# Causal effect of carbon footprint calculators

Developing a time-series model to evaluate causality

Master thesis in Mathematical Sciences

LOUISE HULTÉN

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2022

Master thesis 2022

# Causal effect of carbon footprint calculators

Developing a time-series model to evaluate causality

LOUISE HULTÉN

UNIVERSITY OF GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mathemathical Sciences
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2022

Causal effect of carbon footprint calculators
Developing a time-series model to evaluate causality
LOUISE HULTÉN



Supervisor: Moritz Schauer, Department of Mathematical Sciences
Examiner: Umberto Picchini, Department of Mathematical Sciences

Master Thesis 2022
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2022

Causal effect of carbon footprint calculators
Developing a time-series model to evaluate causality
LOUISE HULTÉN
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

## Abstract

This master thesis aims to answer whether theory on causality and multivariate time series are relevant tools for questions that might arise in the context of different tracking apps. The context is the mobile application Svalna, which is a research-based carbon calculator designed to help people track and reduce their emissions. It has been shown that information provision can impact behavior, so the central question is whether using the Svalna application impacts the users consumption. I introduce a statistical approach to analyse multivariate time series like those gathered through Svalna. I create a data generation model to test the suggested statistical model. As an intermediate check, the model is used to evaluate a data set from Svalnas users. I conclude that the mechanisms of the developed models function in well-behaved data and the model should be seen as a intermediate step towards a model to analyze real data from Svalna. I think it is a useful approach that can contribute to understanding behavioural change and contribute to better app design.

Keywords: causality, time-series, bayes, sampling, stan, carbon footprint calculator, thesis.

# Acknowledgements

I'd like to thank Moritz Schauer for good supervision throughout this thesis. I'd also like to thank David Andersson and Ross Linscott at Svalna AB for all the inspiration and great discussions.

Louise Hultén, Göteborg, October 2022

# Contents

# 1

# Introduction

## 1.1 Background

Different mobile apps for tracking aspects of your life are increasing in popularity. It's possible to track everything from workouts, sleep habits to environmental impact. Tracking apps allow the user to record their actions and track habits. Such large amounts of data have previously been unavailable. The data provides an interesting and easily accessible source for statistical analysis to complement qualitative psychology studies on habits and behaviors. We can use the data to form hypotheses on non-measured intentions and impacts of the usage of the application.

The Svalna application is a research-based carbon calculator mobile application designed to help people track and reduce their emissions. Svalna uses bank statement data, public records, and user input to estimate the impact of consumption on $CO_2$-equivalents. The $CO_2$ consumption is summarized so the user can see which categories relate to a higher $CO_2$ equivalence, compare their consumption to groups similar to themselves etc.[2]. Carbon calculators are a form of *information provision* and can facilitate behavioral change. In other areas, e.g., energy consumption and grocery shopping, information provision has been shown to impact behavior (e.g. [3, 1]).

Svalna has preliminary results, implying that users reviewing their $CO_2$ consumption in the app for the first time coincide with reduced emissions in the ten weeks following [7]. From this observation, one can raise questions that have a causal ring, e.g.: Are users motivated to change their consumption and then download the Svalna app to measure the change? Or reversed, are users curious about their consumption, get information provided in the app, and decide to change their behavior? And even more interestingly, can these questions be answered with *structural causal process* theory?

This report introduces a statistical approach that can be used to analyze multivariate time series like those gathered in different tracking apps. Inspired by McElreath [9] I approached tracking apps, and causal questions by trying to tell a story that explains how certain data arises. Given the particularities of the context and the questions of interest, I explored theory around causality and multivariate time series. Through this report, relevant statistical tools are introduced and used. Given Svalnas' context, a data generation model is developed to test the suggested statistical model and exemplify some strengths and drawbacks. Finally, as an intermediate check, the model is used to evaluate a data set gathered from Svalnas users to see if conclusions can be drawn with this approach.

### 1.1.1 Data collection through the Svalna app

The Svalna app uses three data sources to estimate its users' weekly carbon consumption. The app presents data to the user split into four categories concatenated over one month, six months and one year. The categories are 1) transportation 2) residential energy 3) food 4) goods and services.

The data in the database is summarized and stored in weekly measurements on 27 sub-categories. For each week there is a total spending in SEK and a calculated $CO_2$-equivalence in each consumption category. The $CO_2$ equivalence factors are based upon national registers and the categorization of bank transactions. For example, car fuel and how housing is heated up are fetched from national registers. For further details on data collection and calculation, see [2].

If there is no bank connected i.e. no bank data is available, the $CO_2$ measure is solely based on average estimates and national registers. The app fetches historical bank statement data each time a user connects to their bank. Only if the user connects their bank at least twice with some delay there will be data available on how using the application might change behaviors.

## 1.2 Question

This thesis aims to answer whether causality and time series are relevant tools for the questions that might come up in this context. The goal of the work is an intermediate statistical model that works on well-behaved data to answer questions related to the context around how and if using the Svalna-application impacts behavior. Explicitly stated, the question I revolve around is: Is there a causal connection between app download to $CO_2$-consumption?

# 2

# Theory

## 2.1 Graphical notation and structural causal models

A (partially directed) graph $\mathcal{G}$ for a graphical model consists of *nodes* $V$ representing random variables $X^1$, $X^2$, . . . , and *edges*, $E$, connecting those nodes, (see an example in figure 1). An edge between two nodes can be *undirected* when the direction of the (causal) relationship between the variables is undetermined, or *directed* when it is known.

A node with an edge that points to another node is called a *parent* of that node. In figure 1 the random variable $X^1$ is a parent of the random variable $X^2$ and $X^2$ is a child of $X^1$. Parents of parents are called *ancestors*. From here on, statistical propositions presented apply to *directed acyclic* graphs (DAGs).

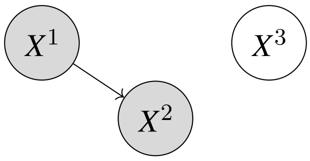



**Figure 1:** A graph with three nodes $V = \{X^1, X^2, X^3\}$ and one edge $E = \{(X^1, X^2)\}$. $X^1$ and $X^2$ are shaded which indicates that they are observed while $X^3$ is white indicating that it is unobserved, also called *latent*. This graph could be drawn from the structural causal model $X^1 = f_1(\epsilon_1)$, $X^2 = f_2(X^1, \epsilon_2)$, $X^3 = f_3(\epsilon_3)$.

A *structural causal model*, (SCM) is a set of variables $V$ and a set of functions $f$ that describe how values are assigned to the variables in $V$. Each SCM has a corresponding graphical model that is useful for communicating and analyzing models. In a graph corresponding to an SCM, the nodes $V$ represent the variables and the edges $E$ describe the qualitative content of the functions $f_v$.

A variable $X$ is a *parent* of Y if X appears in its function $f_Y$, This is denoted by $Y \in pa(X)$.

Using graphical notation for SCM gives that for every $X_v \in V$, and writing $f_v$ for $f_{X_v}$:

$X_v = f_v(pa(X_v), \epsilon_v)$ where $\epsilon_v$ is i.i.d. according to some distribution. If $pa(X_v) = \emptyset$, naturally $X_v = f_v(\epsilon_v)$.

Using the language of graphs one can easily express, draw, the qualitative assumptions of a world in which some data is generated (and measured) and then use that to formulate an SCM. If the underlying causal structure of some data is unknown, we can draw hypotheses graphs, formulate some SCMs and compare them.

## 2.2 Causality

Knowing the causal structure of how some data are produced is the same as understanding the scientific reality and how to manipulate it.

**Definition.** A variable $X$ is a *direct cause* of variable $Y$ if $X$ is a parent of $Y$ in $\mathcal{G}$. $X$ is a *cause* of $Y$ if it appears in the set of ancestors of $Y$.

