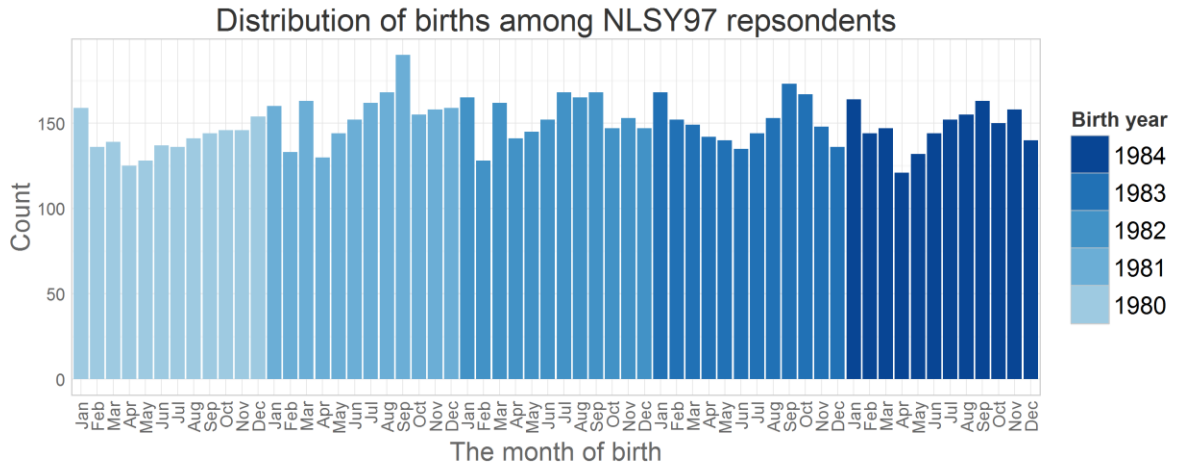


Figure 4.2 Counts of respondents' birth months

The variable **ageyear** records the age in years a respondent reached at the time of the interview. Due to difficulties of administering the survey, time intervals between the waves differed. For example, for one person (**id** = 25) the age was recorded as 21 years for both 2003 and 2004 (see **ageyear**). However, when you examine age in months (**agemon**) you can see this rounding problem disappears once the more precise scale is used (in the table below age is calculated as **agemon**/12). It must be noted however, that the dynamic measure of age was not recorded every year for each respondent and is not of much use, due to its frequent missingness. To avoid numerous missing predictor values, age in years will be calculated as **year - byear**. In this way, we obtain a more consistent measure that could be used in predictive models, although at the expense of some precision. To illustrate the relationship among recorded and computed age variables Table 4.1 lists the complete age data for one respondent.

Tables 4.1 Age data for one respondent in NLSY97.

id	bmonthF	byear	year	agemon	ageyear	age
25	Mar	1983	1997	167	13	13.92
25	Mar	1983	1998	188	15	15.67
25	Mar	1983	1999	201	16	16.75
25	Mar	1983	2000	214	17	17.83
25	Mar	1983	2001	226	18	18.83
25	Mar	1983	2002	236	19	19.67
25	Mar	1983	2003	254	**21**	21.17
25	Mar	1983	2004	261	**21**	21.75
25	Mar	1983	2005	272	22	22.67
25	Mar	1983	2006	284	23	23.67

```
   25      Mar  1983 2007    295       24 24.58
   25      Mar  1983 2008    307       25 25.58
   25      Mar  1983 2009    319       26 26.58
   25      Mar  1983 2010    332       27 27.67
   25      Mar  1983 2011    342       28 28.50
```

Figure 4.3 shows how static age maps onto the dynamic age among the respondents in the wave that was collected in 2000. This graph is not useful in detailed analysis due to the issues mentioned above, however it provides a good snapshot of the age constitution of the sample. The dynamic graph in the appendix animates with frames for each of the rounds of observation.

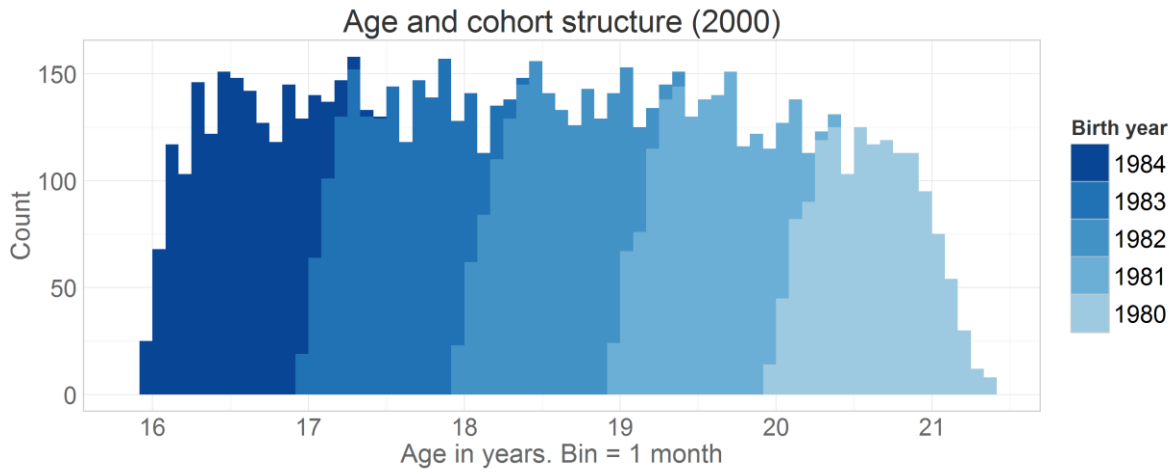


Figure 4.3 Age and cohort structure of NLSY97 respondents in 2000

Church attendance: cross-sectional view

The focal variable of interest is **attend**, the item measuring church attendance for 12 months that preceded the interview date. The questionnaire recorded the responses on an ordinal scale, shown in Figure 4.4.

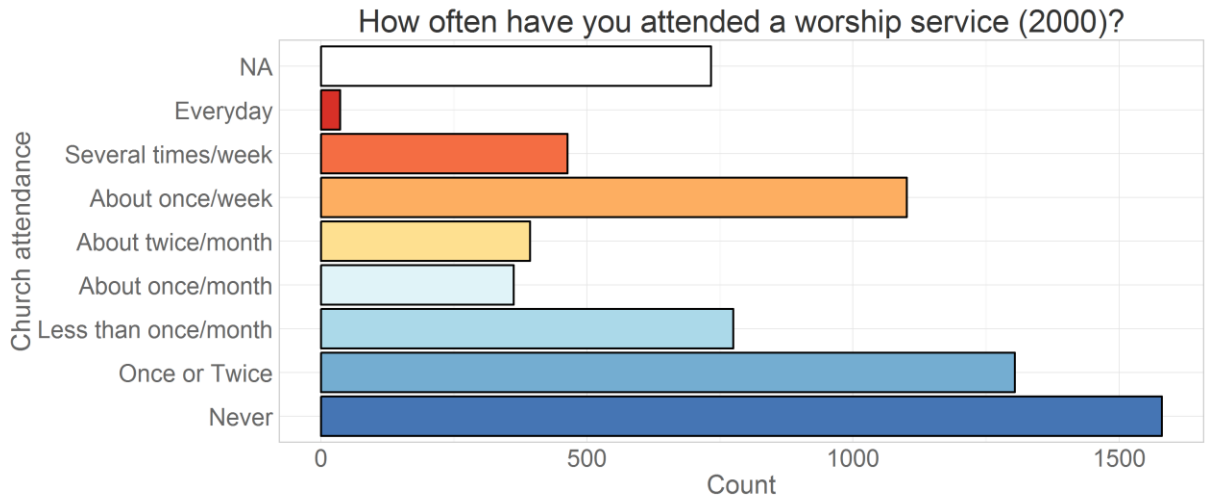


Figure 4.4 Scale for measuring church attendance (8 – Everyday, 1 – Never)

The immediate observation about the focal variable is the bimodal distribution of the responses, with the "Never" or "Once or Twice" response category as one mode and "About once a week" as the other. This makes sense considering the natural regularity of worship services practiced by most religions. Despite the fact that the scale allows for a finer distinction, the distribution of endorsement frequencies invites us to think of going to church more or less as a binary outcome: either you attend church regularly, or you do not. This graph was made using the data only from the member of NLSY97's cross-sectional, representative sample and therefore depicts, with a fair degree of external validity, the religious attitudes and behaviors of the American public in this age range. Considering that more than half of its representation (54.2%) attends church no more than once a month and almost quarter (23.4%) ignores it completely, it won't be an exaggeration to reason that the American young adults are not very religious, at least using church attendance as a criterion.

Church attendance, as discussed in the Methods Section, is one of the standard and among the best measures of religiosity available to longitudinal researchers. Luckily, the NLSY97 tracks it for a period of 12 years, since 2000. Another item about religiosity ("In a typical week, how many days from 0 to 7 do you do something religious as a family such as go to church, pray or read the scriptures together?") was on the questionnaire from 1997 to 2000, overlapping one year with attendance. Theoretically it would interesting to connect these two operationalizations

of religiosity in a longitudinal study, however only a relatively small portion of the sample completed this item prior to 2000 and the available sample size did not afford complex modeling. In addition, numerous missing values in this variable further limited its integration in the present study. Table 4.2 give a two-way frequency count between family religious activity and church attendance. The first columns lists possible responses to the church attendance item, while the first row give possible answers to the question "In a typical week, how many days from 0 to 7 do you do something religious as a family such as go to church, pray or read the scriptures together?"

Table 4.2 Full sample counts in 2000 between family religious activity and church attendance

	0	1	2	3	4	5	6	7	<NA>
Never	914	69	17	16	10	5	6	10	974
Once or Twice	568	196	58	33	17	18	6	15	859
Less than once/month	269	176	30	18	9	9	3	19	523
About once/month	76	136	20	8	6	7	3	8	242
About twice/month	52	147	27	11	9	2	3	10	257
About once/week	74	591	169	40	27	17	17	54	488
Several times/week	36	59	105	78	30	31	14	61	203
Everyday	6	3	1	2	1	2	1	13	24
<NA>	4	1	0	0	0	0	0	1	959

Although the number of the available respondents is small in comparison to the full NLSY97 sample (note the large NA column), the two features of religious involvement, bimodality and prevalence of church avoidance, can nevertheless be recognized in this bivariate data representation, too. Notice a sizable endorsement of "About once/week" by respondents reporting that one day on a typical week their family does something religious together.

Followed over time, the religiosity of adolescent and young adult Americans appears to be declining. Figure 4.5 gives a cross-sectional frequency distribution of the data across the years. Here, missing values are used in the calculation of total number of responses to show the natural attrition of respondents and/or the increased response refusal rate. Assuming that lower rate of response retrieval is not significantly associated with the outcome measure we can remove missing values from the calculation of the total and look at prevalence of response endorsements over time, as Figure 4.6 shows

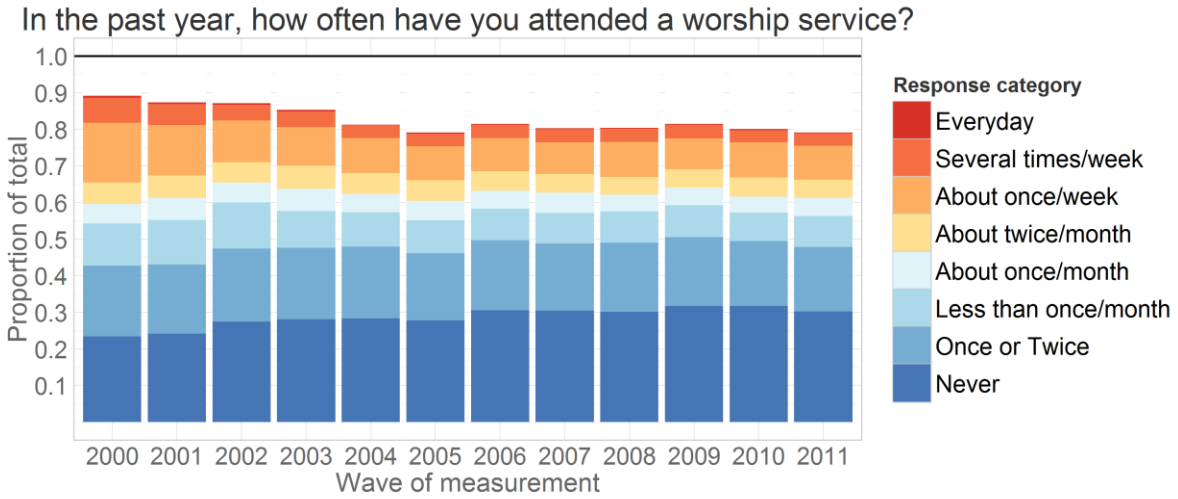


Figure 4.5 Cross-sectional distribution of church attendance categories

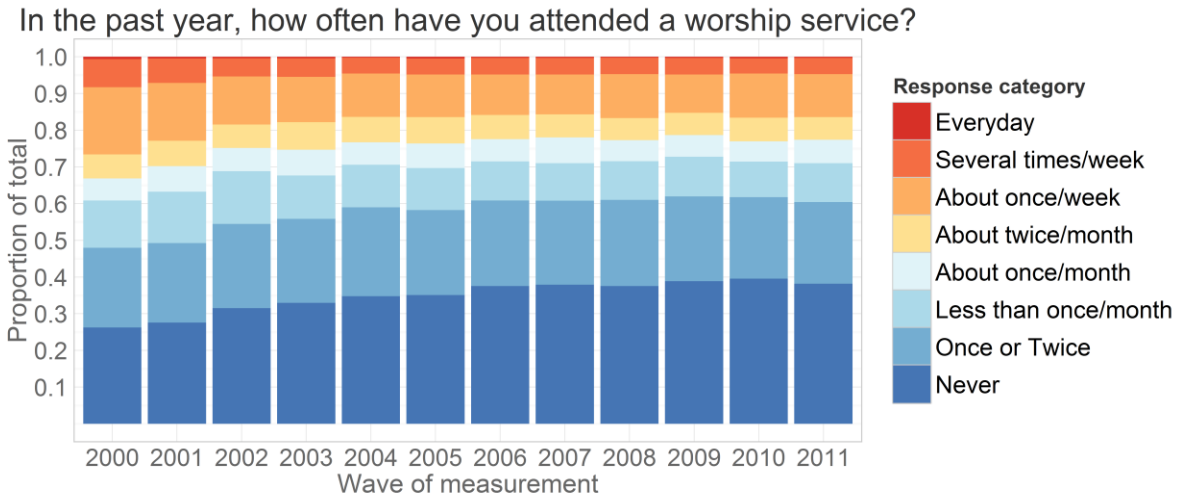


Figure 4.6 Distribution of church attendance as proportions from the total of non-missing values

We see a dominance of blue colors increasing in both views, indicating a change toward a more secular lifestyle. Broad strokes of Figure 4.6 indicate a general decrease in religious involvement in this generation of Americans. To examine the trends with greater precision, Figure 4.7 remaps the same data in a line graph. There we see more clearly how specific categories change over time: "Never" exhibits the sharpest climb, "About once/week" drops rapidly in the first few

rounds for which observation are available but then stabilized around 12%, other categories, as "About twice/month" for example remain relatively stable throughout the years.

