Life Loses A Little Meaning After Losing Religion

Daryl R. Van Tongeren

Chris G. Sibley

Don E Davis

Joseph A. Bulbulia

2025-04-01

Abstract

**KEYWORDS**: *Causal Inference*; *Church*; *Cross-validation*; *Distress*; *Health*; *Longitudinal*; *Machine Learning*; *Religion*; *Semi-parametric*; *Targeted Learning*.

## Introduction

## Method

### Sample

Data were collected as part of the New Zealand Attitudes and Values Study (NZAVS), an annual longitudinal national probability panel assessing New Zealand residents’ social attitudes, personality, ideology, and health outcomes. The panel began in 2009 and has since expanded to include over fifty researchers, with responses from 76,409 participants to date. The study operates independently of political or corporate funding and is based at a university. It employs prize draws to incentivise participation. The NZAVS tends to slightly under-sample males and individuals of Asian descent and to over-sample females and Māori (the Indigenous people of New Zealand). To enhance the representativeness of our sample population estimates for the target population of New Zealand, we apply census-based survey weights that adjust for age, gender, and ethnicity (New Zealand European, Asian, Māori, Pacific) (Sibley 2021). For more information about the NZAVS, visit: [OSF.IO/75SNB](https://doi.org/10.17605/OSF.IO/75SNB). Refer to [Appendix A](#appendix-timeline) for a histogram of daily responses for this cohort.

### Target Population

The target population for this study comprises the cohort of New Zealand residents in New Zealand Attitudes and Values Study wave 10 (years 2018-2019) (Sibley 2021).

### Treatment Indicator

The New Zealand Attitudes and Values Study assesses religious service attendance using the following question:

* *How important is your religion to how you see yourself?*

Ordinal response: (1 = Not Important, 7 = Very Important). This question was only given to those who identify with as religious. Those who did not identify as religious were imputed a value of “1” (Measured developed for the NZAVS.)

For measures, refer to [Appendix B](#appendix-baseline)

### Baseline Covariates

We adjusted for a rich set of demographic, personality, and behavioural indicators measured at the baseline wave, NZAVS time 10 (Wave 2018, years 2018-2019) (see [Appendix B](#appendix-baseline) for full measures). These variables included age, gender, ethnicity, education level, personality traits (Agreeableness, Conscientiousness, Extraversion, Honesty-Humility, Neuroticism, and Openness), household income, employment status, parenting status, relationship status, religious belonging, and health-related behaviours (e.g. smoking, alcohol use, hours spent exercising). We selected only those outcome variables measured in the baseline wave and controlled for these variables. Moreover we controlled for religious identification at baseline (refer to [Appendix B](#appendix-baseline) and [Appendix C](#appendix-confounding)) This strategy of confounding control is powerful because for any confounder to affect subsequent treatments and the outcome, it would need to do so independently of the baseline outcome variables, the baseline exposure, and the rich set of demographic indicators measured at baseline (VanderWeele *et al.* 2020).

### Outcomes

#### Meaning Purpose

*My life has a clear sense of purpose*

Ordinal response (1 = Strongly Disagree to 7 = Strongly Agree). (Steger *et al.* 2006)

#### Meaning Sense

*I have a good sense of what makes my life meaningful.*

Ordinal response (1 = Strongly Disagree to 7 = Strongly Agree). (Steger *et al.* 2006)

### Statistical Estimator

We estimate causal effects of time-varying treatment policies using a Sequential Doubly-Robust (SDR) estimator in the lmtp package (Williams and Díaz 2021). SDR proceeds in two steps. First, we use machine learning to flexibly model relationships among treatments, covariates, and outcomes. This approach captures complex, high-dimensional structures without strict assumptions (Díaz *et al.* 2021). Second, SDR “targets” these initial estimates by incorporating information from the observed data distribution. This step iteratively refines the accuracy of our causal estimates.

The SDR estimator is multiply robust when treatments repeat over multiple waves (Díaz *et al.* 2023; Hoffman *et al.* 2023). This design maintains consistency if either the outcome model or treatment model is correctly specified. The lmtp package relies on the SuperLearner library in R (Polley *et al.* 2023b). We used SL.ranger, SL.glmnet, and SL.xgboost (Chen *et al.* 2023; Polley *et al.* 2023a; Wright and Ziegler 2017) as our base learners. **SL.ranger**: implements a random forest algorithm, capturing non-linear relationships and complex interactions. **SL.glmnet**: provides regularised linear models for high-dimensional data. **SL.xgboost**: uses gradient boosting to capture intricate patterns without over-fitting.

SuperLearner combines these learners adaptively to optimise predictive performance. We created graphs, tables, and output reports with the margot package (Bulbulia 2024a). For more details on targeted learning with lmtp, see (Díaz *et al.* 2021; Hoffman *et al.* 2022, 2023).