This definition referring to graphs translates to SCMs. A variable $X$ is a *direct cause* of variable $Y$ if $X$ it appears as a factor in function $f_Y$.

## 2.3 Some semantics around probability

Hereafter, the concept of probability density functions will be used. For that, some notions around probability will be useful. First, we introduce the concepts of joint probability and conditional probability to allow further derivation. The *joint* probability density $p(x, y)$ is the probability density of a pair of random variables $X, Y$. The *marginal density* of $X$ is $p(x) = \int p(x, y)dy$ (or $= \sum_y p(x, y)$.) The *conditional probability density* $p(x \mid y)$ is the probability density of $X$ *given that* $Y = y$ and is defined as:

$$p(x \mid y) = \frac{p(x, y)}{p(y)} \tag{2.1}$$

Two variables $X$ and $Y$ are *independent* if for their densities $p(x, y) = p(x)p(y)$. They are *conditionally independent given a random variable* $Z$, if

$$p(x, y \mid z) = p(x \mid z)p(y \mid z)$$

for all $z$.

### 2.3.1 Bayes rule

Assume $X_1, \ldots, X_d$ are random variables and let $p(X_1, \ldots, X_d)$ be their joint density function (using *Bayesian notation.*)

Let nodes in $\mathcal{G}$ in figure 1 be an example, so the nodes are random variables. Then a corresponding SCM (and any data generated from this model) would contain $f_{\mathcal{G}} = \{p(x_1), p(x_2 \mid x_1)\}$. By conditional probability (2.1) it's possible to formulate the joint density:

$$p(x_1, x_2) = p(x_2 \mid x_1)p(x_2) \tag{2.2}$$

Equivalently, according to (2.1) we can write

$$p(x_1, x_2) = p(x_1 \mid x_2)p(x_1) \tag{2.3}$$

Solving (2.2) and (2.3) for $p(x_1 \mid x_2)$ gives Bayes rule:

$$p(x_1 \mid x_2) = \frac{p(x_2 \mid x_1)p(x_1)}{p(x_2)} \tag{2.4}$$

Notice how there now is an expression for the "flipped" density compared to what is "readable" of $\mathcal{G}$ in figure 1. Bayes rule provides a formula to determine conditional probabilities (and densities) using the factors on the right-hand side.

As Bayes rule is fundamental to continued theory, I'll introduce the common names of the factors in the rule. The conventional names of the components in the numerator are better suited when Bayes rule is written as the probability of some parameters $\boldsymbol{\theta}$ in a model given data $\boldsymbol{y}$.

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})}$$

The denominator $p(\boldsymbol{y})$ is called the *normalizing constant* as it serves to normalize the numerator to a proper probability distribution. The left-hand side of Bayes rule is called the *posterior* distribution of $\boldsymbol{\theta}$. $p(\boldsymbol{\theta})$ is called *prior distribution* of $\boldsymbol{\theta}$, and $p(\boldsymbol{y} \mid \theta)$ is called *likelihood function*.

### 2.3.2 The backdoor criterion

To identify the causal impact of one variable on another variable, we need to identify dependencies of other variables. This informs whether and how the assessment needs to control for confounding sets. The backdoor criterion is an "algorithm" to, via a graph, identify the correct adjustment set for any causal relation in a graph [5, p.79]. The backdoor criterion is used to identify whether a set is sufficient to identify some causal path $p(y \mid x)$. The idea is to make sure that all backdoors between the variables of interest are "blocked". The definition applies to an *ordered* set of variables $(X, Y)$, ordered means $X$ is an ancestor or parent of $Y$,

**Definition.** Given an ordered pair of variables $(X, Y)$ in a directed acyclic graph, a set of variables $Z$ satisfies the backdoor criterion relative to $(X, Y)$ if

- no node in $Z$ is a descendant of $X$ and
- $Z$ blocks every path between $X$ and $Y$ that contains an arrow into $X$.

If $Z$ satisfies these criteria, the causal impact of $Y$ on $X$ is identifiable *given* $Z$. If there is an open back door, it needs to be blocked, for example by conditioning on a variable on the backdoor path.

## 2.4 Time adapted structural causal model (SCM)

Now, back to the models. Introducing time to a SCM gives a *time-adapted* SCM. Assume a graph where each node $X_v$ in $V$ can be mapped to a time $t \in \mathbb{Z}$. As naturally, there is no causal impact moving backwards in time, the existence of this mapping $V \longrightarrow \mathbb{R}_{\mathbb{Z}}$ reduces the set of possible arrows in E so that

$$(X^i_s, X^j_t) \in E \text{ only if } s < t. \tag{2.5}$$

In SCM notation it holds that for a *time-adapted* SCM $X^j_s \in pa(X^i_t)$ can only hold if $s < t$.

For an edge $(X^i_s, X^j_t)$, the time $\tau$ an edge "passes" $\tau = t - s$ is called lag.

Further, if we assume that every variable is measured at every discrete time-step $t \in \mathbb{Z}$, we get a *structural causal process* (SCP) with variables

$$V = \{\mathbb{X}(t)\}_{t\in\mathbb{Z}} \tag{2.6}$$

### 2.4.1 Time-series graphs and causal stationarity

Time-adapted SCMs can be graphically represented by a *time series graph* TSG where each node only has parents in earlier time steps. Let $V$ be a set of nodes, representing $\mathbf{X}(t)$ a vector of length $N$, $\mathbf{X}_t = \{X^1, X^2, ...X^N\}$. Then a TSG consists of $\mathbb{Z}$ copies of $\mathbf{X}$ with directed edges E only in the temporal direction according to equation (2.5), see an example in figure 2. Where the $i$-th ($i = (1, N)$) variable at time $t$ is noted $X^i_t$.

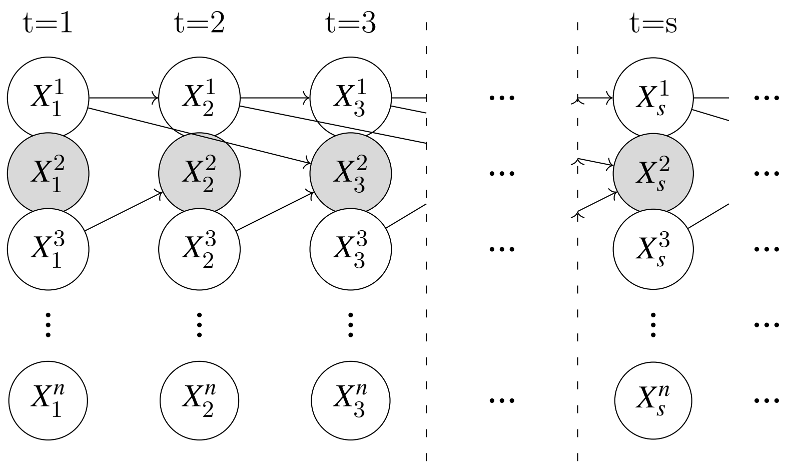



**Figure 2:** An example TSG with copies of $V = \mathbf{X}(t) = \{X^1, ...X^n\}$. The longest edge passes one time step so the lag is $\tau = 2$.

A TSG is called *causally stationary* if it follows from $(X^i_s, X^j_t) \in E$ that $(X^i_{s+\tau}, X^j_{t+\tau}) \in E$ for every $\tau \in \mathbb{Z}$. [10] For *causally stationary* TSGs we can formulate state-equations, a SCP, for all $X^i_t$, $t \in \mathbb{Z}$. As in the stationary case, nodes $V$ correspond to variables and edges to functions. Here, the nodes in each time step correspond to state-vector $\mathbf{X}_t$ and the edges to functions $f_t^n$.