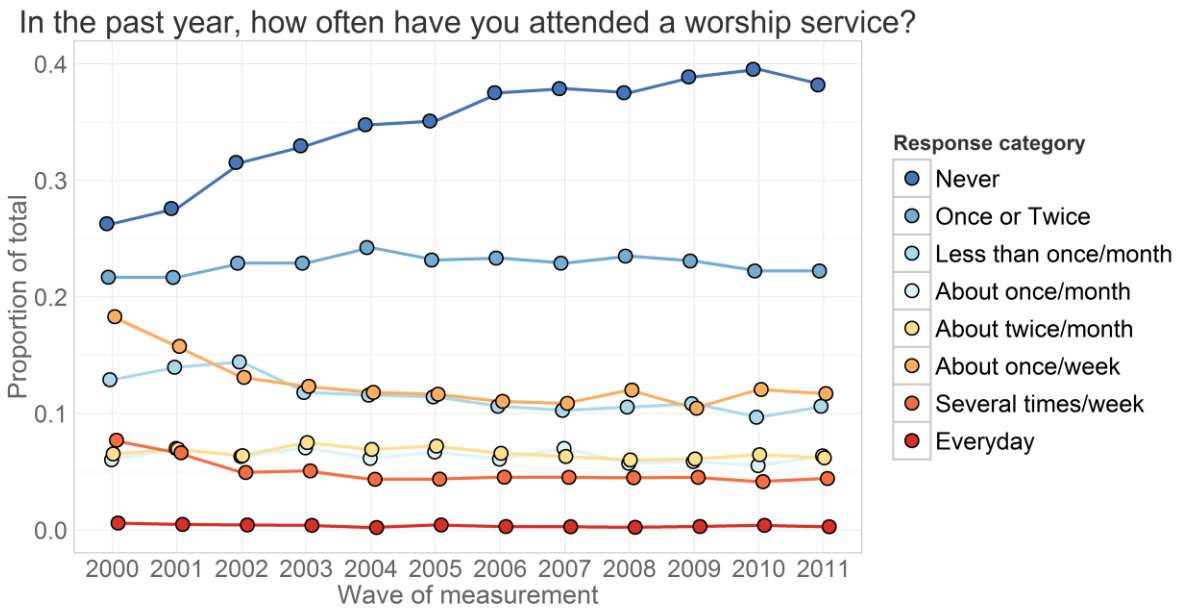


Figure 4.7 Prevalences of church attendance over years. Remapped from 4.6

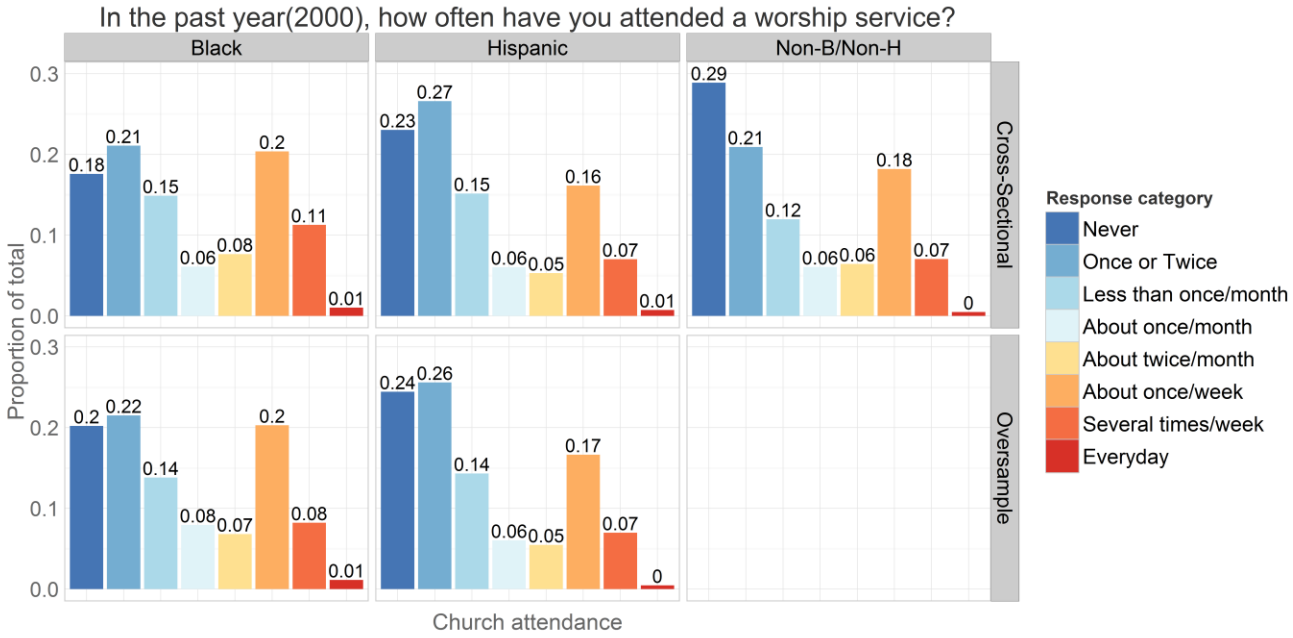


Figure 4.8 Church attendance frequencies across races in 2000

The trend of declining church attendance, however, is not universal. Ethnic groups demonstrate substantial differences in patterns of religious involvement. Figure 4. 8 shows the distribution of responses to the NLSY97 item on church attendance. Supporting the observation that their group is the most polarized, 29% of White (actually non-Black/non-Hispanic) respondents indicated in 2000 that they never went to church in the past year – the highest percentage among racial groups. Both Hispanic and Black seem to be more accepting of nominal attendance (“Once or Twice” is the leading category).

Racial minorities differed substantially not only in the level of initial religious involvement, but also in the rate with which it changed over time. The dynamics of prevalences across racial identifiers are shown in Figure 4.9. The data obtained from respondents, who identified themselves as Non-Black/Non-Hispanic (mostly Whites) make the trend of decreasing

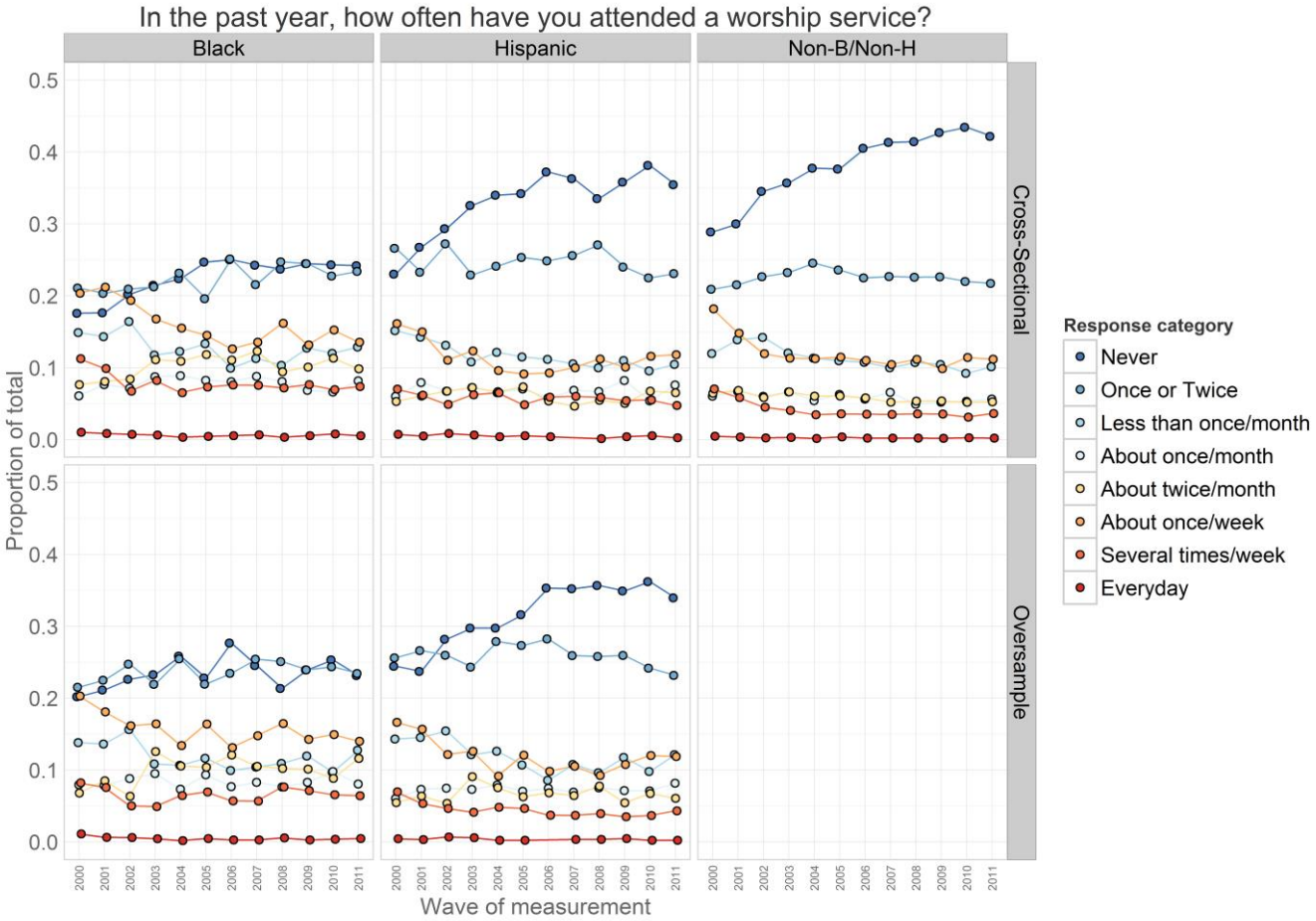


Figure 4.9 Prevalences of church attendance between years 2000 and 2011

church attendance much more clearly pronounced. Between the other two racial categories, Whites are the most polarized in their differences in religious involvement and its dynamics. The gap between those who do not attend church and those who come once or twice a year, while the largest among racial groups at the beginning of the study, only continued to grow, almost tripling by the last round.

To a smaller degree, the same is true of respondents who identified as Hispanic. In contrast to Whites, the difference between these attendance categories ("Never" and "Once or Twice") was reversed among Hispanic respondents, although not by much. The increase in the prevalence of "Never" among Hispanic is not as steep as those of Whites, but similar in pattern and magnitude.

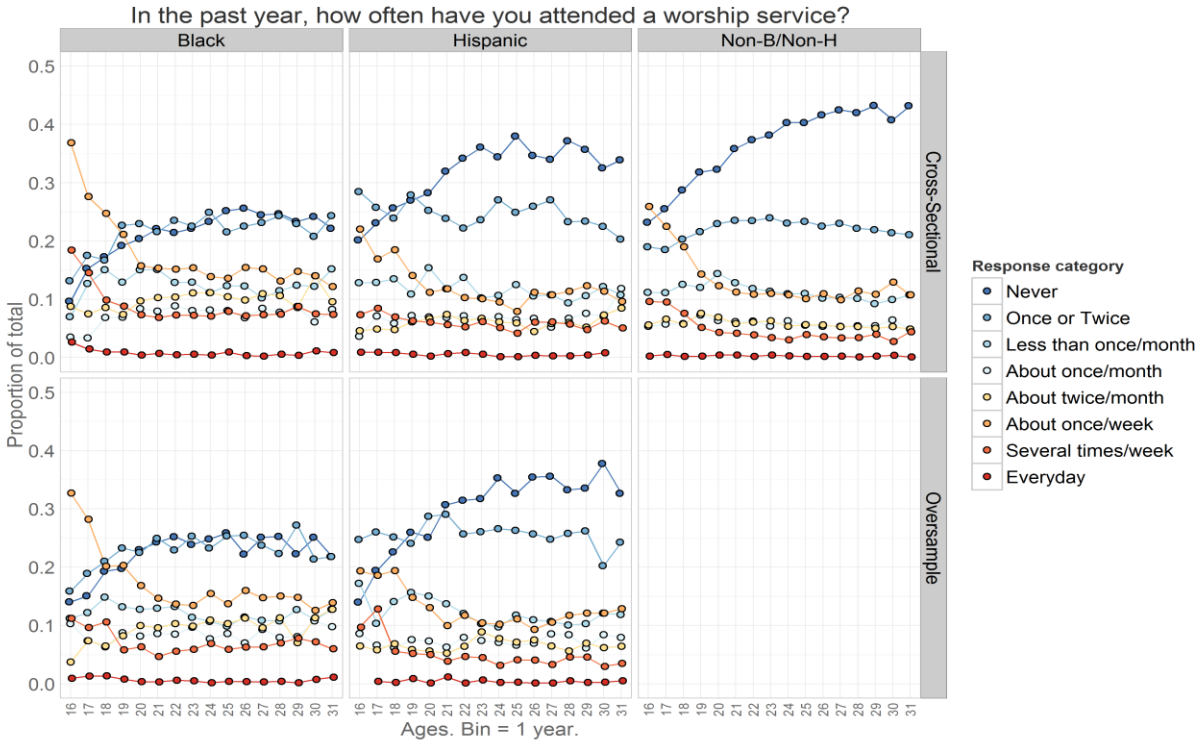


Figure 4.10 Prevalences of church attendance between ages 16 and 32

Blacks seem to exhibit more stability in church attendance than Hispanics and Whites: they experience the smallest surge in the prevalence of endorsing the response "Never", where the trajectory flattens fast and lags behind those of both Hispanic and White respondents.

Another noticeable difference among the races is the curve of the regular church attendance. Hispanic and White respondents descend quickly, reaching the asymptote within around two years. Hispanics demonstrate a more gradual decline in regular attendance, but nothing like Black, whose decline stretches over 7 years . The data collected among the oversample of minorities appears to demonstrate similar patterns, and validates such observations. See temporal animations of these and other graphs to explore the response dynamics.