### Handling of Missing Data

#### Baseline Missingness

We used predictive mean matching from the mice package (Van Buuren 2018) to impute missing baseline values (comprising 1.5929465 of the baseline data). Following (Zhang *et al.* 2023), we performed single imputation using only baseline data. For each column with missing values, we created a binary indicator of missingness so that the machine learning algorithms we employed could condition on missingness information during estimation (see lmtp documentation (Williams and Díaz 2021)).

#### Missingness in Time-Varying Variables

When a time-varying value was missing in any wave but a future value was observed, we carried forward the previous response and included a missingness indicator. Again, this approach let the patterns of missingness inform nonparametric machine learning. If no future value was observed, we considered the participant censored and used inverse probability of treatment weights to address attrition.

#### Outcome Missingness

Finally, to handle confounding and selection bias arising from missing outcomes and panel attrition, we applied inverse probability of censoring weights, estimated via nonparametric machine learning ensembles in the lmtp package (Williams and Díaz 2021).

### Sensitivity Analysis

We perform sensitivity analyses using the E-value metric (Linden *et al.* 2020; VanderWeele and Ding 2017). The E-value represents the minimum association strength (on the risk ratio scale) that an unmeasured confounder would need to have with both the exposure and outcome—after adjusting for measured covariates—to explain away the observed exposure-outcome association (Linden *et al.* 2020; VanderWeele *et al.* 2020).

## Results

### Dones vs Religious

|  |
| --- |
| Figure 1: Dones vs Religious |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Health effects   |  | E[Y(1)]-E[Y(0)] | 2.5 % | 97.5 % | E\_Value | E\_Val\_bound | | --- | --- | --- | --- | --- | --- | | Meaning Purpose | -0.17 | -0.21 | -0.13 | 1.60 | 1.50 | | Meaning Sense | -0.22 | -0.28 | -0.15 | 1.74 | 1.56 | |

#### Meaning purpose

The effect estimate (rd) is -0.168 (-0.208, -0.128). On the original scale, the estimated effect is -0.238 (-0.294, -0.181). E-value lower bound is 1.498, indicating evidence for causality.

#### Meaning sense

The effect estimate (rd) is -0.219 (-0.285, -0.153). On the original scale, the estimated effect is -0.262 (-0.341, -0.183). E-value lower bound is 1.562, indicating evidence for causality.

### Dones vs Secular

|  |
| --- |
| Figure 2: Dones vs Secular |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Effects on Psychological Well-Being   |  | E[Y(1)]-E[Y(0)] | 2.5 % | 97.5 % | E\_Value | E\_Val\_bound | | --- | --- | --- | --- | --- | --- | | Meaning Purpose | 0.20 | 0.16 | 0.24 | 1.69 | 1.59 | | Meaning Sense | 0.13 | 0.08 | 0.18 | 1.51 | 1.38 | |

#### Meaning purpose

The effect estimate (rd) is 0.202 (0.164, 0.241). On the original scale, the estimated effect is 0.286 (0.231, 0.34). E-value lower bound is 1.59, indicating evidence for causality.

#### Meaning sense

The effect estimate (rd) is 0.134 (0.084, 0.184). On the original scale, the estimated effect is 0.16 (0.1, 0.22). E-value lower bound is 1.375, indicating evidence for causality.

### Secular vs Religious

|  |
| --- |
| Figure 3: Secular vs Religious |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Effects on Life-Focussed Well-Being   |  | E[Y(1)]-E[Y(0)] | 2.5 % | 97.5 % | E\_Value | E\_Val\_bound | | --- | --- | --- | --- | --- | --- | | Meaning Purpose | -0.37 | -0.43 | -0.32 | 2.15 | 2.00 | | Meaning Sense | -0.35 | -0.44 | -0.27 | 2.10 | 1.88 | |

#### Meaning purpose

The effect estimate (rd) is -0.37 (-0.426, -0.315). On the original scale, the estimated effect is -0.524 (-0.602, -0.445). E-value lower bound is 1.998, indicating evidence for causality.

#### Meaning sense

The effect estimate (rd) is -0.353 (-0.436, -0.27). On the original scale, the estimated effect is -0.422 (-0.522, -0.323). E-value lower bound is 1.878, indicating evidence for causality.

## Discussion

TBA

### Ethics

The University of Auckland Human Participants Ethics Committee reviews the NZAVS every three years. Our most recent ethics approval statement is as follows: The New Zealand Attitudes and Values Study was approved by the University of Auckland Human Participants Ethics Committee on 26/05/2021 for six years until 26/05/2027, Reference Number UAHPEC22576.

### Data Availability

The data described in the paper are part of the New Zealand Attitudes and Values Study. Members of the NZAVS management team and research group hold full copies of the NZAVS data. A de-identified dataset containing only the variables analysed in this manuscript is available upon request from the corresponding author or any member of the NZAVS advisory board for replication or checking of any published study using NZAVS data. The code for the analysis can be found at [OSF link](https://osf.io/ab7cx/).