A SCP $V = \{\mathbf{X}(t)\}_{t\in\mathbb{Z}}$ is a process specified by functions $\{f_t^i\}$ and random variables $\{\epsilon_t^i\}$ according to

$$X_t^i = f^i(pa(X^i), \epsilon_t^i))$$

Where $\{\epsilon_t^i\}$ are assumed mutually and serially independent.

Parameters of $f_t^i$ can also depend on time. Causally stationary does *not* assume $f_s^i = f_t^i$ for $s \neq t$. For models in this text though assume $f_s^i = f_t^i$ for all $s \neq t \in \mathbb{Z}$ from here on.

A simple example SCP with linear functions $f^i$ and with maximum lag $\tau = 1$ is exemplified in equation (2.7).

$$X_t^i = f^i(\mathbf{X}_{t-1}) = \sum_j q_{ij}\mathbf{X}_{t-1}^i + \epsilon_t \tag{2.7}$$

$\{\epsilon_t^i\}$ are assumed mutually and serially independent. The parameters of the model are the non-zero entries in a square matrix $Q$ (with entries $q_{i,j}$, $i, j \in (1, N)$) where there is an edge in the corresponding TSG. Each node in such a process has parents *only* in the directly preceding time-step.

Note that functions $f^i$ can take any form and contain many parameters. Linear functions allow us to interpret the causal effect as the slope of how one parameter changes another.

## 2.5 Causal discovery

So far, under given assumptions, it's possible to determine independencies among the variables in a given DAG $\mathcal{G}$. In a data-set $\mathcal{D}$ though, there might exist distributions of those variables that show other, additional independence relations when tested. For example, in a linear model, parameters on mediated (i.e not direct) causal paths are such that they cancel each other out. To avoid this trap and still be able to draw conclusions we *assume faithfulness*. A data-set $\mathcal{D}$ and a graph are said to be *faithful* to one another if all and only the conditional independence relations true in $\mathcal{D}$ are entailed by $\mathcal{G}$.

Well put by Spirtes, [11, p.9] "Informally, the faithfulness condition can be thought of as the assumption that conditional independence relations are due to causal structure rather than to accidents of parameter values."

Now, under the assumptions up to here, we can "simply" search for correlations to draw conclusions about whether a causal connection exists.

More intricate methods to estimate the size and shape of the causal effect in an SCP, direct or indirect, exist. However, they are outside the scope of this thesis.

## 2.6 Comparing models

By assuming faithfulness and analyzing the graph for independencies it is possible to estimate a causal effect (and determine its' nonexistence). For an unexplored real-world phenomenon, the true graph and SCM might be unknown. To complement the above described graphical search there are methods of comparing the fit of different models to a given data-set. One is introduced below.

### 2.6.1 Bayes factor

Bayes factor is a measure to compare how well two hypothetical models describe some data. Assume that some data $\boldsymbol{y}$ is generated by one of two models $\mathcal{M}_1, \mathcal{M}_2$. Then the posterior probability of one model being true is:

$$p(\mathcal{M}_1 \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \mathcal{M}_1) \cdot p(\mathcal{M}_1)}{p(\boldsymbol{y} \mid \mathcal{M}_1)p(\mathcal{M}_1) + p(\boldsymbol{y} \mid \mathcal{M}_2)p(\mathcal{M}_2)}$$

This generalizes to more than two models, of course, but in the case of two models, the Bayes factor is defined from the posterior odds, the fraction between the posteriors of the two models. Using Bayes rule from above fraction can be expressed as:

$$\frac{p(\mathcal{M}_1 \mid \boldsymbol{y})}{p(\mathcal{M}_2 \mid \boldsymbol{y})} = \frac{p(\boldsymbol{y} \mid \mathcal{M}_1)}{p(\boldsymbol{y} \mid \mathcal{M}_2)} \cdot \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}$$

Here $p(M_i)$ is the prior belief that each model is true. When assuming that the prior beliefs for each model is equal, the last factor equals 1, then Bayes factor is the ratio of the marginal likelihoods. Bayes factor noted $B_{1,2}$ means is is in favor of $M_1$. [8]

$$B_{1,2} = \frac{p(\boldsymbol{y} \mid \mathcal{M}_1)}{p(\boldsymbol{y} \mid \mathcal{M}_2)} \tag{2.8}$$

The marginal likelihood (or normalizing constant) is obtained by marginalizing the model parameters:

$$p(\boldsymbol{y} \mid \mathcal{M}_1) = \int_{\boldsymbol{\theta}} p(y \mid \boldsymbol{\theta}, \mathcal{M}_1) p(\boldsymbol{\theta} \mid \mathcal{M}_1) d\boldsymbol{\theta}$$

The marginal likelihood is rarely tractable (except in small models).

### 2.6.2 Bridge sampling

Bridge sampling is a method for estimating normalizing constants using a set of samples from a posterior distribution (i.e from a Markov Chain Monte Carlo sampling result). For full details on the algorithm see[4].

Using the bridge-sample results, implemented in the package `bridgesampling` for `R` it's possible to get the Bayes factor to compare models. [4]

### 2.6.3 Sampling and implementation

Estimating the posterior can be done through sampling. There are multiple useful tools with efficiently implemented statistical algorithms and functions. The one used for this thesis is Stan[12] (through rstan). The default sampling algorithm in Stan is the No-U-Turn Sampler (NUTS). NUTS is an extension to Hamiltonian Monte Carlo where the step size and the step length are set through a recursive algorithm, i.e requiring no hand-tuning of sampling parameters at all. See [6] for full details.

#### 2.6.3.1 Evaluating sampling run

To evaluate a sampling run there are useful statistics, R_hat and n_eff output by Stan. The statistics are used in combination with inspecting plots (histograms most often) of the samples and evaluating if the results are expected.

R_hat is a measure of the difference between the sampling chains. Stan recommends that no parameters with R_hat > 1.05 are used as this indicates that the chains have not converged and mixed. This can also be seen by inspecting the sampling chain plots. The plots are also helpful in understanding problems that might have arisen.

n_eff is an estimate of the effective sample size of each parameter. This is estimated by Stan using the finished samples of each chain. Since the samples within each chain correlate, this measure represents the number of independent draws, i.e with sufficient estimation power. Stan's user guide recommends viewing n_eff > 100 as enough for a reliable posterior sample.

# 3

# Model and sampling

## 3.1 Model of reality

This section introduces the variables present in the model of Svalnas' context. See graphs in figure 3 for a stationary view and 5 for the SCP.

Firstly, the Svalna app presents a real value of *kg $CO_2$-equivalents per week* to the user, represented by the node and variable $\boldsymbol{C}$. This is the response variable of interest in this work.

Second, a binary variable, $\boldsymbol{A}$, for whether or not an individual has reviewed their personal carbon footprint by connecting to their bank through the app.

The model includes the variable $\boldsymbol{I}$, representing income. As socioeconomic status determines lifestyle, which correlates with how much miscellaneous consumption a person does.

One could argue that any action taken by a human is due to their motivation in some way. There are intrinsic and extrinsic motivators, but all are mediated by the human carrying out an action. Therefore, an arrow directly from the *information provision*, $\boldsymbol{A}$, to $\boldsymbol{C}$ could not really exist. In this thesis, I assume this is not the case, but that $\boldsymbol{I}$ and $\boldsymbol{A}$ variables *include* the human response of how much they are impacted by knowing them.

There are inevitably some unobserved variables present in this setting, especially due to the unpredictable human in the model. To represent some individual determinants, the model has the latent individual factor *Motivation to change*, $\boldsymbol{M^i}$. One could imagine this being friends talking about environmental issues inspiring a person to change something in their habits. It could also be government subsidies for sustainable energy or transport options. There is no proxy for factors like this available in the current data set.