What drives such dynamics? Time is too easy an answer, because it is confounded with age, period, and cohort effects. Naturally, age and cohort offer richer hypotheses and explanations to developmental researchers. Sociological changes unfold on larger timescales, however, as the NLSY97 stretches the limits of panel studies, both in the resolution of the sample and the span of longitudinal observations.

One of the common ways to untangle the confounded temporal factors is to rescale the metric of time and to re-align the chronology of the study. The patterns of declining church attendance are clearer after changing the metric of time from the rounds of NLSY97 to biological age, as demonstrated in Figure 4.10 and 4.11. Figure 4.10 re-scales the data from Figure 4.9, while Figure 4.11 demonstrates what patterns of church attendance are among 16-yearolds, offering an alternative portrayal to Figure 4.8.

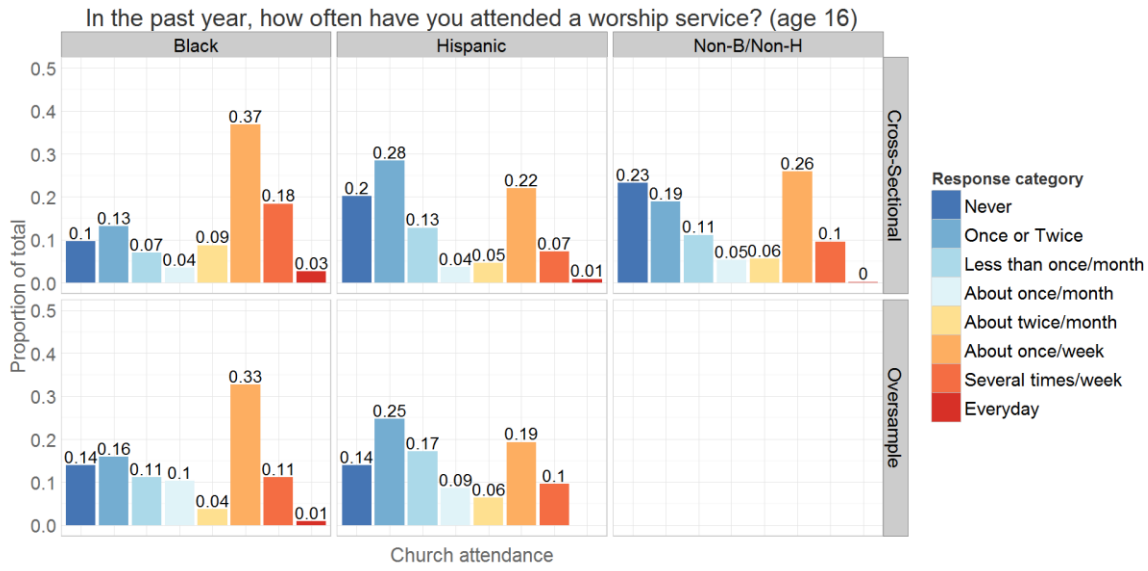


Figure 4.11 Church attendance frequencies across races at age 16

The difference between 4.8 and 4.11 is dramatic: the former describes the group dominated by non-attendance, while the latter gives an opposite picture. The differences among the races, however, preserves the structure we have seen in figures 4.9 and 4.10. Blacks, as the most religiously involved respondents, have a 37% endorsement of weekly church attendance, compared to Whites and Hispanics who are lower by 11 and 15 percentage points, respectively. The difference between these two data views is the age constitution of the selected respondents. Figure 4.8 was produced by data supplied by people of various ages, as young as 16 and as old as 20. In figure 4.11 only the data from 16-year-olds were used, making the group homogeneous in age. Comparing Figure 4.9 with 4.10 and Figure 4.8 with 4.11, we see that patterns of change become more pronounced and more sharply defined. Figure 4.11 especially powerfully demonstrates how much change occurs during the late teen years: the sharp drops in attendance across all races, especially in regular attendance, are steeper than for later ages.

The cross-sectional data described above gave a good overall picture, but left key questions about the change in church attendance unanswered. We can see that the general trend is for respondents to attend church less as time progresses, but it is not clear how individuals contribute to this trend. Does the number of non-goers increase at the expense of fervent churchgoers or those who were only mildly involved in church? Are the prevalences that appear stable across time (see “Several times per week” among Hispanic in Figure 4.8) really stable, or characterized by people moving in and out of this category, creating only an appearance of stability of the prevalence? Is the observed change a result of many individuals changing a little bit, or from drastic changes among a handful of persons? To address these questions we must turn to *longitudinal* data and allow the observed and reproduced intraindividual change to inform the theories about the interindividual differences.

By tracing individuals over time longitudinal methods separate within-person from between-person variability. Figures 3.7 and 3.8 already gave an example of individual trajectories, showing the trajectories of four individuals using the rounds of the study (Figure 3.7) or biological age (Figure 3.8) as the metric of time. Analyzing a large number of such trajectories may be quite challenging. The rest of the chapter demonstrates and discusses the reporting tool developed in this dissertation designed to aid in assessing the interplay between

the interindividual and the intraindividual differences. Building on the existing tools for dynamic reporting, I offer a method to organize, carry out, and communicate the result of a sequence of latent curve models.

Sequence of latent curve models

I opened this work with a discussion of a challenge that has been in development in the last few decades in research and academic circles. Statistical models became so complex that human limitations in attention, perception, and information processing have become relevant to practicing modelers. Much of the challenge facing a modern modeler, however, comes not only from the complexity of statistical structures used to operationalize research theories, but also from organizing and managing their estimation and publication. Despite the wonders of modern computer technology, the time and effort it takes to evaluate a series of models may be onerous, especially in cases involving long sequences and elaborate models.

Recent advances in software technology allowed transforming the modern modeling workflow. Tools such as *knitr* (Xie, 2014) and *pandoc* offered technologies for combining statistical estimation, production of data graphics, and report writing in a single environment of RStudio. The dynamic report developed for this work employs some features of interactive documents to organize the evaluation of LCM sequences. Designed to minimize the information overload, the Interactive system of comparisons and contrasts offers faster and easier evaluation and management of statistical models.

The general form of latent curve models was defined at the end of Chapter 3. I repeat the definition in Figure 4.12 for convenience. This is a general specification; notice that level-1 predictors in the random coefficient equation do not refer to time effects as specifically linear, quadratic, and cubic terms, but rather as some P functions of time, which may include other polynomials, shape factors, piecewise and exponential functions. However, in the more explicit LCM equation, the lambda matrix contains coefficients for polynomial functions specific to current models.

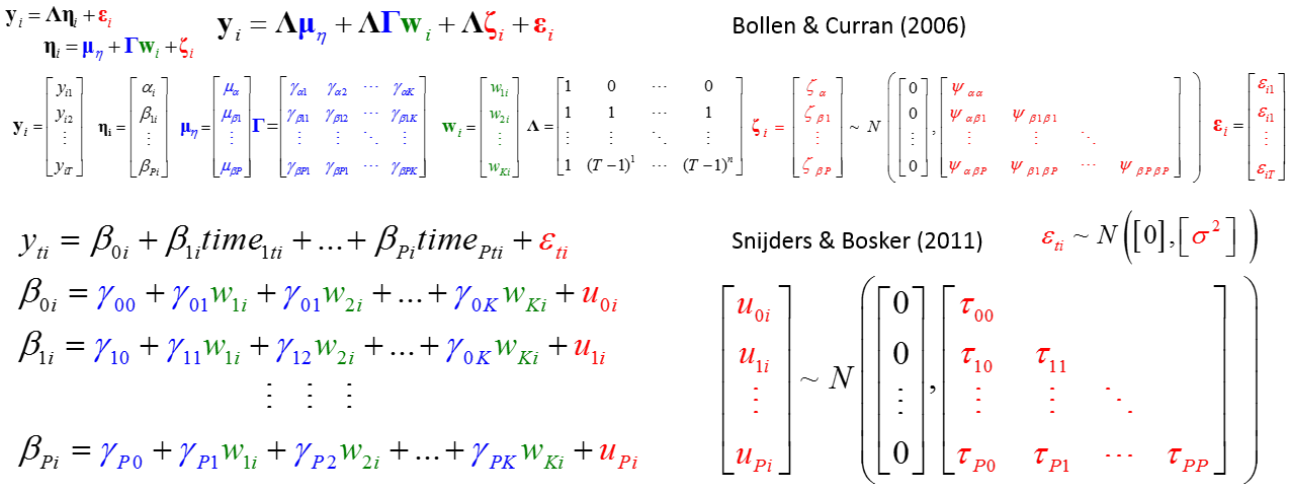


Figure 4.12 General specification of the models used in the study

Fitted models

The primary analytic goal of the study is the examination of how time and the age of respondents interact to explain the observed church attendance in the NLSY97 sample. The models explored three time effects (linear, quadratic, and cubic) which were entered as predictors in the first level of the model. Time was centered at 2000. The second level contained a single predictor *cohort*, which quantified the age difference among the respondents. Cohort predictor was centered at 1984. The model of the maximum complexity contains three functions of time, modeled as random effects, and fixed level-2 predictor *cohort* for each time function:

$$y_{ti} = \beta_{0i} + \beta_{1i} timec_{ti} + \beta_{2i} timec_{ti}^{\,2} + \beta_{3i} timec_{ti}^{\,3} + \varepsilon_{ti}$$
$$\beta_{0i} = \gamma_{00} + \gamma_{01} cohort_i + u_{0i}$$
$$\beta_{1i} = \gamma_{10} + \gamma_{11} cohort_i + u_{1i}$$
$$\beta_{2i} = \gamma_{20} + \gamma_{21} cohort_i + u_{2i}$$
$$\beta_{3i} = \gamma_{30} + \gamma_{31} cohort_i + u_{3i}$$

All other models are nested within this specification and can be derived through incremental deletion of terms. Given this scope, there are a total of 54 distinct models fit to NLSY97 church

attendance data. They can be organized into five groups, according to the number of random terms they contain:

- Group F - models with fixed effects only
- Group R1 - models with 1 random term
- Group R2 - models with 2 random terms
- Group R3 - models with 3 random terms
- Group R4 – models with 4 random terms

Each model will be referred to by a unique name, which will help in navigating the report document and in composing custom sequences of models. The layout in Figure 4.13 helps understand how each of the models was constructed. The columns of the table indicate the terms added to the first level. Thus, the first column contains the intercept-only models, the second column adds linear term, and third and fourth columns add quadratic and cubic terms respectively. The rows indicate what predictors are added to the second level. Thus, the first row contains the models with no predictor on the second level, the second row adds the predictor *cohort* to the equation of the intercept, the thirds, fourth, and fifth rows row each add the indicated predictor to equations of the linear, quadratic, and cubic terms respectively. The star in the name of the model refers to the five groups defined above: F, R1, R2, R3, and R4.

	β_{0i}	$\beta_{1i} timec_{ti}$	$\beta_{2i} timec_{ti}^{2}$	$\beta_{3i} timec_{ti}^{3}$
	m0*	m1*	m2*	m3*
$\gamma_{01} cohort_i$	m*a	m*b	m*f	m4*
$\gamma_{11} cohort_i$		m*c	m*d	m5*
$\gamma_{21} cohort_i$			m*e	m6*
$\gamma_{31} cohort_i$				m7*

Figure 4.13 Name and structure of the models used in the study.

Figure 4.14 lists the models in group F, arranged to fit the pattern in Figure 4.13, with the full names of the models listed. Not every model from Figure 4.14 will be present in groups with a higher number of random components. For example, although one can add a random component to the quadratic term in the model mFc (even though mFc does not include the quadratic time function), such models were omitted in the current analysis. Such models can be added if they present a particular interest to the modeler. The complete list of models is available in the appendix and is included in the dynamic report. Note that the purpose of Figure 4.14 is to show a *collection* of related models: examination of individual specification may be problematic due to small font size, which is necessary to fit all models in a single view. For closer inspection of individual models, the reader is directed to reports in the appendix, which allows zooming in on selected models.

Figure 4.14 Model specifications of the F group

Representing model solution

Each of the models in the sequence corresponds to an array of quantitative descriptors: coefficients, standard errors, residuals, predicted values, etc. The list is extensive and would vary according to model's class and particular specification. Choosing a set of criteria that makes

sense in the evaluation of each model and yet provides enough common ground for meaningful comparisons may be tricky, and certainly is contingent on the class of models and the research agenda. For illustration purposes, only the basic elements of model solutions were reported for this sequence.

Quantitative descriptives of statistical models used in dynamic reports include:

- Estimates and standard errors of fixed effects
- T-value corresponding to the test of each fixed effect
- Standard deviation of each of the random effects
- Covariance matrix of random effects
- Standard error of residuals
- Deviance, BIC, and AIC
- Predicted value of the model at person by time point resolution

Some of these would be available in every model, while other would not (one can think of them as being equal to zero). Each model is processed and reduced into a single complex graph, composed of four elements: model specification, model solution, model prediction, and model fit. Elemental plots are then assembled into the form shown in Figure 4.15a