### Acknowledgements

The New Zealand Attitudes and Values Study is supported by a grant from the Templeton Religious Trust (TRT0196; TRT0418). JB received support from the Max Plank Institute for the Science of Human History. The funders had no role in preparing the manuscript or deciding to publish it.

### Author Statement

TBA

## Appendix A: Daily Data Collection {#appendix-timeline}

**?@fig-timeline** presents the New Zealand Attitudes and Values Study Data Collection (2018 retained cohort) from years 2018-2014 (NZAVS time 10–time 15). (not run)

## Appendix B: Measures and Demographic Statistics

### Measures

### Baseline Variables

#### Age

*What is your date of birth?*

We asked participants’ ages in an open-ended question (“What is your age?” or “What is your date of birth”). (**string\_is?** Developed for the NZAVS.)

#### Agreeableness

*I sympathize with others’ feelings.* *I am not interested in other people’s problems.* *I feel others’ emotions.* *I am not really interested in others (reversed).*

Mini-IPIP6 Agreeableness dimension: (i) I sympathize with others’ feelings. (ii) I am not interested in other people’s problems. (r) (iii) I feel others’ emotions. (iv) I am not really interested in others. (r) (Sibley *et al.* 2011)

#### Alcohol Frequency

*“How often do you have a drink containing alcohol?”*

Participants could chose between the following responses: ‘(1 = Never - I don’t drink, 2 = Monthly or less, 3 = Up to 4 times a month, 4 = Up to 3 times a week, 5 = 4 or more times a week, 6 = Don’t know)’ (Health 2013)

#### Alcohol Intensity

*“How many drinks containing alcohol do you have on a typical day when drinking alcohol? (number of drinks on a typical day when drinking)”*

Participants responded using an open-ended box. (Health 2013)

#### Belong

*Know that people in my life accept and value me.* *Feel like an outsider (reversed).* *Know that people around me share my attitudes and beliefs.*

We assessed felt belongingness with three items adapted from the Sense of Belonging Instrument (Hagerty & Patusky, 1995): (1) “Know that people in my life accept and value me”; (2) “Feel like an outsider”; (3) “Know that people around me share my attitudes and beliefs”. Participants responded on a scale from 1 (Very Inaccurate) to 7 (Very Accurate). The second item was reversely coded. (Hagerty and Patusky 1995)

#### Born Nz Binary

*Where were you born? (please be specific, e.g., which town/city?)*

Coded binary (1 = New Zealand; 0 = elsewhere.) (**string\_is?** Developed for the NZAVS.)

#### Conscientiousness

*I get chores done right away.* *I like order.* *I make a mess of things.* *I often forget to put things back in their proper place.*

Mini-IPIP6 Conscientiousness dimension: (i) I get chores done right away. (ii) I like order. (iii) I make a mess of things. (r) (iv) I often forget to put things back in their proper place. (r) (Sibley *et al.* 2011)

#### Education Level Coarsen

*What is your highest level of qualification?*

We asked participants, “What is your highest level of qualification?”. We coded participans highest finished degree according to the New Zealand Qualifications Authority. Ordinal-Rank 0-10 NZREG codes (with overseas school qualifications coded as Level 3, and all other ancillary categories coded as missing) (**string\_is?** Developed for the NZAVS.)

#### Employed Binary

*Are you currently employed (This includes self-employed of casual work)?*

Binary response: (0 = No, 1 = Yes) (**string\_is?** Stats NZ Census Question)

#### Eth Cat

*Which ethnic group(s) do you belong to?*

Coded string: (1 = New Zealand European; 2 = Māori; 3 = Pacific; 4 = Asian) (**string\_is?** NZ Census coding.)

#### Extraversion

*I am the life of the party.* *I don’t talk a lot (reversed).* *I keep in the background (reversed).* *I talk to a lot of different people at parties.*

Mini-IPIP6 Extraversion dimension: (i) I am the life of the party. (ii) I don’t talk a lot. (r) (iii) I keep in the background. (r) (iv) I talk to a lot of different people at parties. (Sibley *et al.* 2011)

#### Hlth Disability Binary

*Do you have a health condition or disability that limits you and that has lasted for 6+ months?*

We assessed disability with a one-item indicator adapted from Verbrugge (1997). It asks, “Do you have a health condition or disability that limits you and that has lasted for 6+ months?” (1 = Yes, 0 = No). (Verbrugge 1997)

#### Hlth Fatigue

*During the last 30 days, how often did … you feel exhausted?*

Ordinal response: (0 = None Of The Time; 1 = A Little Of The Time; 2= Some Of The Time; 3 = Most Of The Time; 4 = All Of The Time) (Sibley *et al.* 2020)

#### Honesty Humility

*I feel entitled to more of everything (reversed).* *I deserve more things in life (reversed).* *I deserve more things in life (reversed).* *I would get a lot of pleasure from owning expensive luxury goods (reversed).*

Mini-IPIP6 Honesty-Humility dimension: (i) I feel entitled to more of everything. (r) (ii) I deserve more things in life. (r) (iii) I would like to be seen driving around in a very expensive car. (r) (iv) I would get a lot of pleasure from owning expensive luxury goods. (r) (Sibley *et al.* 2011)

#### Hours Children

No information available for this variable.