There are more variables that make up the final $CO_2$-equivalence measure in the app where users have little agency to change the value. For example, each transaction in SEK is automatically categorized and used to calculate the $CO_2$-equivalence. To reduce the error in the future, one could divide $CO_2$ consumption into separate categories and include their factors in the model. Those underlying mechanisms of the Svalna app are left out of scope from this work and I will rely on the summarized $CO_2$-value and treat the rest as measurement noise.

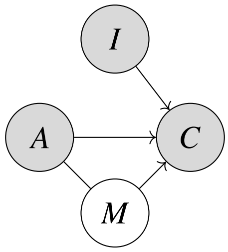
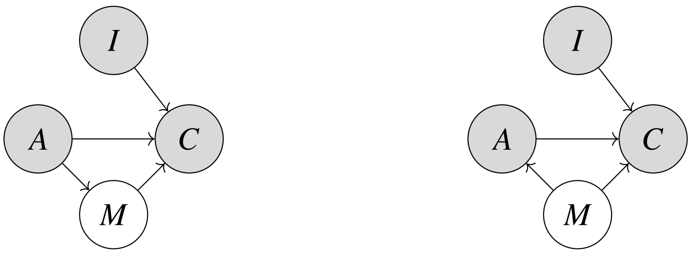



**Figure 3:** Graph with four variables: $I$ for income, $C$ for $CO_2$-*equivalents*, $M$ for motivation and $A$ for app-download. The graph shows the hypothesized relationship between the variables. Note that this figure has a undirected edge between $A$ and $M$ so it is not a DAG.

### 3.1.1 Analysing the graph

The edge of interest in this model is $(A, C)$, i.e the *direct effect* of "using the app" $(A = 1)$ to $C$.

There are two sub-graphs, DAGs, possible by the graph in figure 3, namely those in figure 4. From a causal perspective, if 4a is true where A impacts M the causal effect from $A$ to $C$ could simply be estimated as the joint effect, through M and direct, using regression. The same is true for the case where there is *no connection* between $A$ and $C$.




**(a)** Graph where A causes M. **(b)** Graph where M causes A.

**Figure 4:** Two possible DAGs that could make up the hypothetical model in figure 3.

Based on the interpretation of the model as each variable *includes* the human subjective filter, the second alternative 4b is of interest. It requires a more intricate application of the DAG using the back-door criterion. The analysis follows.

In figure 4b, the back-door path $A - M - C$ is open since $M$ is unobserved. So, regressing $A$ on $C$ would include any effect via $M$. Now, as the data is recorded weekly it's natural to introduce time to this model. Introduction time and assuming it is causally stationary gives the graph in figure 5. The raised index $i$ should now be interpreted as one *individual* person or user of the application.

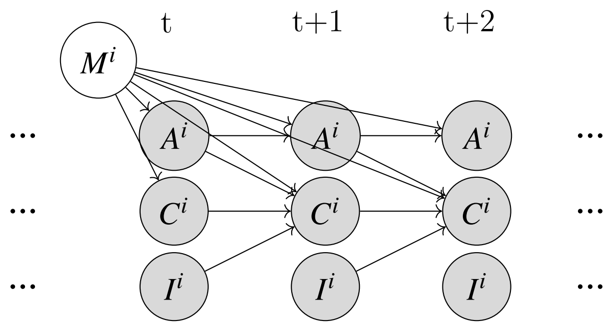



**Figure 5:** TSG of the model used. This graph also shows the true causal model of the simulated data, described in section 3.2.1. $M^i$ does not vary with time as the other variables to. Any exogenous variables and error terms are left out of the graph.

Now, if we interpret the motivation factor as affecting how eager, or likely, an individual is to download the app. Then, once the app is downloaded, the edge $M \rightarrow A$ becomes obsolete since there is no more impact. So, if the data contains records of before and after download, i.e when $A = 0$ switches to $A = 1$ then this mechanism can be programmed into the statistical model and the edge $(A_t, C_{t+1})$ can be estimated.

## 3.2 Data generating model

In order to develop statistical models and to test run them, a simulated data set with known causal structures were created. This is designed to help exemplify causal phenomena that are assumed present in reality. A data-generating model with a "true" causal structure that simulates data is called a *world*. Each *world* simulated here contains $I = 10$ individual time series of $T = 100$ weeks.

### 3.2.1 Parameters and mechanisms of the world

The *world* is developed to allow assumptions of faithfulness and causal stationarity. It's limited to maximum lag $\tau = 1$ and contains variables corresponding to those described in section 3.1 above. The graph corresponding to this world is the same as the model (figure 4b) above.

The true state-vector $Y_t^i = (A_t^i, C_t^i, I_t^i, M^i)$ in the world includes observed variables and the "latent" motivation factor.

For each individual, the initial value $C_{t=0}^i$ is drawn from a normal distribution with mean 0 and standard deviation 1.

$I_0^i$ is drawn from a uniform distribution $U(-1, 1)$ and repeated for every time point until some $t$. To represent some change in an individual's income, a random time point $t$ is drawn, where a $I$ is drawn and repeated through the rest of the time series.

$M^i$ is drawn from a uniform distribution between $U(-1, 0)$ and remains the same throughout the time series.

To let $A_t$ depend on $M^i$, the inverse logit-link (3.1) function is used to map a linear transformation $\beta + M^i$ to a probability (a value between $(0, 1)$) which is used to draw from a Bernoulli distribution. Once $A = 1$, i.e. when the app is downloaded, this relationship is obsolete.

$$\text{logit}^{-1}(\alpha) = \frac{1}{1 + \exp(\alpha)} \tag{3.1}$$

The initial $A_{t=0} = 0$ and the continued time-series of A is simulated as follows:

$$A_t = f_A(A_{t-1}, M, \epsilon_A) = \begin{cases} \sim \text{Bern}(\text{logit}^{-1}(\beta + M)) & \text{if } A_{t-1} = 0 \\ 1 & \text{if } A_{t-1} = 1 \end{cases}$$

The parameter $\beta = -2$ was chosen to spread the times of App download throughout the time series. Since $M^i = (-1, 0)$ this gives Bernoulli parameter range: $\text{logit}^{-1}(\beta + M) = [0.047, 0.11]$. The Bernoulli parameter is the probability of $A_t = 1$ at each time-step where $A_{t-1} = 0$.

To simulate the rest of the time-series, where $C_t$ depends linearly on the variables from the previous time-step $C_t = f_C(Y_{t-1}, \epsilon_C))$. A transition vector is used containing entries representing the slope of the impact.

$$\begin{cases} C_t^i = & \pi_1 A_{t-1}^i + \pi_2 C_{t-1}^i + \pi_3 I_{t-1}^i + \pi_4 M^i + \sigma_C \epsilon_t \\ \\ \boldsymbol{\pi} = & (0.5, 0.5, 0.5, 1)^\top \\ \sigma_C = & 0.33 \end{cases}$$

where $\epsilon$ is drawn independently from a normal distribution with mean 0 and standard deviation 1. $\sigma_C$ is a scale parameter that would allow exploration of different magnitudes of noise. 0.33 was chosen as a level of noise since other variables are near standardized values, this allows us to visually see the impact of for example $I$. Entries in $\boldsymbol{\pi}$ should stay in range $(-1, 1)$ to avoid variables going to infinity. 0.5 was chosen to be able to visually identify then impact.

See a plot of the final generated data set in figure 6.