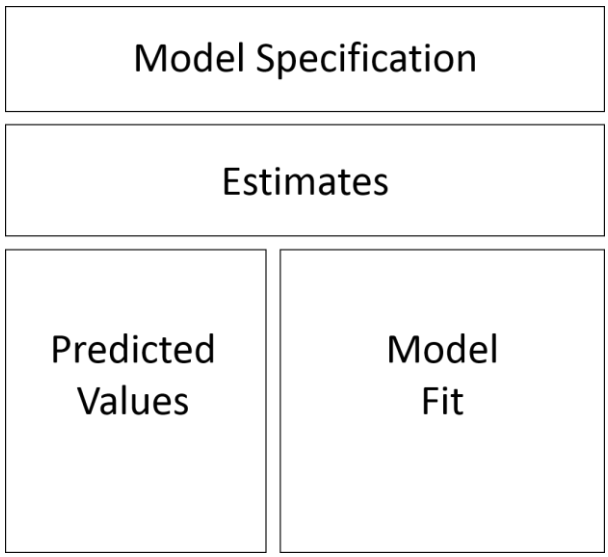


Figure 4.15a Layout of the complex graph describing model solutions

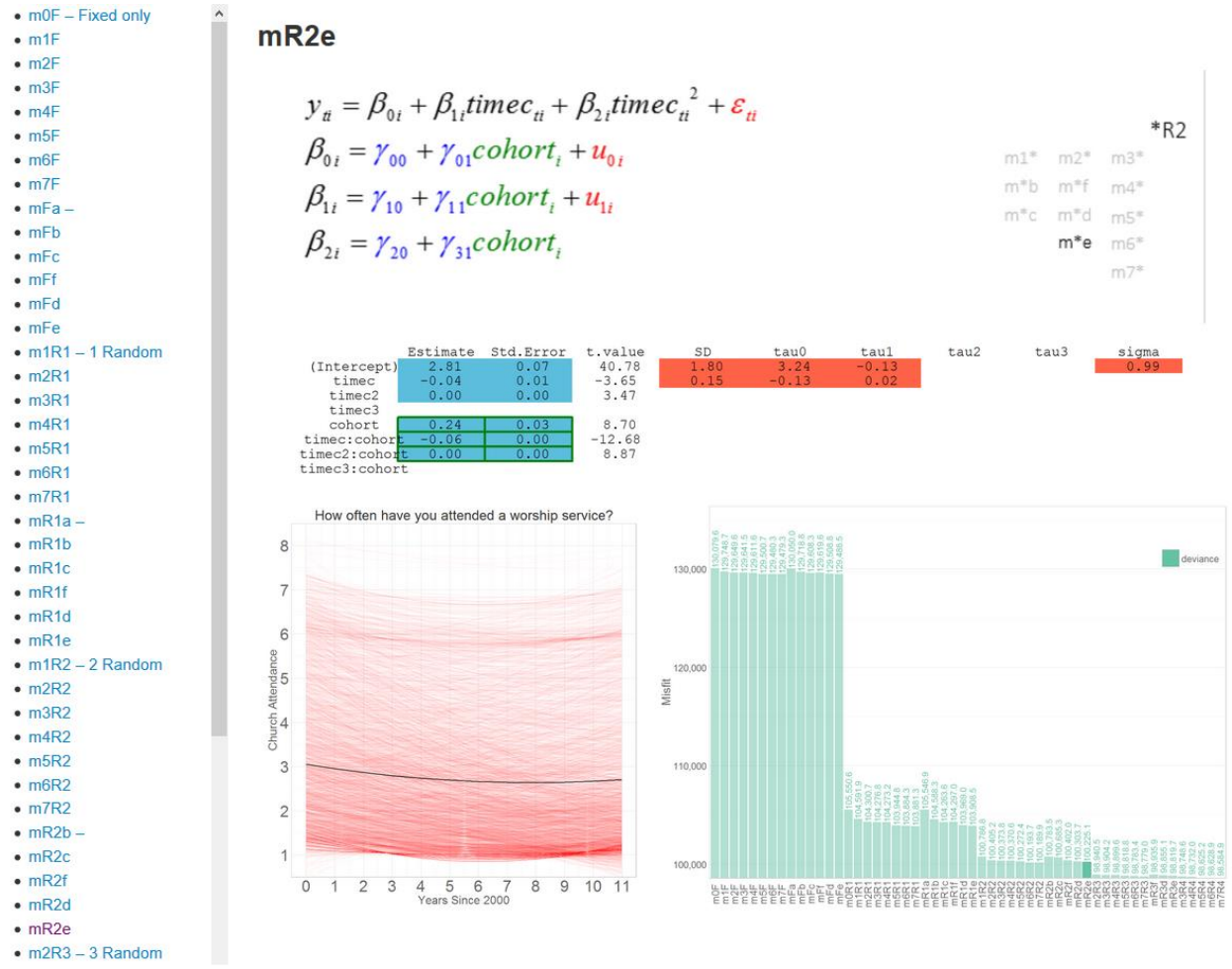


Figure 4.15b Screenshot of the sequence report

Compare Figure 4.15a to Figure 4.15b containing the screenshot of the sequence report. The box with model specification contains current model specification in multilevel notation. In the box below, the estimated values of the coefficients, test statistics, and the standard error of residuals are placed in a grid graph. The cells of this graph are colored to match the colors of the elements in the specification equation. Thus, estimates of the fixed coefficients are marked by blue, and random coefficients are marked by red. This makes for a quicker connection between the specification and the actual values of the coefficients estimated from fitting the model to the data.

The bottom row of the complex graph contains the graph of modeled individual trajectories (left) and the graph of model information indices (right). Predicted trajectories are

represented by semi-transparent red lines. The solid black line gives the common trajectory, estimated from fixed effects. The bar graph contains BIC, AIC, and raw deviance for each model, highlighting the bar with the fit of the current model, in this case mR2e. The graph adjusts the limits of the y scale to accommodate the lowest and the highest values plotted. As immediately obvious, such a graph is not the best for scrutinizing the difference in fit among the models, due to drastically different values of fit indices associated with each group. On the other hand, such a view is very useful in comparing the groups of models: one can immediately see what effect the decision to model a particular level-1 term has on model performance. A separate report contains bar graphs of fit scaled and subset for each modeling group, as demonstrated later.

The interactive table of contents on the left lists the models available in the report – clicking the model name will take you to the results of the corresponding model. Such a layout allows quick retrieval of the individual model solution, without losing track of the context of the sequence. The entire sequence is estimated by a single call in RStudio and printed into the document, containing the processed and organized results for each model. Knitr and rmarkdown packages allow generating dynamic reports in both web (html) and print (pdf) formats from the same source code. The full reports of the sequence is available as an appendix for both metrics of time: rounds of observation and biological age.

Model selection criteria

Although it is tempting to choose a single model fit index as the guide in selecting the "optimal" model, using multiple criteria offers a better perspective on model comparison. The graphs of model fit offer three indices for each model: deviance, AIC, and BIC. Deviance is a −2log likelihood of the misfit function; it is the measure of the total discrepancy between all observed data points and model predictions for them. In confirmatory mode, this quantity is used to carry out significance tests. However, as sample size increases, the value of significance test deteriorates. AIC and BIC represent two adjustments to the raw deviance that account for model complexity and sample size, respectively.

AIC penalizes for model complexity, increasing the value of deviance by 2q, where q is the number of estimated parameters. AIC reflects the difference between implied and observed

models adjusted for parsimony. It has no meaning on its own and must be interpreted in term of the differences among and between the models. A lower AIC is better, including negative values. If a more complex model of the pair has lower AIC, its increase in complexity from the less complex model of the pair is considered justified. The greater the difference in AIC, the more efficient (per degree of freedom) is the model with lower AIC. For each additional parameter to estimate, the deviance must decrease by at least 2, to offset the parsimony penalty.

BIC, in addition to model complexity, also penalizes for sample size. It increases the value of deviance by q*ln(N), where q is the number of estimated parameters and N is the number of data points. BIC is more conservative than AIC, giving greater penalty for model complexity than AIC, and favoring parsimonious models with fewer parameters. A lower BIC is better. If a more complex model of the pair has lower BIC, its increase in complexity from the less complex model of the pair, adjusted for sample size, is considered justified. The greater the difference in BIC between the model in the pair, the more efficient the model (per degree of freedom, accounting for sample size) with lower BIC is.

The deviance will always be lower in models that are more complex, so it is not an ideal criterion for model comparison, but it provides a useful basis for perceiving the levels of misfit among the groups of models. In typical cases, AIC will be larger than the deviance on which it is based, and BIC will be larger than AIC. Graphs that combine all three can point to the location in the sequences where the additions to model complexity become counterproductive. The selection of the "optimal" model from a particular span, therefore, should be based on the behavior of model indices. When a modeling step produces an increase in AIC we have the first hint at the counterproductive increase in complexity; the increase in BIC offers another, taking sample size into consideration. The full reports on model performance is available in the appendix for both metrics of time: rounds of observation and biological age.

Model analysis and synthesis

As was demonstrated in the first section of this chapter, racial groups exhibit substantial heterogeneity in church attendance, both in cross-sectional and longitudinal views. In light of

this, the current demonstration will use the data only from respondents, who identified themselves as White and provided the response on the focal variable at every time point.

The estimation of the models in the F group generated fit statistics shown in Figure 4.16. Notice that the order of bars in the graph corresponds to the order laid out in Figure 4.14 if the elements are read sequentially by **rows**, starting with the top left position. The model m0F (the first bar in the graph) gives the reference point for relative improvements of fit with each added term. Adding a linear function of time results in a substantial reduction of misfit, as would be expected from data that have a heavy longitudinal structure. The curvature of the quadratic term in m2F further reduces the misfit, however the cubic term in m3F shows only a slight reduction in the absolute deviance and a slight increase in BIC, which penalized model complexity. The next four bars (mFa, mFb, mFf, and m4F) correspond to the same progression of models , but with the predictor *cohort* entered into the second level equation of the intercept. The reduction in misfit follows a similar pattern: substantial drop after the linear term is added,

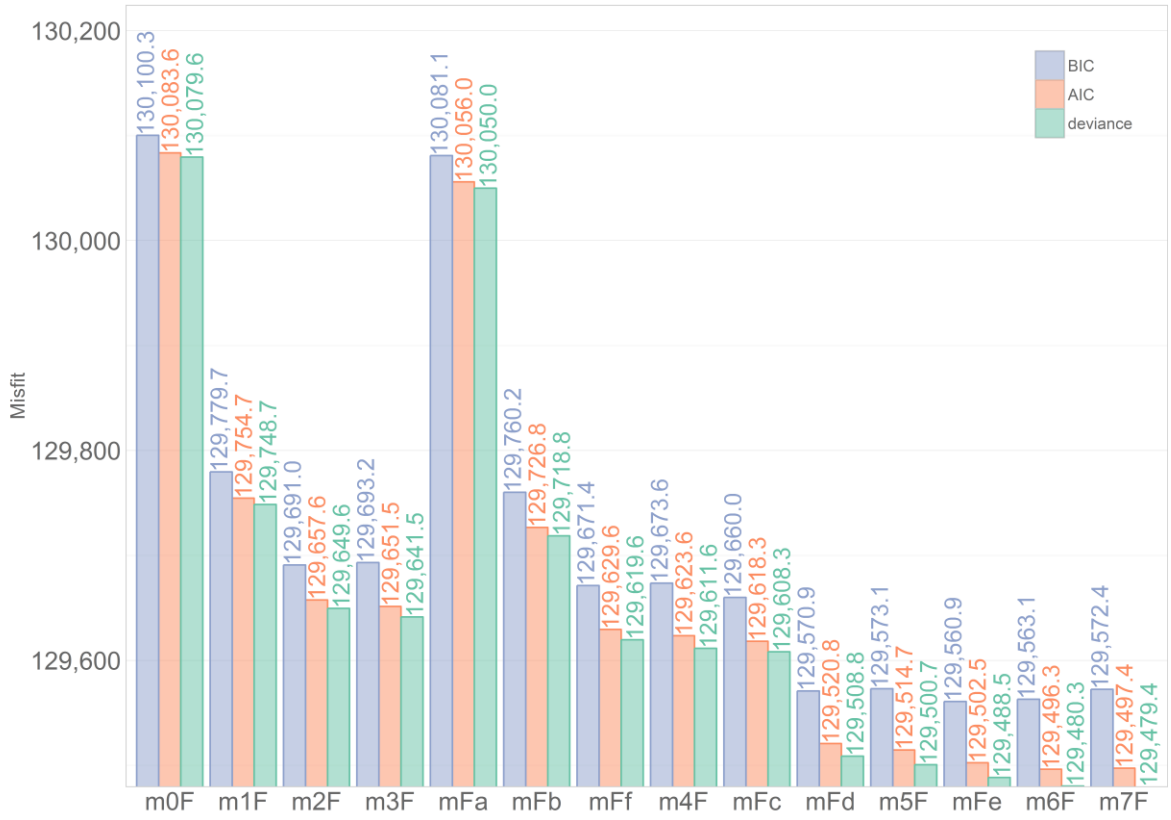


Figure 4.16 Fit of models in the F group: view by rows in the group specification.

noticable decrease following the introduction of the quadratic term, and a similar reaction to the cubic term: minor decrease in deviance and AIC with a minor increase in BIC.

The next bar corresponds to the model at the beginning of the third row of the model group specification. This model (mFc) is not nested with m4F, but offers an interesting observation: a better fit can be achieved by extending the common ancestor mFb with a predictor *cohort* to the linear term, than by adding quadratic and cubic terms to the first level equation. The sizable decrease of misfit in the next bar (mFf), however, indicates that quadratic term is much more valuable if enhanced by the presence of the predictor *cohort* in the equations of the first two time effects.

As is apparent, the interplay between first level and second level predictors can be more conveniently explored by organizing the bars in a different order. Figure 4.17 arranges the fit bars in the order layed out in Figure 4.14 if the elements are read sequentially by columns, starting with the top left position. Such arrangement allows looking at the effect of adding second level predictors among the models with the same number of time effects.

Comparing m0F with the adjacent mFa shows that using the age difference does not improve the model much, which is not surprising given the longitudinal structure of the data. The next three bars show the decrease in model misfit when the predictor cohort is added to the model with a linear term. The drop from mFb to mFc reiterates the finding gleaned from Figure 4.16: *cohort* improves the model substantially when added to the equation of the linear term and makes the quadratic term much more valuable (mFd). The next two bars (m2F and mFf) demonstrate that bare quadratic and cubic terms cannot compensate for the absence of cohort in the second level, even when it is entered into the intercept equation (mFf). Adding cohort to the equation of the quadratic term (mFe) further decreases both absolute and adjusted fit of the model, although not as drastically as adding cohort to the equation of the linear term in mFd.