#### Hours Commute

*Hours spent…travelling/commuting.*

(**string\_is?** Developed for the NZAVS.)

#### Hours Exercise

No information available for this variable.

#### Hours Housework

No information available for this variable.

#### Household Inc

*Please estimate your total household income (before tax) for the year XXXX.*

#### Kessler Latent Anxiety

*During the past 30 days, how often did…you feel restless or fidgety?* *During the past 30 days, how often did…you feel that everything was an effort?* *During the past 30 days, how often did…you feel nervous?*

Ordinal response: (0 = None Of The Time; 1 = A Little Of The Time; 2= Some Of The Time; 3 = Most Of The Time; 4 = All Of The Time) (Kessler *et al.* 2002)

#### Kessler Latent Depression

*During the past 30 days, how often did…you feel hopeless?* *During the past 30 days, how often did…you feel so depressed that nothing could cheer you up?* *During the past 30 days, how often did…you feel you feel restless or fidgety?*

Ordinal response: (0 = None Of The Time; 1 = A Little Of The Time; 2= Some Of The Time; 3 = Most Of The Time; 4 = All Of The Time) (Kessler *et al.* 2002)

#### Log Hours Children

*Hours spent…looking after children.*

We took the natural log of the response + 1. (Sibley *et al.* 2011)

#### Log Hours Commute

*Hours spent…travelling/commuting.*

We took the natural log of the response + 1. (**string\_is?** Developed for the NZAVS.)

#### Log Hours Exercise

*Hours spent…exercising/physical activity.*

We took the natural log of the response + 1. (Sibley *et al.* 2011)

#### Log Hours Housework

*Hours spent…housework/cooking.*

We took the natural log of the response + 1. (Sibley *et al.* 2011)

#### Log Household Inc

*Please estimate your total household income (before tax) for the year XXXX.*

We took the natural log of the response + 1. (**string\_is?** Developed for the NZAVS.)

#### Male Binary

*We asked participants’ gender in an open-ended question: “what is your gender?”*

Here, we coded all those who responded as Male as 1, and those who did not as 0. (Fraser *et al.* 2020)

#### Neuroticism

*I have frequent mood swings.* *I am relaxed most of the time (reversed).* *I get upset easily.* *I seldom feel blue (reversed).*

Mini-IPIP6 Neuroticism dimension: (i) I have frequent mood swings. (ii) I am relaxed most of the time. (r) (iii) I get upset easily. (iv) I seldom feel blue. (r) (Sibley *et al.* 2011)

#### Not Heterosexual Binary

*How would you describe your sexual orientation? (e.g., heterosexual, homosexual, straight, gay, lesbian, bisexual, etc.)*

Open-ended question, coded as binary (not heterosexual = 1). (Greaves *et al.* 2017)

#### Nz Dep2018

*New Zealand Deprivation - Decile Index - Using 2018 Census Data*

Numerical: (1-10) (Atkinson *et al.* 2019)

#### Nzsei 13 l

*We assessed occupational prestige and status using the New Zealand Socio-economic Index 13 (NZSEI-13).*

This index uses the income, age, and education of a reference group, in this case, the 2013 New Zealand census, to calculate a score for each occupational group. Scores range from 10 (Lowest) to 90 (Highest). This list of index scores for occupational groups was used to assign each participant a NZSEI-13 score based on their occupation. (Fahy *et al.* 2017)

#### Openness

*I have a vivid imagination.* *I have difficulty understanding abstract ideas (reversed).* *I do not have a good imagination (reversed).* *I am not interested in abstract ideas (reversed).*

Mini-IPIP6 Openness to Experience dimension: (i) I have a vivid imagination. (ii) I have difficulty understanding abstract ideas. (r) (iii) I do not have a good imagination. (r) (iv) I am not interested in abstract ideas. (r) (Sibley *et al.* 2011)

#### Parent Binary

*If you are a parent, in which year was your eldest child born?*

Parents were coded as 1, while the others were coded as 0. (**Developed?** for the NZAVS.)

#### Partner Binary

*What is your relationship status? (e.g., single, married, de-facto, civil union, widowed, living together, etc.)*

Coded as binary (has partner = 1). (**string\_is?** Developed for the NZAVS.)