## 3.3 Sampling models

Each SCM can be translated to a statistical model in `R` for which sampling is done, called *sampling model*. The sampling model used corresponds closely to the data generating model described above. The assumed observed variables $\{C, I, A\}$ are

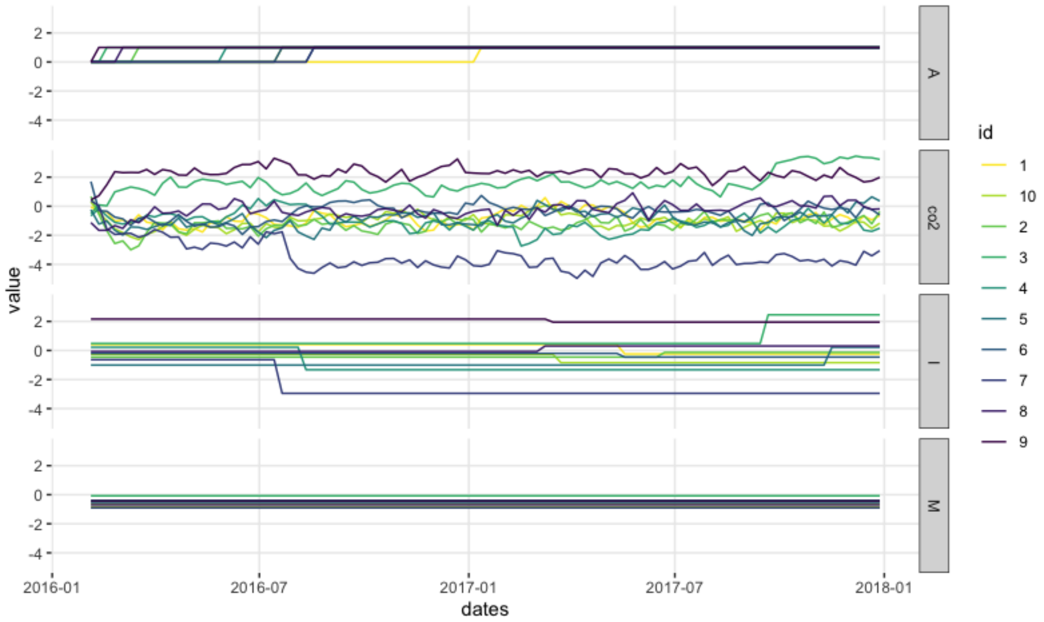



**Figure 6:** Plot of the generated data set. The data set contains $I = 10$ individual time series, each of 4 variables spanning $T = 100$ time points, representing weeks. The week-representation is unimportant and only due to matching the structure of generated data and real data so that any implementations work for both sets of data.

passed as data along with important constants describing the data, which are needed for the sampling to run smoothly $(I = 10, k = 4, T = 100)$.

### 3.3.1 Statistical model

To model the arrows going to $C_t$, i.e the linear function $f_C$, a Normal distribution is used. The mean is the linear function and the standard deviation represents the scale factor $\sigma_C$. The parameter vector $\boldsymbol{\pi}$ represents the $\boldsymbol{\pi}$. See here:

$$C_t^i \sim \mathcal{N}(\tilde{\pi}_1 A_{t-1}^i + \tilde{\pi}_2 C_{t-1}^i + \tilde{\pi}_3 I_{t-1}^i + M^i, \quad \tilde{\sigma}_C) \tag{3.2}$$

Note that $\tilde{\pi}_4$ is omitted. This is to simplify the estimation of $M^i$ since otherwise, the starting point of the sampling would decide the balance of those factors. $\pi_4 = 1$ in the simulated world so $\tilde{\pi}_4$ is excluded here.

Further, the mechanism of $(M_t, A_{t+1})$ is programmed the same way as it is in the world. $\tilde{\beta}$ is the corresponding parameter to $\beta$ above.

$$\begin{aligned}&\text{if } A_{t-1} = 0 :\\&\qquad A_t \sim \text{Bern}(\text{logit}^{-1}(\tilde{\beta} + M^i))\end{aligned}$$

When sampling, the parameters to be estimated are:

- $\tilde{\pi}_{1,2,3}$
- $\tilde{\beta}$

- $\tilde{\sigma}_C$
- $\tilde{M}^i$ for all $i \in I$.

Each parameter has a specified prior distribution for which the sampling is initiated in each chain. The priors are defined as follows:

$$\begin{aligned}
\tilde{\pi}^j &\sim \mathcal{N}(0, 10) \text{ for } j = \{1, 2, 3\} \\
\tilde{M}^i &\sim \mathcal{N}(0, 10) \text{ for all } i \in I \\
\tilde{\beta} &\sim \mathcal{N}(0, 10) \\
\tilde{\sigma}_C &\sim \mathcal{N}(0, 10)
\end{aligned}$$

### 3.3.2 Running generated data and sampling model

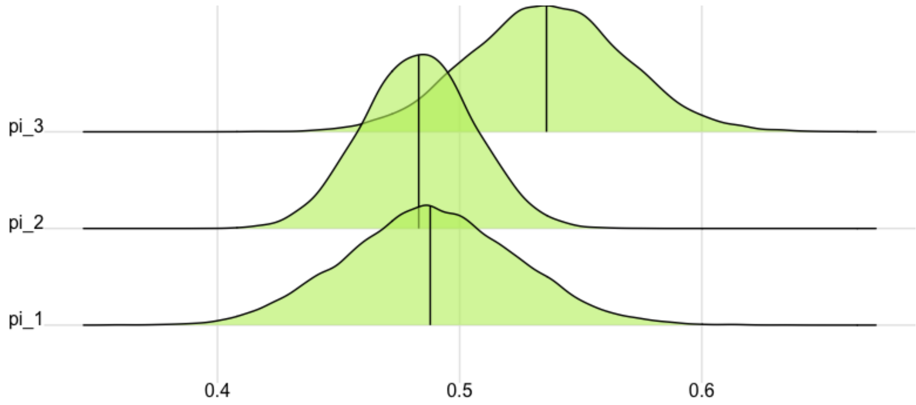
Below are samples plotted as densities from running the sampler in Stan using two chains of 15000 samples each. See the summary table of all variables in Appendix A. All the chains converged, indicated by all Rhat values are 1, and the effective number of samples span between 254 and 28398.

The shape of the samples is typical curves of normal distribution. $\tilde{\pi}_1$ and $\tilde{\pi}_2$ means are below the real $\pi$-values while $\tilde{\pi}_3$ is higher. $\pi_1$ and $\pi_2$ are the weights of $A_{t-1} \rightarrow C_t$ and $I_{t-1} \rightarrow C_t$ respectively, so $\pi_3$ is $C_{t-1} \rightarrow C_t$.

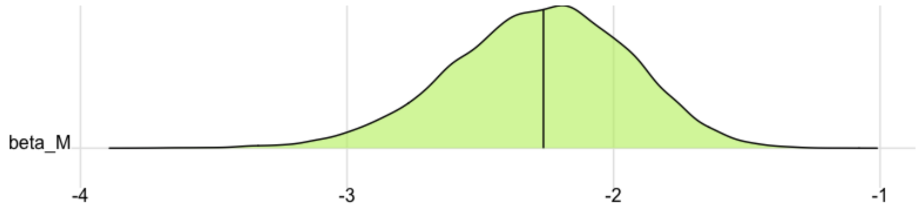
The scale variable of the noise, $tilde\sigma_C$ has a sample mean matching the value of $\sigma_C$, with a standard deviation of 0.01.

The samples of $\tilde{\beta}_M$ have a standard deviation of 0.33 and the sample mean $-2.28$ is slightly lower than the real corresponding value $\beta = -2$. This would estimate the probability parameter of the Bernoulli distribution modeling $M_i$ to be smaller than the real value.