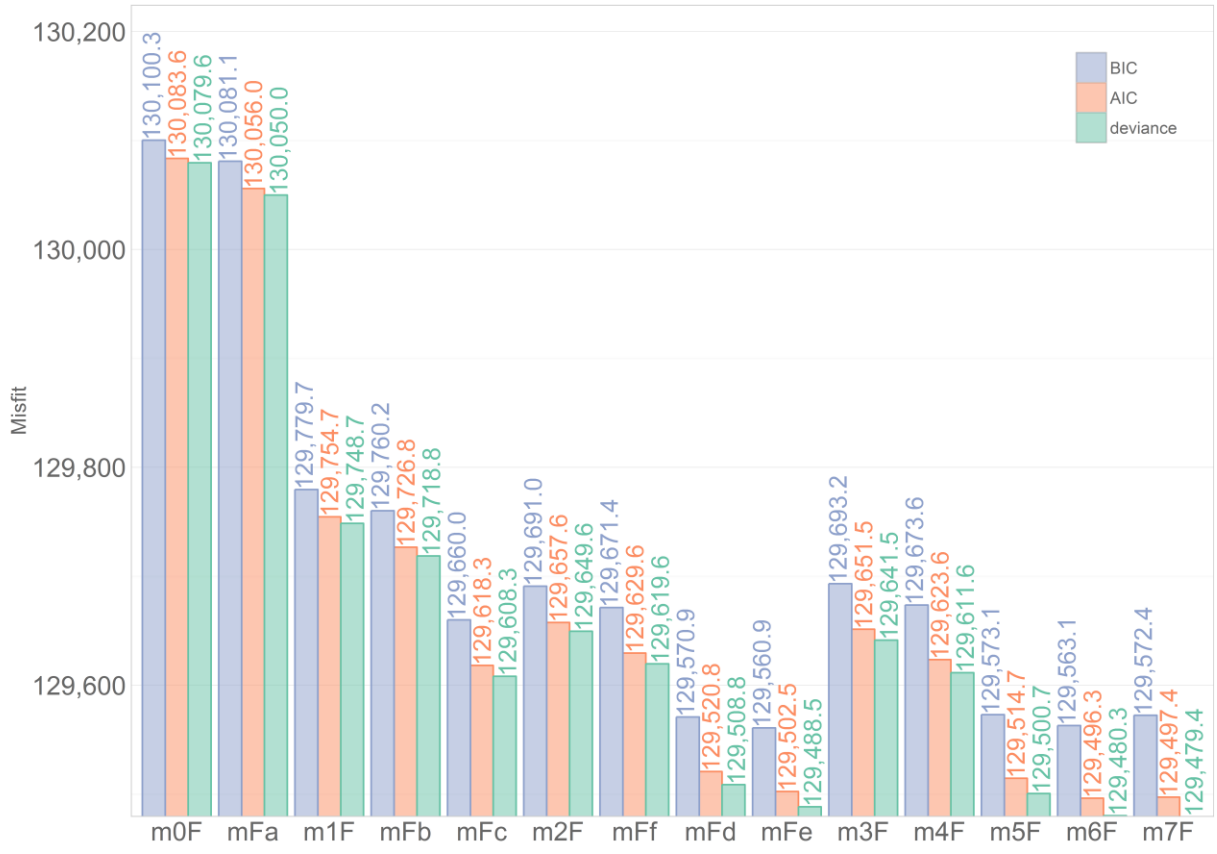


Figure 4.17 Fit of models in the F group: view by columns in the group specification.

The last column of the group F specification starts with m3F. The presence of the cubic term at the first level cannot compensate for the absence of cohort at the second level of the model equation. Looking at the change in misfit from m4F to m5F, we once again recognize the importance of having *cohort* as the predictor of the linear term. Minor fit reduction is associated with using cohort to predict quadratic term (m6F), but adding it to the cubic term (m7F) begins to increase AIC and BIC, indicating that the gains in misfit reduction are not justified by the increase of model complexity.

The key finding from this brief analysis is the exceptional role that *cohort* plays in predicting the linear term. Only after it is added to the linear term does the model fully capitalize on the presence of quadratic and cubic terms at the first level. Model complexity

beyond mFe (see fig. 4.16) however does not offer improvements in adjusted fit, indicating that after cohort enters the model for the first three time effects, modeling additional curvature may be unnecessary.

Fit graphs show how models *perform*, but they do not describe what the models *are* in terms of the numeric solution of their coefficients. To examine how changes in model specification affect the model coefficients and the reproduced patterns of data, we turn to the interactive feature of the model sequence report. Open appendix containing the report (I recommend Firefox browser for a more consistent performance). The top of the document contains the definition of the data used in the model, expressed through dplyr syntax.

The null model (m0F) estimated the grand mean to be 2.80 with standard error of 0.01, resulting in the residual of 1.96 (standard deviation). The graphical interpretation of this model is a straight line passing through the y coordinate 2.80, depicted in the predicted value plots in the bottom right corner. We interpret the intercept as the grand mean of church attendance over all time points and individuals. Although the estimated value does not have a direct quantitative interpretation, referring back to Figure 4.4 we see that value 3 on the scale with which church attendance was measured corresponds to response category “Less than once a month”, while value of 2 corresponds to “Once or twice a year”.

The table of contents on the right lists the models available for viewing. Instead of scrolling down to see the results of the next model, click on the corresponding TOC item starting with m1F. The graph of the predicted values reflects the changes in the model: the new slope of the predicted trajectory is estimated to be -0.06 and the intercept increases to 3.12. The time variable on the x-axis is centered at 2000, thus the slope can be interpreted as the average change in church attendance for every additional year past 2000. Clicking through models m2F, m3F walks us through the models until the predictor is introduced in the second level. The curvature added by the quadratic term in m2F convex the line and moves the intercept even higher to 3.32 and increases the magnitude of the slope to 0.17. The cubic term, although invisible at the two decimal point resolution, is clearly visible in the shape of the line, continuing

the trend of the quadratic term: intercept increase to 3.27, the slope accrues magnitude and becomes 0.24. The quadratic term also increases from 0.01 in m2F to 0.03.

Model m4F adds cohort as predictor of the intercept. For every year difference in age, younger respondents are expected to have the intercept higher by 0.04 from the grand mean of 3.27. This indicates that respondents are expected to attend church more often the younger they are, which concurs with the cross-sectional descriptives. Addition of this term also introduces the thin red lines in the predicted value plots. Each of them represents the conditional prediction for every value of *cohort*. The difference among the five birth cohorts become more visible after moving to m5F. The intercepts for each cohort spread out on the y-axis, with higher values for younger cohorts. Now for every year in age difference, the younger respondents are expected to have the intercept higher by 0.18 from the grand mean of 3.00. This indicates that younger respondents have higher attendance at the beginning of their trajectories.

Models m6F and m7F continue the trend of increasing the difference among the cohorts: the intercepts for *cohort* changes from 0.18 in m5F to 0.24 and 0.25 respectively. The increases of *cohort* coefficients in the equations of linear and quadratic terms also help to spread out the predicted lines. The coefficient associated with predicting the linear term from cohort membership γ_{11} changes from -0.02 in m5F to -0.06 and -0.08 in m6F and m7F respectively. This indicates that younger respondents undergo more change than older respondents, which is congruent with their higher initial attendance.

The path from m0F to m7F that we just walked is one of the many that can be found in the current span of the modeling space. The choice of the path is arbitrary and can be changed. Consulting the layout in Figure 4.14 we may choose a different path (e.g. m0F -> mFa -> mFb -> mFf -> mFd -> mFe -> m6F -> m7F), which might better suit our analytical interests or demonstration purposes. For example, to explore the role of cohort in models with three time effects we can use the following series of steps: m2F -> mFf -> mFd -> mFe, which takes us down the third column in group F specification from Figure 4.14.

Instead of examining each model in this sequence (m2F -> mFf -> mFd -> mFe) and using your mouse to navigate between the models, click through the entire sequence first. Now using keyboard keys (Alt + arrow key) traverse this sequence back and forth. This frees up the resource of attention it takes to identify the model name in the table of contents. With the sequence loaded, focus on the t statistics in the grid graph. Notice how t values of the time effects change with the introduction of *cohort* into the second level of model equation. This indicates that the predictor *cohort* absorbs the variability, "stealing" it from the terms in the first level. With your keyboard buttons go to the furthest point in this sequence (mFe) and extend the path by clicking m6F and m7F. We see that m6F regains some of the "stolen" significance from the first level term, but the cohort predictor in m7F reverses this, dropping all of the time effects except for the intercept below significance level.

Naturally, observations like these could have been drawn from scrutinizing the model outputs provided by the estimation software. However, it would be probably not be the most efficient use of the analyst's attentional resources.

Other custom sequences

Sequences similar to those demonstrated with fixed effects models can be replicated with their counterparts from random effect groups, but due to space limitation they will not be narrated here. Instead, to further demonstrate the utility of the interactive model sequence I give a brief example of comparing models from different model groups.

Using m5* as the model present in all five groups we define the sequence as m5F -> m5R1 -> m5R2 -> m5R3 ->m5R4. Click through the models to load the sequence. Alternating the views between the first pair with the keyboard keys we see how modeling the intercept term as random changes the t-values for each term. Time:cohort interaction changes its t-value from -10.58 to -18.17, indicating an increased importance of this term in the model of this configuration. This value, however, goes down as we add more random terms: -10.01, -9.12, and -10.34 in m5R2, m5R3, and m5R4, respectively.

The clear advantage of using m5R1 over m5F is evident in the drastic reduction of the residual from 1.94 to 1.13. It progressively decreases, arriving at 0.90 in m5R4. The residual is

not the only reason to prefer m5R4 over its counterparts. Observing the changes in the covariance matrix of the random effects while walking this sequence, we can see that m5R4 has the lowest variances of the random terms. In addition, the covariance between the intercept and linear terms, reaching as high as -0.37 in m5R3, has the lowest value in m5R4 of any other preceding model. Although we can potentially interpret such covariance, a matrix with lower covariances offers simpler and more straightforward interpretation.

Extending the sequence to include m6R4 and m7R4 we see that entering cohort into the equation of the quadratic and cubic terms does not offer us any reduction in the residual variance. However, when examining model performance in Figure 4.18, m7R4 does come on top, with the lowest absolute and adjusted fit.

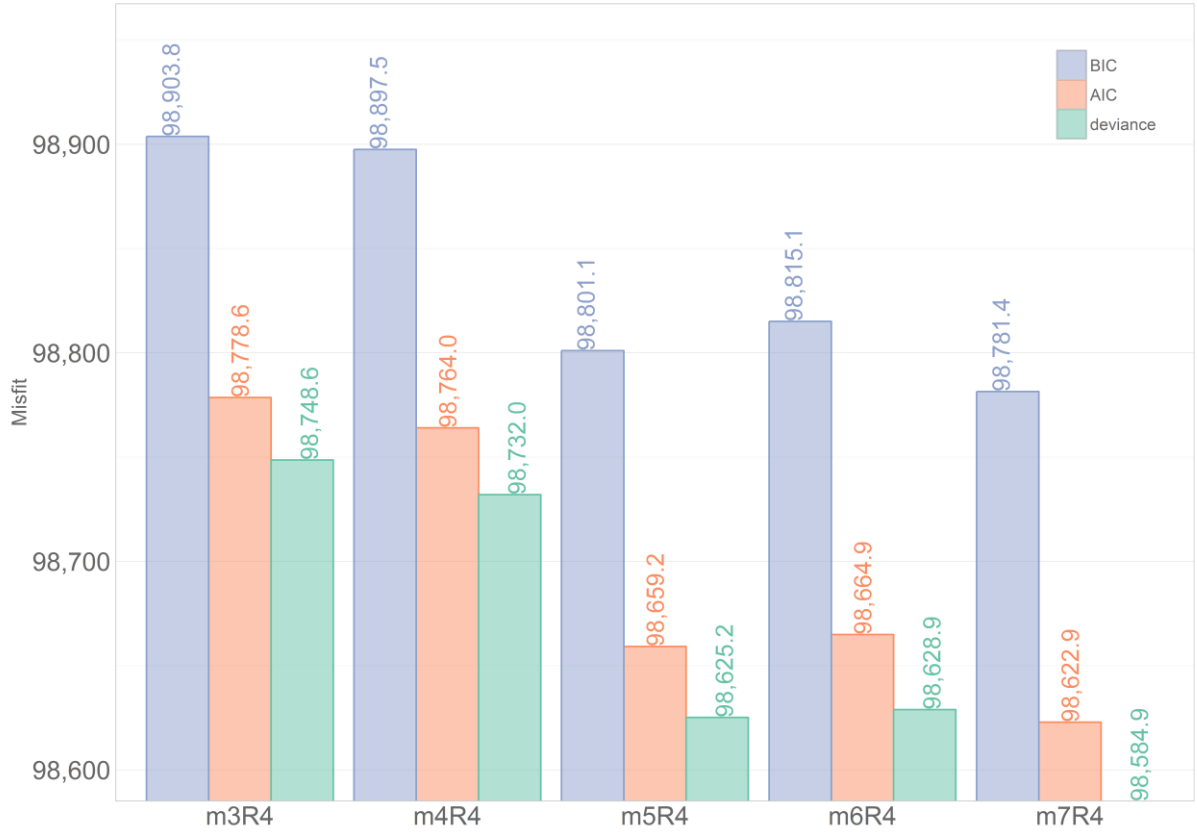


Figure 4.18 Fit of models in group R4

Changing the metric of time

The models above use cohort as the second level predictor that accounts for age differences among individuals. Another way to explore the role of age in defining the trajectory of church attendance is to change the metric of time from the wave of observation to the biological age of respondents. Due to the dynamic nature of the reports, this can easily be accomplished by changing the definitions of the variables timec, timec2, and timec3: timec was previously computed as *timec* = *year* – 2000, now we use *timec* = *age* – 16, centering it about 16 years of age. Figure 4.19 reflects this change. The x-axis of the graph of predicted trajectories now counts the number of years past the age of 16. Compare Figure 4.19 to Figure 1.2, where the same model was estimated using wave of measurement as the metric of time. The interactive report is included as a separate appendix.