#### Political Conservative

*Please rate how politically liberal versus conservative you see yourself as being.*

Ordinal response: (1 = Extremely Liberal, 7 = Extremely Conservative) (Jost 2006)

#### Rural Gch 2018 l

*High Urban Accessibility = 1, Medium Urban Accessibility = 2, Low Urban Accessibility = 3, Remote = 4, Very Remote = 5.*

“Participants residence locations were coded according to a five-level ordinal categorisation ranging from Urban to Rural.” (Whitehead *et al.* 2023)

#### Rwa

*It is always better to trust the judgment of the proper authorities in government and religion than to listen to the noisy rabble-rousers in our society who are trying to create doubt in people’s minds.* *It would be best for everyone if the proper authorities censored magazines so that people could not get their hands on trashy and disgusting material.* *Our country will be destroyed some day if we do not smash the perversions eating away at our moral fibre and traditional beliefs.* *People should pay less attention to The Bible and other old traditional forms of religious guidance, and instead develop their own personal standards of what is moral and immoral.* *Atheists and others who have rebelled against established religions are no doubt every bit as good and virtuous as those who attend church regularly.* *Some of the best people in our country are those who are challenging our government, criticizing religion, and ignoring the “normal way” things are supposed to be done (reversed).*

#### Sample Frame Opt in Binary

*Participant was not randomly sampled from the New Zealand Electoral Roll.*

Code string (Binary): (0 = No, 1 = Yes) (**string\_is?** Developed for the NZAVS.)

#### Sdo

*It is OK if some groups have more of a chance in life than others.* *Inferior groups should stay in their place.* *To get ahead in life, it is sometimes okay to step on other groups.* *We should have increased social equality (reversed).* *It would be good if groups could be equal (reversed).* *We should do what we can to equalise conditions for different groups (reversed).*

#### Short Form Health

*In general, would you say your health is…*

Ordinal response: (1 = Poor, 7 = Excellent) (Instrument Ware Jr and Sherbourne 1992)

#### Smoker Binary

*Do you currently smoke tobacco cigarettes?*

Binary smoking indicator (0 = No, 1 = Yes). (**string\_is?** Developed for NZAVS.)

#### Support

*There are people I can depend on to help me if I really need it.* *There is no one I can turn to for guidance in times of stress (reversed).* *I know there are people I can turn to when I need help.*

Ordinal response: (1 = Strongly Disagree, 7 = Strongly Agree) (Cutrona and Russell 1987)

### Exposure Variable

#### Religion Identification Level

*How important is your religion to how you see yourself?*

Ordinal response: (1 = Not Important, 7 = Very Important) (**string\_is?** Developed for the NZAVS.)

### Outcome Variables

#### Meaning Purpose

*My life has a clear sense of purpose*

Ordinal response (1 = Strongly Disagree to 7 = Strongly Agree). (Steger *et al.* 2006)

#### Meaning Sense

*I have a good sense of what makes my life meaningful.*

Ordinal response (1 = Strongly Disagree to 7 = Strongly Agree). (Steger *et al.* 2006)

### Sample Demographic Statistics

[Table 4](#tbl-baseline) presents sample demographic statistics.

|  |
| --- |
| Table 4: Demographic statistics for New Zealand Attitudes and Values Cohort waves 2018. |

### Exposure Variable: Religious Identification

[Table 5](#tbl-sample-exposures) presents sample statistics for the exposure variable, religious identification, during the baseline and exposure waves. This variable was not measured in part of NZAVS time 12 (years 2020-2021) and part of NZAVS time 13 (years 2021-2022). To address missingness, if a value was observed after NZAVS time 14, we carried the previous observation forward and created and NA indicator. If there was no future observation, the participant was treated as censored, and inverse probability of censoring weights were applied, following our standard method for handling missing observations (see mansucript **Method**/**Handling of Missing Data**). Here, our carry-forward imputation approach may result in conservative causal effect estimation because it introduces measurement error. However, this approach would not generally bias causal effect estimation away from the null because the measurement error is unsystematic and random and unrelated to the outcomes.

|  |
| --- |
| Table 5: Exposure descriptive statistics by wave. |

### Outcome Variables

|  |
| --- |
| Table 6: Outcome Variables at baseline (NZAVS time 10, years 2018-2019, and time 15, years 2023-2024). |

## Appendix C: Confouding Control

For confounding control, we employ a modified disjunctive cause criterion (VanderWeele 2019), which involves:

1. Identifying all common causes of both the treatment and outcomes.
2. Excluding instrumental variables that affect the exposure but not the outcome.
3. Including proxies for unmeasured confounders affecting both exposure and outcome.
4. Controlling for baseline exposure and baseline outcome, serving as proxies for unmeasured common causes (VanderWeele *et al.* 2020).

Additionally, we control for time-varying confounders at each exposure wave (Bulbulia 2024b; Richardson and Robins 2013; Robins and Hernan 2008).

The covariates included for confounding control are described in Rosa *et al.* (2024).

Where there are multiple exposures, causal inference may be threatened by time-varying confounding (Bulbulia 2024b).