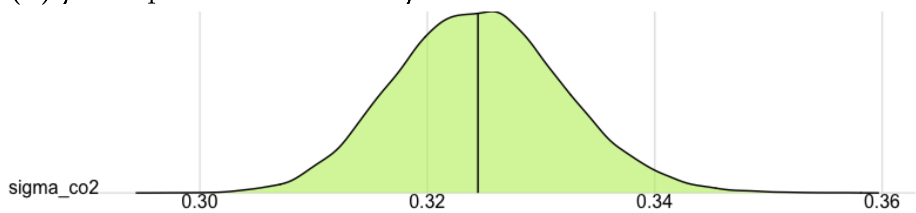
The "true" values of $M^i$-factors are consistently within one standard deviation of the sampled mean of each $\tilde{M}^i$. There is no obvious pattern of over vs underestimation.



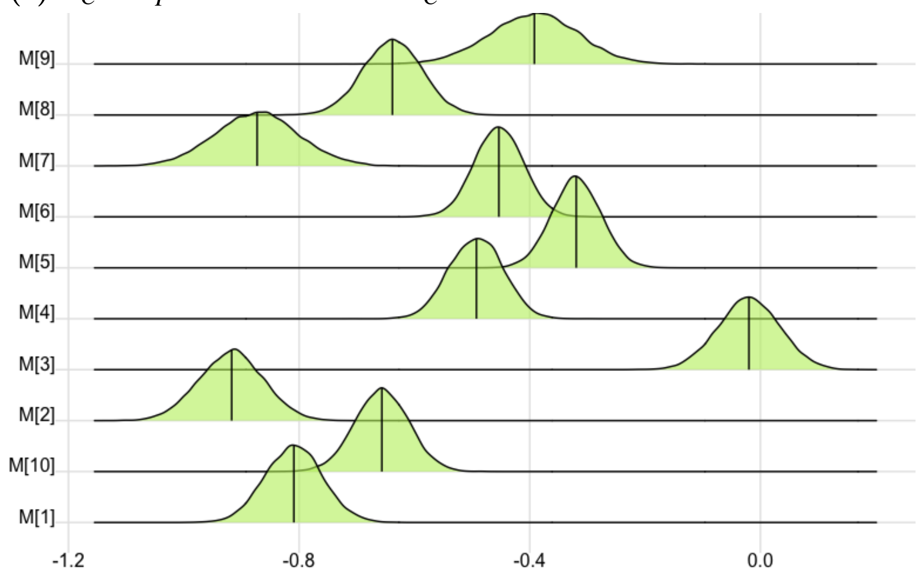

**(a)** $\tilde{\pi}$ samples. True values are $\pi_{1,2,3} = 0.5$.




**(b)** $\tilde{\beta}$ samples. True value is $\beta = -2$.




**(c)** $\sigma_C samples.$True value is $\epsilon_C = 0.33$




**(d)** $\tilde{M}^i$ samples. True values are $M_1 = -0.84, M_2 = -0.91, M_3 = -0.07, M_4 = -0.49, M_5 = -0.39, M_6 = -0.50, M_7 = -0.89, M_8 = -0.69, M_9 = -0.43, M_{10} = -0.64$.

**Figure 7:** Density plots of 15000 samples (after warm up).

### 3.3.3 Alternative world and statistical model

In an experimental setting to evaluate an app, it's common to recruit users to download the app and ask for feedback. Data from such an experiment would *not* contain any causal arrow between $M$ and $A$. To investigate if the sampling model $\mathcal{M}_1$ from above can distinguish between such settings, an alternative world and alternative model $\mathcal{M}_2$ was created.

Let's call this the *world 2* and it's simulated in the same way as the world above with the exception that the time of App download is random (drawn from a uniform distribution $U(0, T)$) and not dependent $M$. See the DAG in figure 8.

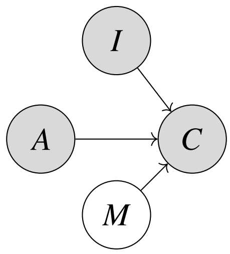



**Figure 8:** Graph showing *world 2*. This world has no edge between $A$ and $M$.

Model 2 only contains the parameters of the linear function causing $C$, according to:

$$C_t^i \sim \mathcal{N}(\pi_1 A_{t-1}^i + \pi_2 C_{t-1}^i + \pi_3 I_{t-1}^i + M^i, \quad \tilde{\sigma}_C) \tag{3.3}$$

The parameters estimated are:
- $\pi_{1,2,3,4}$
- the scale of the noise represented by $\tilde{\sigma}_C$
- $M^i$ for all $i \in I$.

### 3.3.4 Bayes factors

For comparing the models using Bayes factor a bridge sampler was used to obtain an estimate of the marginal likelihood. Running both models on *world 2* -data, the Bayes factor results in the following.

Estimated Bayes factor in favor of *model 2* over *model 1* for the alternative data (*world 2*)

$$\mathcal{W}_2 : \quad \log B_{\mathcal{M}_2, \mathcal{M}_1} = 51.38$$

Running both models on *world 1* results in Bayes factor in favor of $\mathcal{M}_1$:

$$\mathcal{W}_1 : \quad B_{\mathcal{M}_1, \mathcal{M}_2} = 0$$

# 4
# Real data

The real data differ from the simulated data in a few ways. In the real data there are obviously no negative *income*-values or $CO_2$-*equivalence*-values, whereas the maximum values are high and sporadic leaving a skewed representation. See a line-plot of the real data series in Appendix C, Figure B.1 and compare to figure 6 for which we know that the model works well.

It's still a relevant reality check to see how the statistical model acts on the real data, results are described in this chapter.

## 4.1 Data preparation

A set time period was chosen with start date 2017-06-05 and end date 2020-02-24 giving a length of each series $T = 143$. This choice had three reasons: 1) To avoid handling rugged data (e.i to have the same length time-series for all users) and 2) to have the time-series seasonally lined up 3) to avoid unrepresentative change in behavior due to the covid-19 pandemic.

As the app download is essential for this analysis, only users who connected their app three weeks "away" from the start and end dates are included. I.e in $t = (3, T-3)$. The filtered data set consists of $I = 320$ unique users.

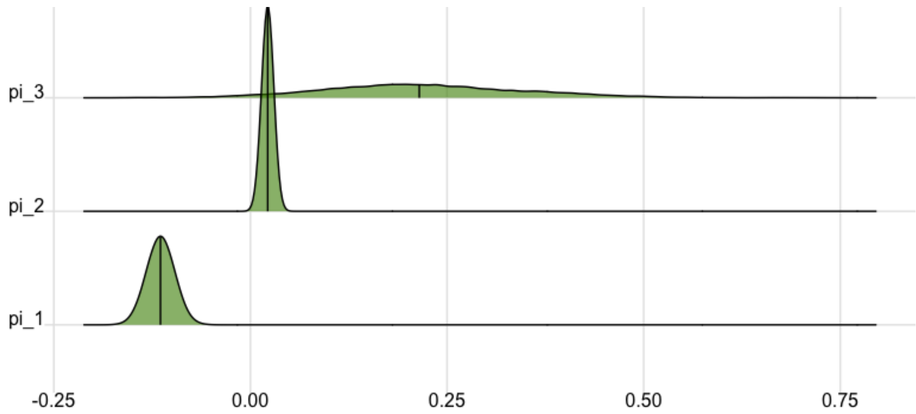
Columns for total $CO_2$-*equivalence*, $C$ and *income*, $I$, are standardised using z-score (4.1). This results in the total data set having a standard deviation of 1 and mean 0.

$$z' = \frac{z - \text{mean}(z)}{\text{std}(z)} \tag{4.1}$$

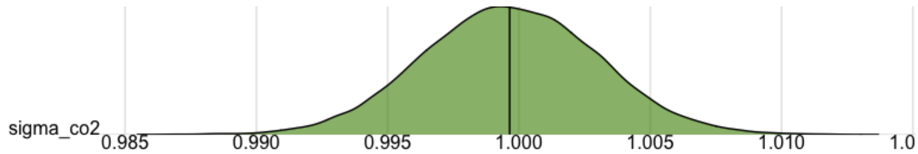
## 4.2 Statistical model and results

The statistical model used for sampling was exactly the one described in 3.3.1. Summarized sampling results for the global variables ($\pi_{1,2,3}$, $\beta$, $\tilde{\sigma}_C$) are printed in table C.2 in appendix C and plotted in figure 4.2 below. Results for all the 320 $M^i$ variables are summarized in Table C.3 in Appendix C.