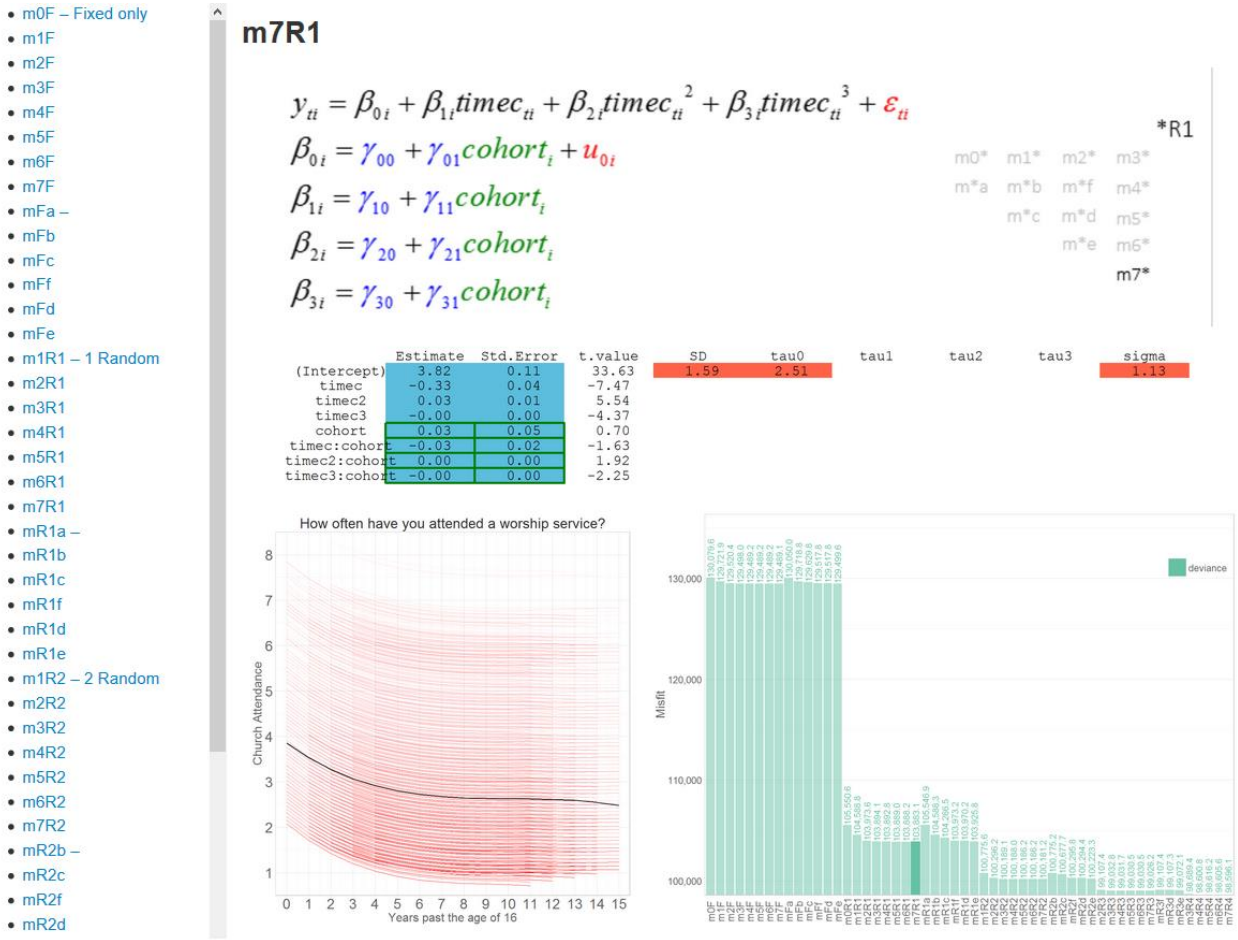


Figure 4.19 Screen shot of model sequencer with age as the metric of time

Changing the metric of time validates the key role that biological age plays in defining the trajectory of church attendance. When entered as the predictor on the first level, age explains the trajectories better, when equivalent specifications with different metrics are compared side by side. This makes sense, because age on the first level contains some of the age difference previously entered as a level-2 predictor. We can observe in Figure 4.20 that adding *cohort* to the second level does not result in relative misfit decrease: AIC and BIC begin increasing when *cohort* is added. Similar behavior of AIC and BIC is observed when time effects are modeled as random.

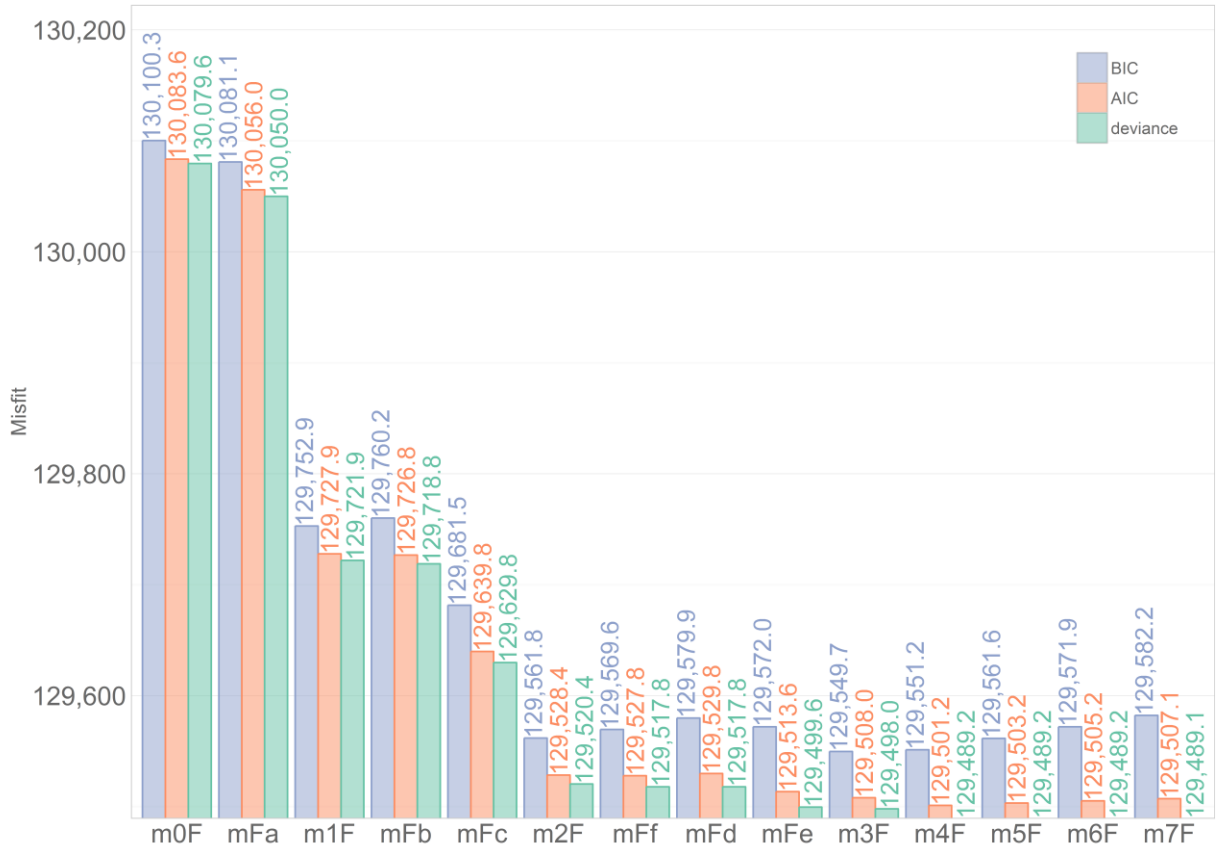


Figure 4.20 Fit F group models: view by columns in the group specification. Time metric: age

Conclusions

Is m7R4 the “winning” model? When time metric is the wave of measurement fit indices indicate that the answer is “yes.” However, the value of such a modeling exercise will be curtailed by simply reporting and interpreting the best fitting model. Frequently, a fuller picture emerges from studying how the model reacts to the introduction or removal of particular terms. As was demonstrated, using cohort to predict linear and quadratic time effects results in substantial misfit reduction, especially in models with random effects. This implies that individual age differences are the key factor in explaining the observed trajectories of church attendance. This conclusion is verified by the change of time metric.

The power of the demonstrated method for examining statistical models proves to be useful both in searching for the optimal model to report and in gleaning a deeper understanding of the studied phenomenon. Focusing on the behavior of models, as terms are being added to or removed from them, offers an opportunity to explore various scenarios of model development that may not be evident at the beginning of the analysis and to test hypotheses about the role of individual predictors that emerge from ongoing analysis. Most importantly, reporting the entire span of models, as opposed to a few with the highest fit, invites the reader to participate in the analysis and delivers a richer opportunity for insight.

CHAPTER V

DISCUSSION

In this section, I will discuss several interpretational and summarizing topics. First, I will review what has been learned through application of the graphical modeling method from the analysis of the NLSY97 religious data. Next, I'll discuss how the graphical modeling method can be used in broader settings, by various audiences and to various ends. Following, I will review weaknesses and limitations of the current research, and indicate possible future directions.

Dynamics of church attendance

The respondents in the NLSY97 survey demonstrated that the dynamics of the frequency of church attendance heavily depends on age. As respondents grew older, they generally attended church less. The common trajectory for this pattern is captured with predicted values plots from model m7F. Figure 5.1 shows the frequency of responses to church attendance item of NLSY97 at the first and last rounds of observation, and the common trajectory line from model m3F that describes the change between these two time points.

Figure 5.1 General trajectory of change in church attendance among Whites.

The sample from NLSY97 included five birth cohorts and, as latent curve models indicated, the age difference was a significant factor in modeling the dynamics of church attendance. Figure 5.2 shows predicted lines, when age difference was used to predict how time affects the modeled trajectory. Each red line gives the predicted trajectory for respondents in the same birth cohort. Younger cohorts had steeper declines, while older cohorts tended to have trajectories with smaller slopes and curvatures. An important caveat relates to the age period during which the observations were taken. As Figure 4.10 indicated, church attendance tended to become relatively stable as subjects reach the age of about 20-21. Had data not included observations from respondents before that age, the effect of age on church attendance would have been much harder to detect.

Figure 5.2 Common trajectory lines for each cohort.

Although inclusion of age differences helped in explaining the observed dynamics, respondents in each cohort were far from homogenous in their trajectories of church attendance. The bimodality of the response distribution, observed during cross-sectional analysis, was also evident in the longitudinal perspective. After the inclusion of random effects, which models individual trajectories, it became evident that particular types of trajectories were especially common. As seen in Figure 5.3, which reproduces the predicted value graph from model m3R4, a consistently low rate of attendance (or not attending church at all) throughout the rounds of observation is the most salient cluster of individual trajectories. Another easily

detectable cluster of trajectories is tracing a regular attendance, at the level of around 6 on the outcome scale, which corresponds to attending church at least once a week.

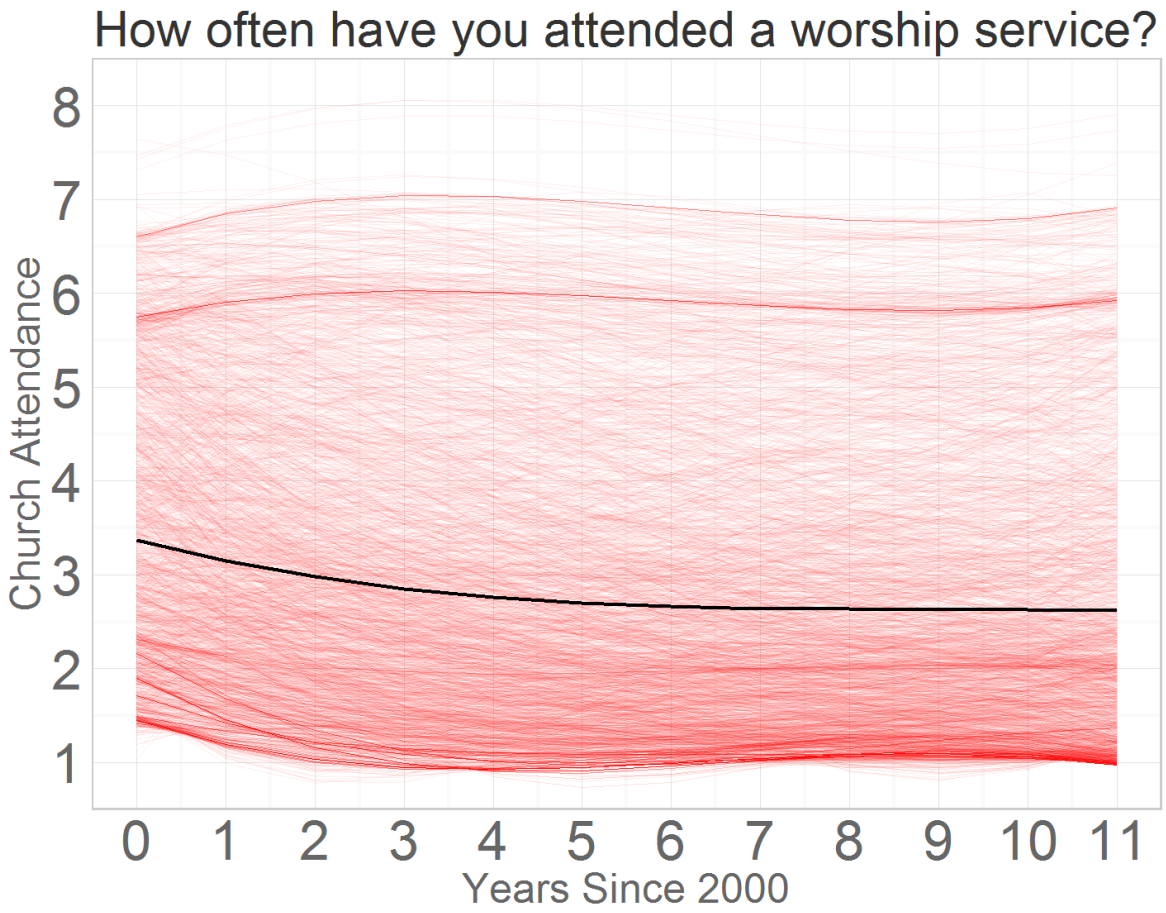


Figure 5.3 Predicted value plots of church attendance in model m3R4.

Further evidence of the importance of age differences in explaining individual trajectories can be found by juxtaposing the predicted value plots from models m3R4 and m7R4, given in Figure 5.4. In m3R4, while allowing every time effect to vary across individuals, no age difference was entered into predicting the effect of time functions. This forced the predicted trajectories into pronounced clusters that defined typical dynamics, disregarding the age differences. In m7R4, however, we can see that, while preserving recognizable clusters, trajectories become more evenly distributed on the graph canvas, implying that accounting for age differences permits more accurate representation of individual trajectories.

$$y_{ti} = \beta_{0i} + \beta_{1i}timec_t + \beta_{2i}timec^2_t + \beta_{3i}timec^3_t + \varepsilon_{ti}$$

$$\beta_{0i} = \gamma_{00} + u_{0i}$$
$$\beta_{1i} = \gamma_{10} + u_{1i}$$
$$\beta_{2i} = \gamma_{20} + u_{2i}$$
$$\beta_{3i} = \gamma_{30} + u_{3i}$$

$$y_{ti} = \beta_{0i} + \beta_{1i}timec_t + \beta_{2i}timec^2_t + \beta_{3i}timec^3_t + \varepsilon_{ti}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}cohort_i + u_{0i}$$
$$\beta_{1i} = \gamma_{10} + \gamma_{11}cohort_i + u_{1i}$$
$$\beta_{2i} = \gamma_{20} + \gamma_{21}cohort_i + u_{2i}$$
$$\beta_{3i} = \gamma_{30} + \gamma_{31}cohort_i + u_{3i}$$

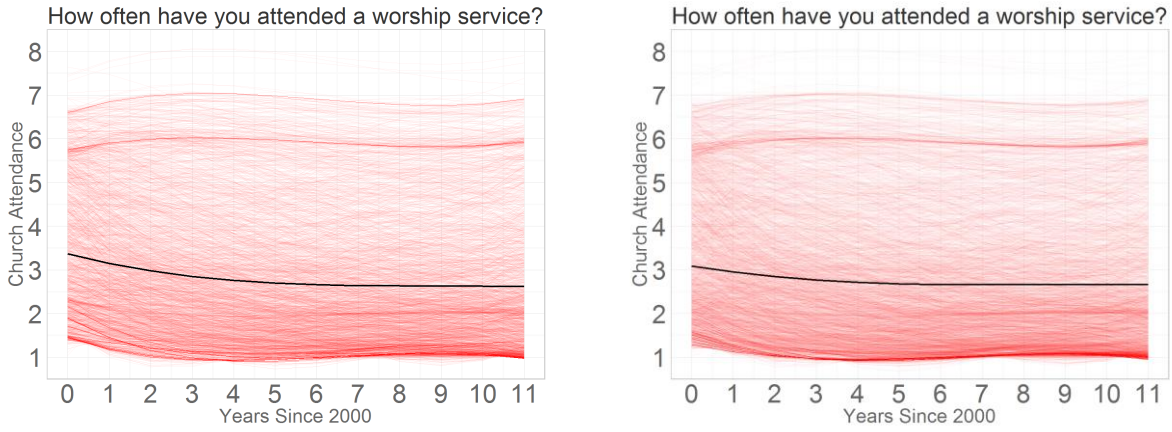


Figure 5.4 Effect of age difference on modeled individual trajectories.