## Appendix D: Causal Contrasts and Causal Assumptions

### Notation

* : Observed religious identification at Wave , for .
* : Outcome measured at the end of the study (Wave 5).
* : Confounders measured at baseline (Wave 0).
* : Time-varying confounders measured at Wave (for ).

### Shift Functions

Let represent the **regular attendance** treatment sequence and the **no attendance** treatment sequence, where the interventions occure at each wave . Formally:

#### Steady Religious

#### Steady Secular

Here, is the observed attendance at Wave . The shift function “nudges” to a target level (four times per month or zero) only if the current value is below (for regular attendance) or above (for no attendance) that target. Across the four waves, these shifts form a sequence , which defines a complete intervention regime.

### Dones

This policy:

1. Applies the steady religious (nudging up to 7) during waves 1 and 2.
2. Applies the steady secular rule (nudging down to 1) during waves 3 and 4.

### Causal Contrast

We focus on primarily on two causal contrasts. The difference between de-identification and steady religion:

$$
\text{ATE}^{\text{done}}
\;=\;
\mathbb{E}
\Bigl[
Y\_\tau\!\bigl(\boldsymbol{\text{d}}(a^+)\bigr)
\;-\;
Y\_\tau\!\bigl(\boldsymbol{\text{d}}(a^{+/-})\bigr)|W\_0, L\_k
\Bigr].
$$

and the difference between de-identification and steady secular:

$$
\text{ATE}^{\text{residue}}
\;=\;
\mathbb{E}
\Bigl[
Y\_\tau\!\bigl(\boldsymbol{\text{d}}(a^-)\bigr)
\;-\;
Y\_\tau\!\bigl(\boldsymbol{\text{d}}(a^{+/-})\bigr)|W\_0, L\_k
\Bigr].
$$

### Assumptions

To estimate this effect from observational data, we assume:

1. **Conditional Exchangeability:** Once we condition on and each , the interventions or are effectively random with respect to potential outcomes.
2. **Consistency:** The potential outcome under a given regime matches the observed outcome when that regime is followed.
3. **Positivity:** Everyone has a non-zero probability of receiving each level of attendance (i.e., a chance to be “shifted” up or down) given their covariates. The positivity assumption is the only causal assumption that can be evaluated with data. We evaluate this assumption in [Appendix E](#appendix-transition)).

Mathematically, for conditional exchangeability, we write:

$$
\Bigl\{
Y\bigl(\boldsymbol{\text{d}}(a^+)\bigr),
\;
Y\bigl(\boldsymbol{\text{d}}(a^-)\bigr),
\;
Y\bigl(\boldsymbol{\text{d}}(a^{+/-})\bigr)
\Bigr\}
\coprod
A\_k |
W\_0,
L\_k
$$

That is, we assume the potential outcomes under each treatment regime are independent of each treatment at every time point, conditional on baseline confounders and time-varying confounders.

Under these assumptions, our statistical models permit us to estimate and e from observational data. We define the target population as the New Zealand Population from 2019-2024, the years in which measurements were taken, had no one been censored/lost to follow up.

## Appendix E: Transition Matrix to Check The Positivity Assumption

These transition matrices capture shifts in states between consecutive waves. Each cell represents the count of individuals transitioning from one state to another. The rows correspond to the initial state (From), and the columns correspond to the subsequent state (To). **Diagonal entries** (in **bold**) correspond to individuals who remained in the same state. **Off-diagonal entries** correspond to individuals who transitioned to a different state.