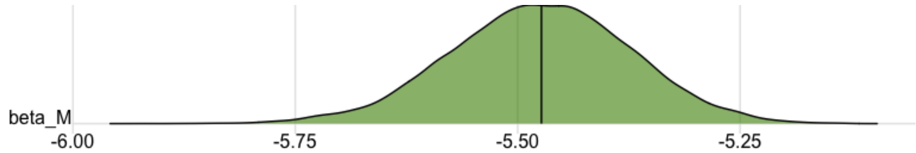
All chains converged, as indicated by the R-hat values all are 1. Further the effective sampling size for $M^i$ span between $n\_eff = (254, 28398)$ (out of 40000 iterations).




**(a)** $\pi$ samples.




**(b)** $\sigma_{CO_2}$ samples.




**(c)** $\beta$ samples.

**Figure 9:** Results from using the real data to sample parameters in the statistical model. Density plots of 15000 samples (after warm up).

The value $\tilde{pi}_3$ estimated mean 0.22 would be the indicator weight quantifying a causal relationship between $A \rightarrow C$ . To analyze the discrepancy of the statistical model and the real data I use the above sampled values as parameters in the data-simulation program described in 3.2.1 and create time-series for ten individuals. Results from this simulation is shown in Appendix C, figure B.2. Comparing figure B.2 and figure B.1 the periodicity or heavy tailed pattern of the real data is clearly not being represented in the model - thereby I refrain from drawing any conclusions on weather there is a causal relationship here.

# 5
# Conclusion

## 5.1 Discussion

The starting point of this thesis was literature on causality, which I explored guided by the context of the Svalna app and focused on investigating the presumed relationship from App-download to $CO_2$-consumption. Using an understanding of the context, statistical tools were selected and gradually built together into a statistical model. The final statistical model, *model 1*, above was preceded by simpler ones, mostly ordered by increasing complexity. The generated data set provides a means to get to know the mechanisms of the theory put together. So far, the mechanisms function together and recognize the patterns in the data as I expect.

See the result presented above as an intermediate step towards reaching a "fitting" statistical model for the real context and data. There might be other mechanisms beyond what is treated in section 2 or just a more complex statistical model that fits depending on the exact goals. More complex could entail using more aspects of the data *and/or* more parameters. In this chapter, some notions worth addressing in the next evolution of the model are discussed.

### 5.1.1 Time

The questions of interest in this context have a natural aspect of time in them. To develop the Svalnas service it might be interesting to know *How long do habit-changes induced by the app hold? When do effects seem to fade?* Runge et al. [10], from whom I have used semantics and definitions, provided the idea to assume the model to be causally stationary. This served as a convenient way to analyze time series but kept the model to a reasonable complexity. In the same paper, they continue to introduce ways to analyze time-windowed effects of events (*impulses*). With this opportunity in mind, choosing to treat the data as a time series seems to be the right choice. Additional to, of course, deciphering the graph in figure 3.1.

### 5.1.2 Choice of variables and functions

So far, I did not use any independence test or data-informed choices but depended solely on literature and understanding of the domain. To continue this work with the goal to estimate behavior changes in $CO_2$ consumption related to app download and app use, I would start from the data and meet half-ways with the domain knowledge approach.

In the next evolution of the model I would consider changing the mechanism of the relationship $(\boldsymbol{CO}_{2_{t-1}}, \boldsymbol{CO}_{2_t})$. The linear relationship used until now passed as a good starting point while analyzing the graph and testing the model and implementation. I would op to use something that fluctuates from high to low with some periodicity. If a person spent less $CO_2$ (or money) one week we'd expect to see an increase in spending soon, and vice versa. An alternative could be to analyze the data grouped by month, I'd expect that to be more stable. Further the current model is missing the fact that the real data is skewed to low values, everyday normal consumption tend to stay at similar values month by month and then there are spikes for bigger purchases. Think of such spikes as for example the yearly payment for summer vacation, or house purchase. Simulation of the $\boldsymbol{C}$-time-series should result in something like a heavy tailed truncated normal distribution, whereas its a *normal* normal distribution in the current statistical model and simulated world.

Currently, the motivation factor is a variable with a weak scientific meaning. Attempting to represent intrinsic motivation and external factors is relevant in contexts like this. I believe those variables will often show up as confounds to the edge between app engagement and behavior. Above, I've kept the Motivation-factor latent. In the future, to increase the intricacy of the model picking up on more separate un-measured factors that impact human behavior finding some proxy for the factors in the data. An idea for the available data set could be that app engagement (which is measured to some extent) could serve as a proxy for curiosity and motivation.

### 5.1.3 Curse of dimensionality

There is basically no limit to the complexity of a model used to evaluate the context in question here. Mechanisms that have been nearby throughout the work have been to control for seasonality in spending or to allow trends in the motivation factor. Most of the alternatives and developments, both my ideas and those suggested in the literature, would add to the number of parameters in the model. More parameters increase the demand for high qualitative data to draw any conclusion, and, the risk of overfitting. In the case of the Svalna data - the variability among users can be a good way to counter overfitting, as long as the number of individual parameters is kept down.

### 5.1.4 New insights to the Svalna context

The insight into the data from section 4.2 indicate some causal relationship between the variables ($\boldsymbol{A}$, $\boldsymbol{C}$) and ($\boldsymbol{I}$, $\boldsymbol{C}$). I withdraw from drawing any conclusions on the effects as I think some variables in the data should be included in the analysis first - for example, the prevalence of manually updated categorizations, $CO_2$ intensity per SEK spent, etc.

The social aspect of the app, e.g. leaderboards, comparisons to friends, etc. do give an incentive for the user to re-categorize and alter transactions to show lower $CO_2$ footprints. These are all factors to be considered in an analysis of the data. Granted, they may not be bad for guiding people to reduce their consumption but complicate any statistical analysis. Results from biased data should be interpreted exactly as such. Also, the $CO_2$ measure in the app obviously only represents the aspects of life that the user chose to share with the app. For example, if a user has multiple bank accounts, one with which they pay for expensive things like vacations (flight tickets) this will not show up.

## 5.2 Conclusion

The model developed in this thesis can identify parameters as expected in well-behaved data. The sampling of multiple (individual) time series estimating parameters of a mixed model (with individual and global parameters) works. As discussed above, I think the mechanisms used until now are meaningful and will (should) be kept in future models meant to analyze the real data. Although creating a statistical model for the true data and daring to conclude to inform the field of carbon calculators, it likely requires a few more mechanisms and especially to be informed by available data.

I think causality is a relevant perspective to analyze habits and behaviors. Multivariate time series are already commonly used in psychology. Using methods e.g. time-windowed effect estimation presented by Runge et. al. can contribute to understanding behavioral change and contribute to better design of apps to fulfill their goals, as long as the data-collection is well understood in its context.

# Bibliography

[1] Hunt Allcott and Todd Rogers. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10), 2014.

[2] David Andersson. A novel approach to calculate individuals' carbon footprints using financial transaction data – App development and design. *Journal of Cleaner Production*, 256, 2020.

[3] Toke R ; Fosgaard, Alice ; Pizzo, and Sally Sadoff. Do People Respond to the Climate Impact of their Behavior? The Effect of Carbon Footprint Information on Grocery Purchases. *IFRO Working Paper*, 2021.