The last example allows illustrating what insights can be achieved from using an interactive, graphical model report, unlikely otherwise. Consider sequence m4R4 -> m5R4 -> m6R4 -> m7R4. All models result in residual variance, indistinguishable under the second decimal point precision. The bar graph of model performance indicates increasing fit, however, more complex models are expected to fit better, and the large sample size might complicate the detection of statistical significance of the added terms. The changing values of the coefficients are useful, but do not describe how the model recreates the data. However, by focusing on the graph of the predicted values while walking the sequence we can observe the direct effect of the added terms on recreating individual trajectories: using age difference to predict the linear term results in the most visible changes in the modeled trajectories.

Although it was demonstrated that changing the metric of time to biological age helps seeing the patterns in the dynamics of church attendance clearer, the former metric should not be discarded as inferior. When waves of measurement are used as the metric of time, we can express the influence of age through gamma estimates, which may be preferable depending on

the particular research question at hand. Another advantage of using rounds of observation as the metric of time is a better view of the period effect, concealed by repositioned trajectories when age is used.

Uses and Applications

Although rooted in empirical research, the focus of this dissertation is on the design of a dynamic reporting method for statistical modeling in general and latent curve models in particular. The rise of complexity in statistical modeling, discussed in the first chapter, pushes practitioners towards adopting more specialized software solutions that could offset the increasing complexity of modeling projects. Although at least some of the model analysis and synthesis demonstrated in this dissertation could have been accomplished with traditional means of reporting, automating the most taxing tasks involved in modeling frees up the attentional resources of the analyst, affording more cognitive energy to be spent on perceiving and interpreting the difference among the models. The uses of the interactive model reporting, demonstrated in this dissertation, can suit various audiences of researchers, depending on their particular needs.

Analysis and Synthesis

One of the most obvious use of the interactive dynamic reports demonstrated in this dissertation is the organization of the modeling workflow that permits maximum flexibility of model development within a defined span of models. Although in many cases, the analysis of data is preceded by formulating specific research hypotheses to be tested with appropriate statistical operationalizions, rarely are they specific enough to correspond to a particular sequence of model specifications. For example, while a hypothesis "Age is a significant predictor of church attendance trajectory" is well formulated for human understanding, it can describe an entire array of models, each answering the question with the precision defined by its particular specification. Which specification should be chosen to be tested? Both "*Age difference has a significant effect on predicting the quadratic time function of the trajectory of church attendance when accounting for individual differences in the intercept and linear time function*" and "*Age difference has a significant effect on predicting the cubic time function of the trajectory of church*

attendance when accounting for individual differences in the intercept, linear, and quadratic time functions" seem appropriate, among many other possible models.

When model reporting is costly in terms of organizing the estimation and reviewing the output, researchers are discouraged from testing many relevant specifications. When, however, a span of models is specified, estimated, and reported at a relatively low cost, the analyst may evaluate the hypotheses s/he did not anticipate due to their high specificity. Walking through various paths, which connect the extremes of model complexity within the span, offers opportunities for insight that might not be as readily available with traditional model reports. Instead of organizing the workflow around estimation and analysis of individual models, this dissertation demonstrates the advantage of higher order units of analysis – model sequences, collections of models, joined in a deliberate and meaningful order to provide custom views on the modeling space.

Reproduction

The reports presented and discussed here are fully reproducible and can be downloaded and adapted for personal use from the GitHub hosting service. This functionality offers the interested researcher a good starting point if models need to be explored in greater detail or if the defined span does not include a specific model. For example, the demonstrated report can be reproduced using a modified lambda matrix, which instead of polynomial time functions encodes piecewise shapes or some other exotic time functions.

Another powerful feature of the provided templates is their custom-made wrappers (also known as adapters) to the popular estimation packages. An adapter creates an environment in which imported objects are transformed to fit common, specially designed syntax of interaction. The wrapper in this work was optimized for model reporting, rather model estimation. For the span of models used in the present work, *nlme* and *gls* packages of R proved sufficient to carry out all estimations. Other packages can also be added without disturbing the integrity of the code that compiles the dynamic report.

Given that modeling is almost always a cyclical process, with analyses and estimations replicated once insights are drawn from the previous run, having the ability to reproduce the

report with a slightly different input can be a powerful factor in reaching deeper understanding of the studied phenomenon. For example, the demonstrated report was compiled using the models that relied on the data provided only by respondents who identified themselves as White and who had no missing observations on the outcome. This decision was informed by the obvious heterogeneity among the racial groups in the patterns of their church attendance. Would the conclusions hold if another racial group were chosen as focal, or if analyses were conducted across races? Replicating the report for different race categories is as easy as changing the one character in the code that selects the appropriate observations from the data.

Alternatively, one may wish to enter race as a predictor into the model to quantify its effect on the model's solution and performance . This also can be easily achieved by adding relevant model formulas to the provided slot in the code (see "./Models/LCM/LCModels.R" file at the GitHub repository). The general form and the interactive functionality of the report will not be affected by such modifications. Naturally, the graph's contents and aesthetics are highly customizable, powered by *ggplot2* syntax.

Communication

Simplifying scientific communication serves an important role. Although the practice of reporting the "winning" model has its fervent adherents, there are also plenty of researchers who insist on reporting several models, each giving a slightly idiosyncratic take on recreating the data and explaining its patterns. The insights gathered from such multiple winners are argued to be superior over those relying on a single winner. However, similar to the controversy surrounding NHST, the debates surrounding the "legality" of reporting a single model as the winner is misdirected: The issue is not whether a tool should exist, but rather how and to what end it should be used. Reporting a single model gives a tangible, concrete, and relatively simple statement about how the world works, without the vagueness of stipulations and contingencies that would invariably accompany any modeling exercise and research projects, but which may not be particularly useful in communicating the nuances and breadth of one's research findings to the rest of the research community.

On the other hand, omitting the "failed" models, while beneficial to the simplicity of communication, is detrimental to the thoroughness and transparency of the published results. Scientific publications are meant not only to preserve the conclusions of the study, but also to recreate the path of reasoning that led the researcher to the particular conclusion. To verify the validity of a study's conclusions frequently requires accessing the models typically omitted from the reports due to either space limitation of the publication medium or simply because the researcher finds them "uninteresting." Reporting all the models fitted in the study and pointing to the sequences that provide the basis for the reported conclusions avoids a version of the ubiquitous "file drawer" problem in academic research, in which only significant findings are reported to the research community and nonsignificant ones are stored unseen in the filing cabinet. Knowing what doesn't work sometimes can be just as important as knowing what does.

Once a report is generated within the current system, no further programming intervention is necessary to employ its analytical utility. Moreover, each interactive model report is created as a self-standing webpage (or a PDF document) that can be shared through data storage device or deployed to a remote website. This feature makes the results highly sharable, allowing a wider audience to be involved in its analysis and interpretation. In a research team, where programming is sometimes accomplished by a dedicated member, the adaptation of dynamic and interactive reports such as one demonstrated here offers a useful specialization of labor: production and analysis can be easily divided among members with different skillsets.

Finally, such reports can also be used to present the research findings to a live audience. After studying, analyzing, and synthesizing the models, a presenter may "record" a particular sequence that demonstrates his/her point, and have the entirety of supplemental materials readily available to answer follow-up questions and raised concerns. In this sense, the method becomes truly dynamic, and new results could be "discovered" by an engaged audience interacting with the system.

Limitations and Future Directions

I will discuss limitations of the analysis of the NLSY97 religious attendance data, and also limitations and future directions within the system itself. Each will be treated separately.

Obviously, only a few "pathways through the data" have been presented in the current document. The sequences of models from group F (fixed effects only) that were demonstrated in detail could be repeated for group R1 (random intercept) to verify the finding from their counterparts in group F. Some of the models have corresponding specification in other groups: for example m3* has an expression in all four. By progressing through the sequence m3F -> m3R1-> m3R2 -> m3R3 -> m3R4 we can see how accounting for individual differences in the intercept, linear, quadratic, and cubic terms affects predicted trajectories in that specific type of model (m3*).

Another possible exploration capitalizes on the location of certain models in the span. For example, mR2f is one modeling step away from 6 other models. Loading and progressing through the sequence mR2f -> m2R2 -> mR2f -> mR2b -> mR2f -> mRd -> mR2f -> m4R2 -> mR2f -> mR3f -> mR2f -> mR1f allows identifying promising directions of development for mR2f and deciding whether and how this model should be reduced or extended. Alternatively, we can pick any model in the span as the starting point and explore what developments offer the greatest insights.

As far as the specific results presented here, the clusters found in the individual trajectories recreated by models' predictions suggest that while LCM is an effective method to describe the overall trend, the data may contain latent classes that should be accounted for. Although the general trend of the observed individual trajectories indicates a decrease in church attendance, it is but an average. It makes sense to assume that individuals increase their attendance as they age, but their contribution to the general trend may be concealed by the weight of the majority. Other applications of the graphical modeling tool developed here would undoubtedly tell slightly (or even substantially) different stories.

Another limitation of the statistical analysis has to do with the scale on which church attendance was measured. Originally ordinal, it was transformed and treated as continuous for the purpose of fitting latent curve models. A more precise account of variability can be achieved by the models adapted to categorical data, such as survival analysis or Markov chains. Applying

growth mixture models and Markov chains to the same data may validate the findings of LCM and offer new insights into the dynamics of church attendance.

To this end, future directions of developing dynamic model sequence reports include accommodation of other modeling methods, capable of working with data on different scales of measurement and operationalizing different type of research theories. Developing new wrappers for R packages responsible for their estimation promises to extend the use the demonstrated graphical methods of model synthesis to the wider audience of researchers.

Any new system will of course have weaknesses, which require development over time, informed by the experience of users of the system. One challenge will be moving this work into the hands of researchers. The latest version of R package *rmarkdown* allows uniting multiple reports and deploying the project as a static website (see examples at http://rmarkdown.rstudio.com) using Jekyll site generator. This immediately places the research results within reach of a wide audience, but may be challenging to implement without certain programming skills. GitHub, from which the project can be downloaded for reproduction, while a powerful social coding platform, may also present somewhat of a learning curve for uninitiated users. However, given the rising popularity of R and RStudio in the research community and the creative momentum of the RStudio team, who continues to develop both functionality and the user interface of the latter, it is reasonable to expect that the skills needed to implement the system demonstrated in this dissertation will penetrate a continuingly wider audience of researchers in the immediate future. Another challenge will be coordinating and compiling fixes and improvements. However, like the internet, this kind of system does not necessarily require a single -- or even a few -- organizers. The system can ideally become a dynamical and changing method itself, as new models are developed within the context illustrated here.

Conclusions

In conclusion, I would like to revisit the metaphor that opened this work. Given the crucial role *seeing* plays in human cognitive process, the ability to visually inspect the object of analysis cannot be underestimated. This dissertation made three important advancements to that end, relating to modeling workflow. First, I offered an example of how all crucial

components of a model (specification, solution, fit) could be represented visually in a single graphical object. Second, I provided a technology for organizing collections of models and a system for navigating among them that outsources detection of differences to visual processing, reducing cognitive strain and freeing attentional resources. Third, I demonstrated how the these two advances can be used creatively to compile custom sequences – perhaps even unforeseen initially – empowering both the analyst in testing custom hypotheses and the target audience of the report in joining in the analysis and interpretation. When looking down the microscope, a researcher does not notice its lens or pay attention to its focus knob, but attends to the objects it magnifies. In the same way, the methods and techniques of model comparison presented here place the attentional focus where it is due: the model and the data.

As a forest can be hidden behind the trees, so a model can hide behind its estimates. While statistical modeling at the individual model level is valuable in its own right and forms the foundation of many projects, *model synthesis* offers a far richer opportunity for insight and discovery. Bringing many models together to define, illuminate, and critique each other creates a frame of reference in which the meaning of individual models become clearer and richer, offering something greater than an isolated analysis or a comparison to only a few rivals could. The word “rivals” in describing the models competing to be chosen as a “winner,“ deemphasizes the importance of “losing” models to enhance our understanding of the modeled phenomenon, and perhaps should be used with discretion. Understandably, model synthesis is a more involved process with challenges distinct from those in model analysis. This dissertation offered a set of tools and examples to popularize this practice.

All models are wrong, but *any* of them can be made useful.

REFERENCES

Arnett, J. J. (1995). Adolescents' uses of media for self-socialization. *Journal of Youth and Adolescence, 24*(5), 519-533.

Astin, A. W., & Astin, H. S. (2003). Spirituality in College Students: Preliminary Findings from a National Study.

Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development*, 145-152.

Bengtson, V. L., Copen, C. E., Putney, N. M., & Silverstein, M. (2009). A longitudinal study of the intergenerational transmission of religion. *International Sociology, 24*(3), 325-345.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588-606.

Bock, R. (1989). *Multilevel analysis of educational data.* Location: Academic Press.

Boker, S. M., & Graham, J. (1998). A dynamical systems analysis of adolescent substance abuse. *Multivariate Behavioral Research, 33*(4), 479-507.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research, 37*(1), 127-160.

Bollen, K. A. (1986). Sample size and Bentler and Bonett's nonnormed fit index. *Psychometrika, 51*(3), 375-377.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*(3), 303-316.

Bollen, K. A. (2007). On the origins of latent curve models. In: Cudeck R, MacCallum R, editors. Factor analysis at 100. Mahwah, NJ: Lawrence Erlbaum Associates; 2007. pp. 79–98.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research, 32*(3), 336-383.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467): Wiley.com.

Braskamp, L. A. (2008). The Religious and Spiritual Journeys of College Students. In R. H. J. D. Jacobsen (Ed.), *The American University in a Postsecular Age* (Vol. 1, pp. 117-135). New York: Oxford University Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. *Review of Research in Education, 8*, 158-233.