A higher number on the diagonal relative to off-diagonal entries indicates greater stability in a state. Conversely, higher off-diagonal numbers suggest more frequent shifts between states.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 7: Transition Matrix From Wave 2018 to Wave 2019   | From / To | State 1 | State 2 | State 3 | State 4 | State 5 | State 6 | State 7 | Total | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | State 1 | 21923 | 328 | 147 | 251 | 165 | 92 | 104 | 23010 | | State 2 | 632 | 372 | 112 | 167 | 62 | 13 | 5 | 1363 | | State 3 | 287 | 169 | 149 | 190 | 84 | 13 | 4 | 896 | | State 4 | 353 | 140 | 164 | 478 | 299 | 80 | 42 | 1556 | | State 5 | 250 | 71 | 104 | 360 | 727 | 322 | 113 | 1947 | | State 6 | 111 | 13 | 28 | 104 | 351 | 656 | 389 | 1652 | | State 7 | 111 | 5 | 4 | 38 | 114 | 460 | 2772 | 3504 | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 8: Transition Matrix From Wave 2019 to Wave 2020   | From / To | State 1 | State 2 | State 3 | State 4 | State 5 | State 6 | State 7 | Total | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | State 1 | 17854 | 299 | 181 | 314 | 200 | 86 | 101 | 19035 | | State 2 | 375 | 223 | 132 | 125 | 40 | 5 | 5 | 905 | | State 3 | 156 | 91 | 100 | 120 | 74 | 17 | 6 | 564 | | State 4 | 248 | 124 | 134 | 366 | 276 | 78 | 21 | 1247 | | State 5 | 145 | 40 | 82 | 252 | 554 | 267 | 73 | 1413 | | State 6 | 84 | 11 | 25 | 60 | 271 | 543 | 297 | 1291 | | State 7 | 86 | 3 | 11 | 29 | 93 | 377 | 2099 | 2698 | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 9: Transition Matrix From Wave 2020 to Wave 2021   | From / To | State 1 | State 2 | State 3 | State 4 | State 5 | State 6 | State 7 | Total | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | State 1 | 15718 | 236 | 99 | 182 | 124 | 47 | 52 | 16458 | | State 2 | 285 | 188 | 94 | 74 | 33 | 8 | 6 | 688 | | State 3 | 164 | 82 | 137 | 141 | 58 | 12 | 5 | 599 | | State 4 | 238 | 84 | 154 | 358 | 214 | 51 | 16 | 1115 | | State 5 | 139 | 34 | 72 | 246 | 564 | 195 | 59 | 1309 | | State 6 | 68 | 6 | 10 | 48 | 270 | 521 | 250 | 1173 | | State 7 | 80 | 2 | 4 | 24 | 65 | 298 | 1816 | 2289 | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 10: Transition Matrix From Wave 2021 to Wave 2022   | From / To | State 1 | State 2 | State 3 | State 4 | State 5 | State 6 | State 7 | Total | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | State 1 | 12386 | 258 | 98 | 203 | 104 | 68 | 80 | 13197 | | State 2 | 206 | 152 | 65 | 62 | 37 | 11 | 3 | 536 | | State 3 | 115 | 98 | 70 | 96 | 50 | 11 | 0 | 440 | | State 4 | 177 | 87 | 107 | 245 | 171 | 46 | 25 | 858 | | State 5 | 119 | 35 | 67 | 217 | 372 | 217 | 80 | 1107 | | State 6 | 49 | 5 | 4 | 38 | 180 | 376 | 246 | 898 | | State 7 | 52 | 3 | 2 | 20 | 42 | 166 | 1497 | 1782 | |

## References

Atkinson, J, Salmond, C, and Crampton, P (2019) *NZDep2018 index of deprivation, user’s manual.*, Wellington.

Bulbulia, JA (2024a) *Margot: MARGinal observational treatment-effects*. doi:[10.5281/zenodo.10907724](https://doi.org/10.5281/zenodo.10907724).

Bulbulia, JA (2024b) Methods in causal inference part 2: Interaction, mediation, and time-varying treatments. *Evolutionary Human Sciences*, **6**, e41. doi:[10.1017/ehs.2024.32](https://doi.org/10.1017/ehs.2024.32).

Chen, T, He, T, Benesty, M, … Yuan, J (2023) *Xgboost: Extreme gradient boosting*. Retrieved from <https://CRAN.R-project.org/package=xgboost>

Cutrona, CE, and Russell, DW (1987) The provisions of social relationships and adaptation to stress. *Advances in Personal Relationships*, **1**, 37–67.

Díaz, I, Williams, N, Hoffman, KL, and Schenck, EJ (2021) Non-parametric causal effects based on longitudinal modified treatment policies. *Journal of the American Statistical Association*. doi:[10.1080/01621459.2021.1955691](https://doi.org/10.1080/01621459.2021.1955691).

Díaz, I, Williams, N, Hoffman, KL, and Schenck, EJ (2023) Nonparametric causal effects based on longitudinal modified treatment policies. *Journal of the American Statistical Association*, **118**(542), 846–857. doi:[10.1080/01621459.2021.1955691](https://doi.org/10.1080/01621459.2021.1955691).

Fahy, KM, Lee, A, and Milne, BJ (2017) *New Zealand socio-economic index 2013*, Wellington, New Zealand: Statistics New Zealand-Tatauranga Aotearoa.

Fraser, G, Bulbulia, J, Greaves, LM, Wilson, MS, and Sibley, CG (2020) Coding responses to an open-ended gender measure in a New Zealand national sample. *The Journal of Sex Research*, **57**(8), 979–986. doi:[10.1080/00224499.2019.1687640](https://doi.org/10.1080/00224499.2019.1687640).

Greaves, LM, Barlow, FK, Lee, CH, et al.others (2017) The diversity and prevalence of sexual orientation self-labels in a New Zealand national sample. *Archives of Sexual Behavior*, **46**, 1325–1336.

Hagerty, BMK, and Patusky, K (1995) Developing a Measure Of Sense of Belonging: *Nursing Research*, **44**(1), 9–13. doi:[10.1097/00006199-199501000-00003](https://doi.org/10.1097/00006199-199501000-00003).

Health, Ministry of (2013) *The New Zealand Health Survey: Content guide 2012-2013*, Princeton University Press.

Hoffman, KL, Salazar-Barreto, D, Rudolph, KE, and Díaz, I (2023) Introducing longitudinal modified treatment policies: A unified framework for studying complex exposures. doi:[10.48550/arXiv.2304.09460](https://doi.org/10.48550/arXiv.2304.09460).

Hoffman, KL, Schenck, EJ, Satlin, MJ, … Díaz, I (2022) Comparison of a target trial emulation framework vs cox regression to estimate the association of corticosteroids with COVID-19 mortality. *JAMA Network Open*, **5**(10), e2234425. doi:[10.1001/jamanetworkopen.2022.34425](https://doi.org/10.1001/jamanetworkopen.2022.34425).

Instrument Ware Jr, J, and Sherbourne, C (1992) The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, **30**(6), 473–483.

Jost, JT (2006) The end of the end of ideology. *American Psychologist*, **61**(7), 651–670. doi:[10.1037/0003-066X.61.7.651](https://doi.org/10.1037/0003-066X.61.7.651).

Kessler, R C, Andrews, G, Colpe, L J, … Zaslavsky, A M (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, **32**(6), 959–976. doi:[10.1017/S0033291702006074](https://doi.org/10.1017/S0033291702006074).

Linden, A, Mathur, MB, and VanderWeele, TJ (2020) Conducting sensitivity analysis for unmeasured confounding in observational studies using e-values: The evalue package. *The Stata Journal*, **20**(1), 162–175.

Polley, E, LeDell, E, Kennedy, C, and Laan, M van der (2023a) *SuperLearner: Super learner prediction*. Retrieved from <https://CRAN.R-project.org/package=SuperLearner>

Polley, E, LeDell, E, Kennedy, C, and van der Laan, M (2023b) *SuperLearner: Super learner prediction*. Retrieved from <https://github.com/ecpolley/SuperLearner>

Richardson, TS, and Robins, JM (2013) Single world intervention graphs: A primer. In, Citeseer. Retrieved from <https://core.ac.uk/display/102673558>

Robins, J, and Hernan, M (2008) Estimation of the causal effects of time-varying exposures. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*, 553–599.

Rosa, PA de la, Cowden, RG, Bulbulia, JA, Sibley, CG, and VanderWeele, TJ (2024) Effects of screen-based leisure time on 24 subsequent health and wellbeing outcomes: A longitudinal outcome-wide analysis. *International Journal of Behavioral Medicine*, 1–20. doi:<https://doi.org/10.1007/s12529-024-10307-0>.

Sibley, CG (2021) [*Sampling procedure and sample details for the New Zealand Attitudes and Values Study*](https://doi.org/10.31234/osf.io/wgqvy).

Sibley, CG, Afzali, MU, Satherley, N, … others (2020) Prejudice toward muslims in New Zealand: Insights from the New Zealand Attitudes and Values Study. *New Zealand Journal of Psychology*, **49**(1).

Sibley, CG, Luyten, N, Purnomo, M, … Robertson, A (2011) The Mini-IPIP6: Validation and extension of a short measure of the Big-Six factors of personality in New Zealand. *New Zealand Journal of Psychology*, **40**(3), 142–159.

Steger, MF, Frazier, P, Oishi, S, and Kaler, M (2006) The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, **53**(1), 80–93. doi:[10.1037/0022-0167.53.1.80](https://doi.org/10.1037/0022-0167.53.1.80).

Van Buuren, S (2018) *Flexible imputation of missing data*, CRC press.

VanderWeele, TJ (2019) Principles of confounder selection. *European Journal of Epidemiology*, **34**(3), 211–219.

VanderWeele, TJ, and Ding, P (2017) Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, **167**(4), 268–274. doi:[10.7326/M16-2607](https://doi.org/10.7326/M16-2607).

VanderWeele, TJ, Mathur, MB, and Chen, Y (2020) Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science*, **35**(3), 437–466.

Verbrugge, LM (1997) A global disability indicator. *Journal of Aging Studies*, **11**(4), 337–362. doi:[10.1016/S0890-4065(97)90026-8](https://doi.org/10.1016/S0890-4065(97)90026-8).

Whitehead, J, Davie, G, Graaf, B de, … Nixon, G (2023) Unmasking hidden disparities: A comparative observational study examining the impact of different rurality classifications for health research in aotearoa new zealand. *BMJ Open*, **13**(4), e067927.

Williams, NT, and Díaz, I (2021) *lmtp: Non-parametric causal effects of feasible interventions based on modified treatment policies*. doi:[10.5281/zenodo.3874931](https://doi.org/10.5281/zenodo.3874931).

Wright, MN, and Ziegler, A (2017) ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, **77**(1), 1–17. doi:[10.18637/jss.v077.i01](https://doi.org/10.18637/jss.v077.i01).

Zhang, J, Dashti, SG, Carlin, JB, Lee, KJ, and Moreno-Betancur, M (2023) Should multiple imputation be stratified by exposure group when estimating causal effects via outcome regression in observational studies? *BMC Medical Research Methodology*, **23**(1), 42.