Cannister, M. W. (1999). Mentoring and the spiritual well-being of late adolescents. *Journal article by Mark W. Cannister; Adolescence, 34*.

Card, N. A., & Little, T. D. (2007). Longitudinal modeling of developmental processes. *International Journal of Behavioral Development, 31*(4), 297-302.

Casey, D. M., Williams, R. J., Mossière, A. M., Schopflocher, D. P., El-Guebaly, N., Hodgins, D. C. . . . Wood, R. T. (2011). The role of family, religiosity, and behavior in adolescent gambling. *Journal of Adolescence*.

Cattell, R. B. (1966). Handbook of multivariate experimental psychology. Chicago: Rand-McNall

Cattell, R. B. (1988). Psychological theory and scientific method, *Handbook of multivariate experimental psychology,* pp. 3-20.

Cheung, C., & Yeung, J. W. (2011). Meta-analysis of relationships between religiosity and constructive and destructive behaviors among adolescents. *Children and Youth Services Review, 33*(2), 376-385.

Ching, C., Fok, T., & Ramsay, J. O. (2006). Periodic Trends, Non-Periodic Trends, and Their Interactions in Longitudinal and Functional Data. In T. A. Walls & J. L. Schafe (Eds.), Models for intensive longitudinal data. New York: Oxford University Press.

Clark, L. S. (2002). US adolescent religious identity, the media, and the "funky" side of religion. *Journal of communication, 52*(4), 794-811.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual review of psychology, 57*, 505-528.

Cox, D., & Lewis, P. (1966). The statistical analysis of series of events. London: Chapman & Hall, 1966.

Cox, D. R. (1972). Regression models and life tables. *JR stat soc B, 34*(2), 187-220.

Cressie, N. C. (1991). Statistics for Spatial Data: Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons.

Cronbach, L. J. (with the assistance of Deken, J. E., & Webb, N.). Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional paper of the Stanford Evaluation Consortium, Stanford University, California, July 1976.

Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual review of psychology, 58*, 615-637.

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological methods, 7*(1), 41.

Cumsille, P. E., Sayer, A. G., & Graham, J. W. (2000). Perceived exposure to peer and adult drinking as predictors of growth in positive alcohol expectancies during adolescence. *Journal of consulting and clinical psychology, 68*(3), 531.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38*(4), 529-569.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development, 11*(2), 121-136.

Curran, P. J., & Willoughby, M. T. (2003). Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology, 15*(3), 581-612.

Day, R. D., Jones-Sanpei, H., Price, J. L. S., Orthner, D. K., Hair, E. C., Moore, K. A., & Kaye, K. (2009). Family processes and adolescent religiosity and religious practice: View from the NLSY97. *Marriage & Family Review, 45*(2-3), 289-309.

DeHaan, L. G., Yonker, J. E., & Affholter, C. (2011). More than enjoying the sunset: Conceptualization and measurement of religiosity for adolescents and emerging adults

and its implications for developmental inquiry. *Journal of Psychology and Christianity, 30*(3), 184.
Desmond, S. A., Soper, S. E., & Kraus, R. (2011). Religiosity, peers, and delinquency: Does religiosity reduce the effets of peers on delinquency? *Sociological Spectrum, 31*(6), 665-694.
Desrosiers, A., & Miller, L. (2007). Relational spirituality and depression in adolescent girls. *Journal of clinical psychology, 63*(10), 1021-1037.
Diggle, Liang, K., & Zeger, S. (1994). *Analysis of longitudinal data.* Oxford, England: Clarendon Press.
Diggle, P. J., & Diggle, P. J. (1983). Statistical analysis of spatial point patterns.
Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological methods, 1*(3), 236.
Fan, J., & Gijbels, I. (1996). *Local polynomial modeling and its applications: monographs on statistics and applied probability 66* (Vol. 66): CRC Press.
Fitzmaurice, G., & Molenberghs, G. (2009). Advances in longitudinal data analysis: an historical perspective. *Longitudinal Data Analysis*, 3-30.
Gibbons, R. D., Hedeker, D., & DuToit, S. (2010). Advances in analysis of longitudinal data. *Annual review of clinical psychology, 6*, 79.
Goldstein, H. (1987). *Multilevel models in education and social research*: Oxford University Press.
Gunnoe, M. L., & Moore, K. A. (2002). Predictors of religiosity among youth aged 17–22: A longitudinal study of the National Survey of Children. *Journal for the Scientific Study of Religion, 41*(4), 613-622.
Hakin Orman, W., North, C., & Gwin, C. (2009). *Mom and Dad Took Me to Church*.
Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: an overview of methodological issues. *Psychology and aging, 18*(4), 639.
Hill, J. P. (2011). Faith and understanding: Specifying the impact of higher education on religious belief. *Journal for the Scientific Study of Religion, 50*(3), 533-551.
Hood Jr, R. W., Hill, P. C., & Spilka, B. (2009). *The psychology of religion: An empirical approach*. New York: The Guilford Press.
Inglehart, R. (2004). *Human beliefs and values: A cross-cultural sourcebook based on the 1999-2002 values surveys*. Mexico City: Siglo XXI.
King, P. E., & Boyatzis, C. J. (2004). Exploring Adolescent Spiritual and Religious Development: Current and Future Theoretical and Empirical Perspectives.
Kuljanin, G., Braun, M. T., & DeShon, R. P. (2011). A cautionary note on modeling growth trends in longitudinal data. *Psychological methods, 16*(3), 249.
Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963-974.
Langeheine, R. (1994). Latent variables Markov models. In A. von Eye & C. C. Clogg (Eds.),*Latent variables analysis: Applications for developmental research*. Thousands Oaks: Sage.
Langeheine, R., & Van de Pol, F. (2002). In J.A. Hagenaars & A.L. McCutcheon (eds.), *Applied latent class analysis* (pp. 304-341). Cambridge, UK: Cambridge University Press.
Lanza, S. T., & Collins, L. M. (2002). Pubertal timing and the onset of substance use in females during early adolescence. *Prevention Science, 3*(1), 69-82.

Lanza, S. T., Collins, L. M., Schafer, J. L., & Flaherty, B. P. (2005). Using data augmentation to obtain standard errors and conduct hypothesis tests in latent class and latent transition analysis. *Psychological methods, 10*(1), 84.

Lanza, S. T., Flaherty, B. P., & Collins, L. M. (2003). Latent class and latent transition analysis. *Handbook of psychology*.

Lewis, P. A. (1972). *Stochastic point processes: statistical analysis, theory, and applications*: Wiley-Interscience Toronto.

Li, R., Root, T., & Shiffman, S. (2006). A local linear estimation procedure for functional multilevel modeling. In *Models for Intensively Longitudinal Data*, (T. Walls and J. Schafer eds), 63-83. Oxford University Press.

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development, 31*(4), 357-365.

Longford, N. T. (1993). *Random coefficient models*. Oxford, England: Clarendon Press.

MacArthur, S. S. (2008). *Adolescent Religiosity, Religious Affiliation, and Premarital Predictors of Marital Quality and Stability*. Unpublished Dissertation.

Martin, T. F., White, J. M., & Perlman, D. (2003). Religious socialization: A test of the channeling hypothesis of parental influence on adolescent faith maturity. *Journal of Adolescent Research, 18*(2), 169.

Mason, W. A., & Spoth, R. L. (2011). Thrill seeking and religiosity in relation to adolescent substance use: Tests of joint, interactive, and indirect influences.

McArdle, J. J. (2005). Five steps in latent curve modeling with longitudinal life-span data. *Advances in life course research, 10*, 315-357.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual review of psychology, 60*, 577-605.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 58, 110-133.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 137–175). Washington, DC: American Psychological Association.

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological bulletin, 107*(2), 247-255.

McNamara Barry, C., Nelson, L., Davarya, S., & Urry, S. (2010). Religiosity and spirituality during the transition to adulthood. *International journal of behavioral development, 34*(4), 311-324.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107-122.

Milevsky, I. M., Szuchman, L., & Milevsky, A. (2008). Transmission of religious beliefs in college students. *Mental Health, Religion and Culture, 11*(4), 423-434.

Miyazaki, Y., & Raudenbush, S. W. (2000). Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological methods, 5*(1), 44-63.

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class–latent

growth modeling. In Collins, L. M. & Sayer, A. (Eds.), *New methods for the analysis of change* (pp. 291-322). Washington, D.C.: APA.
Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research, 24*(6), 882-891.
Muthén , B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological methods, 2*(4), 371-402.
Myers, S. M. (1996). An interactive model of religiosity inheritance: The importance of family context. *American Sociological Review*, 858-866.
Nagin, D., & Nagin, D. (2005). *Group-based modeling of development*: Harvard University Press.
Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological methods, 4*(2), 139-157.
Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychological methods, 6*(1), 18-34.
Nelson, L. J., & Barry, C. M. N. (2005). Distinguishing features of emerging adulthood. *Journal of Adolescent Research, 20*(2), 242-262.
Orathinkal, J., & Vansteenwegen, A. (2006). Religiosity and marital satisfaction. *Contemporary family therapy, 28*(4), 497-504.
Pardun, C. J., & McKee, K. B. (1995). Strange Bedfellows. *Youth & Society, 26*(4), 438-449.
Pearl, J. (2000). *Causality: models, reasoning and inference* (Vol. 29): Cambridge Univ Press.
Petts, R. J. (2009). Trajectories of religious participation from adolescence to young adulthood. *Journal for the Scientific Study of Religion, 48*(3), 552-571.
Puffer, K. A., Pence, K. G., Graverson, T. M., Wolfe, M., Pate, E., & Clegg, S. (2008). Religious doubt and identity formation: Salient predictors of adolescent religious doubt. *Journal of Psychology and Theology, 36*(4), 270-284.
Ramsay, J. O. (2006). In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 176-194). New York: Oxford University Press.
Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics* 14:1–17.
Rathbun, S. L., Shiffman, S., & Gwaltney, C. J. (2006). In *Models for intensive longitudinal data*, T. A. Walls & J. L. Schafer (Eds.), 219–253. New York: Oxford University Press.
Raudenbush, S. W. (2001a). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual review of psychology, 52*(1), 501-525.
Raudenbush, S. W. (2001b). Toward a Coherent Framework for Comparing Trajectories of Individual Change. In C. Horn (Ed.).
Regnerus, M., Smith, C., & Fritsch, M. (2003). Religion in the lives of American adolescents.
Regnerus, M. D., Smith, C., & Smith, B. (2004). Social context in the development of adolescent religiosity. *Applied Developmental Science, 8*(1), 27-38.
Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist, 65*(1), 1-12. doi: 10.1037/a0018326
Rohrbaugh, J., & Jessor, R. (1975). Religiosity in youth: A personal control against deviant behavior. *Journal of Personality, 43*(1), 136-155.
Rostosky, S. S., Wilcox, B. L., Wright, M. L. C., & Randall, B. A. (2004). The impact of religiosity on adolescent sexual behavior. *Journal of Adolescent Research, 19*(6), 677-697.

Rowthorn, R. (2011). Religion, fertility and genes: a dual inheritance model. *Proceedings of the Royal Society B: Biological Sciences, 278*(1717), 2519.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology, 66*(5), 688-701.

Sanchez, Z. M., Opaleye, E. S., Chaves, T. V., Noto, A. R., & Nappo, S. A. (2011). God Forbids or Mom Disapproves? Religious Beliefs That Prevent Drug Use Among Youth. *Journal of Adolescent Research, 26*(5), 591-616.

Schettino, J. R., Olmos, N. T., Myers, H. F., Joseph, N. T., Poland, R. E., & Lesser, I. M. (2011). Religiosity and treatment response to antidepressant medication: a prospective multi-site clinical trial. *Mental Health, Religion & Culture, 14*(8), 805-818.

Schwartz, K. D. (2006). RESEARCH: Transformations in Parent and Friend Faith Support Predicting Adolescents' Religious Faith. *The International Journal for the Psychology of Religion, 16*(4), 311-326.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*: Wadsworth Cengage learning.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*: Oxford university press.

Smith, C. (2003). Theorizing religious effects among American adolescents. *Journal for the Scientific Study of Religion, 42*(1), 17-30.

Smith, C., Denton, M. L., Faris, R., & Regnerus, M. (2002). Mapping American adolescent religious participation. *Journal for the Scientific Study of Religion, 41*(4), 597-612.

Smith, C., & Snell, P. (2009). *Souls in transition: The religious and spiritual lives of emerging adults*. New York: Oxford University Press.

Snijders, T., & Bosker, R. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modeling. *New York*.

Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in cognitive sciences, 9*(2), 69-74.

Stoppa, T. M., & Lefkowitz, E. S. (2010). Longitudinal changes in religiosity among emerging adult college students. *Journal of Research on Adolescence, 20*(1), 23-38.

Theus, M. (2003). Interactive data visualization using Mondrian. *Journal of Statistical Software, 7*(11), 1-9.

Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika, 23*(1), 19-23.

Uecker, J. E., Regnerus, M. D., & Vaaler, M. L. (2007). Losing my religion: The social sources of religious decline in early adulthood. *Social Forces, 85*(4), 1667-1692.

Vaidyanathan, B. (2011). Religious resources or differential returns? Early religious socialization and declining attendance in emerging adulthood. *Journal for the Scientific Study of Religion, 50*(2), 366-387.

Vaughan, E. L., de Dios, M. A., Steinfeldt, J. A., & Kratz, L. M. (2011). Religiosity, alcohol use attitudes, and alcohol use in a national sample of adolescents. *Psychology of Addictive Behaviors, 25*(3), 547-553.

Velleman, P. F. (1989). *Data Desk: Handbook, Volume 1 (1)*: Data Description, Inc.

Weeden, J., Cohen, A. B., & Kenrick, D. T. (2008). Religious attendance as reproductive support. *Evolution and Human Behavior, 29*(5), 327-334.

Wilcox, W. B. (2002). Religion, convention, and paternal involvement. *Journal of Marriage and Family, 64*(3), 780-792.

Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595

Young, F. W., & Bann, C. M. (1996). ViSta: the visual statistics system: Technical Report 94-1 (c).

APPENDIX

The following is the list of available appendices located at http://statcanvas.net/thesis/appendix

Reports

1. Derive dataset
2. Metrics
3. Descriptives
4. Attendance in focus

Illustrations

5. Specification of LCM sequence
6. Descriptives animated over time
7. Databox

Guides

8. Data Manipulation Guide
9. Guide to lmer() syntax

Sequence reports

Time metric: Year

10. Whites, complete trajectories
11. Whites, complete trajectories: Fit

Time metric: Age

12. Whites, complete trajectories
13. Whites, complete trajectories: Fit