[4] Quentin F. Gronau, Alexandra Sarafoglou, Dora Matzke, Alexander Ly, Udo Boehm, Maarten Marsman, David S. Leslie, Jonathan J. Forster, Eric Jan Wagenmakers, and Helen Steingroever. A tutorial on bridge sampling. *Journal of Mathematical Psychology*, 81, 2017.

[5] Christopher Hitchcock and Judea Pearl. *Causality: Models, Reasoning and Inference*, volume 110. 2001.

[6] Matthew D. Hoffman and Andrew Gelman. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 2014.

[7] Jakob Enlund. Final Seminar Jakob Enlund, Gothenburg University, 3 2022.

[8] Robert E. Kass and Adrian E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430), 1995.

[9] Richard McElreath. *Statistical Rethinking : a Bayesian course with examples in R and Stan.* Boca Raton : CRC Press/Taylor & Francis Group, 11 2016.

[10] Nicolas-Domenic Reiter, Andreas Gerhardus, and Jakob Runge. Causal inference for temporal patterns. *Preprint: arXiv:2205.15149 [stat.ME]*, 5 2022.

[11] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search.* MIT press, 2000.

[12] Stan Development Team. Stan User's Guide, Version 2.30, 2022.

# A

# Appendix A

## Generated data and statistical model 1

| | mean | se_mean | sd | 2.5% | 25% | 50% | 75% | 97.5% | n_eff | Rhat | True |
|---|---|---|---|---|---|---|---|---|---|---|---|
| pi_1 | 0.49 | 0.00 | 0.04 | 0.42 | 0.46 | 0.49 | 0.51 | 0.56 | 3300.52 | 1.00 | 0.50 |
| pi_2 | 0.48 | 0.00 | 0.02 | 0.44 | 0.47 | 0.48 | 0.50 | 0.53 | 4269.21 | 1.00 | 0.50 |
| pi_3 | 0.54 | 0.00 | 0.03 | 0.47 | 0.51 | 0.54 | 0.56 | 0.60 | 4437.87 | 1.00 | 0.50 |
| sigma_co2 | 0.32 | 0.00 | 0.01 | 0.31 | 0.32 | 0.32 | 0.33 | 0.34 | 12297.28 | 1.00 | 0.50 |
| beta_M | -2.28 | 0.00 | 0.33 | -2.97 | -2.50 | -2.26 | -2.04 | -1.67 | 11814.94 | 1.00 | 2.00 |
| M[1] | -0.81 | 0.00 | 0.05 | -0.91 | -0.84 | -0.81 | -0.78 | -0.71 | 4479.26 | 1.00 | -0.84 |
| M[2] | -0.92 | 0.00 | 0.06 | -1.03 | -0.96 | -0.92 | -0.88 | -0.81 | 4011.41 | 1.00 | -0.91 |
| M[3] | -0.02 | 0.00 | 0.05 | -0.12 | -0.06 | -0.02 | 0.02 | 0.08 | 5793.42 | 1.00 | -0.07 |
| M[4] | -0.49 | 0.00 | 0.05 | -0.58 | -0.52 | -0.49 | -0.46 | -0.40 | 5816.91 | 1.00 | -0.49 |
| M[5] | -0.32 | 0.00 | 0.04 | -0.40 | -0.35 | -0.32 | -0.29 | -0.24 | 6686.60 | 1.00 | -0.39 |
| M[6] | -0.45 | 0.00 | 0.04 | -0.54 | -0.48 | -0.45 | -0.42 | -0.37 | 5738.42 | 1.00 | -0.50 |
| M[7] | -0.87 | 0.00 | 0.07 | -1.01 | -0.92 | -0.87 | -0.83 | -0.73 | 5418.93 | 1.00 | -0.89 |
| M[8] | -0.64 | 0.00 | 0.05 | -0.74 | -0.67 | -0.64 | -0.60 | -0.54 | 4236.89 | 1.00 | -0.69 |
| M[9] | -0.39 | 0.00 | 0.08 | -0.55 | -0.44 | -0.39 | -0.34 | -0.24 | 4795.40 | 1.00 | -0.43 |
| M[10] | -0.66 | 0.00 | 0.05 | -0.75 | -0.69 | -0.66 | -0.62 | -0.56 | 4693.39 | 1.00 | -0.64 |
| lp__ | 579.81 | 0.04 | 2.73 | 573.63 | 578.15 | 580.12 | 581.82 | 584.21 | 6011.17 | 1.00 | |

**Table A.1:** Results using data described in 3.2.1 and statistical model described in 3.3.1. Results are from 2 chains, each with iter=15000; warmup=7500; thin=1; post-warmup draws per chain=7500, total post-warmup draws=15000.

# B

# Appendix B

## Plot of standardized real data

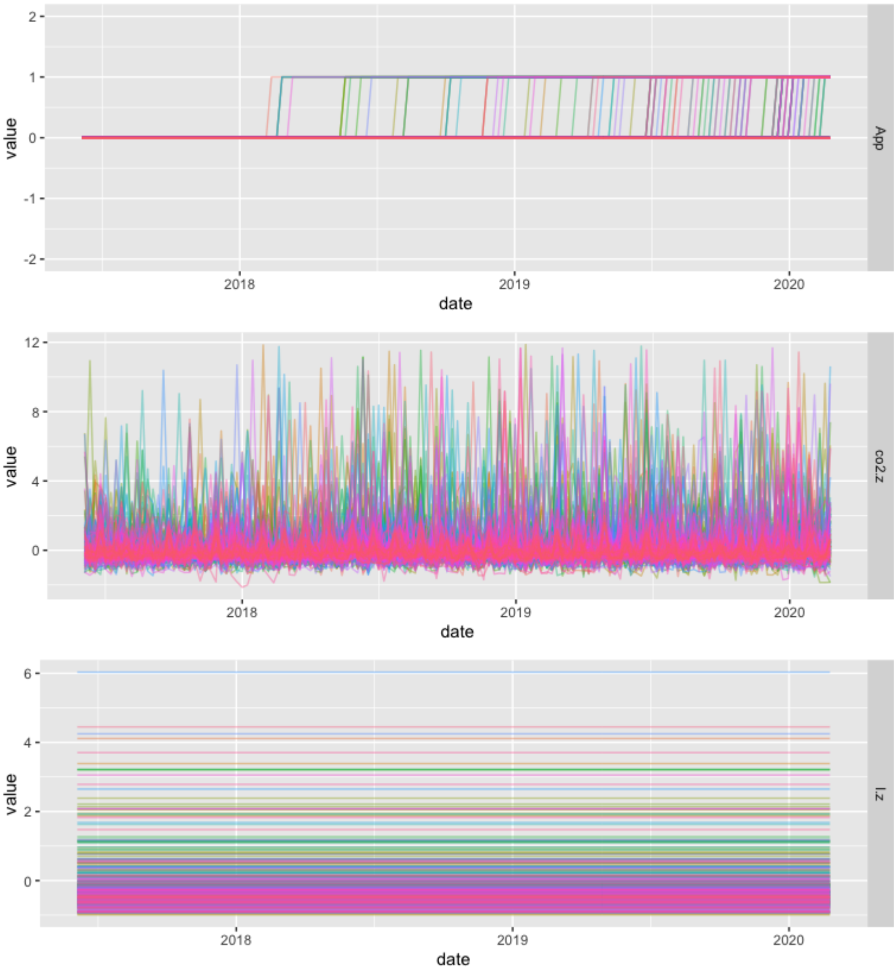



**Figure B.1:** Line-plot of the selected and standardized real data.

## Plot of re-simulated world based on real data sample results.

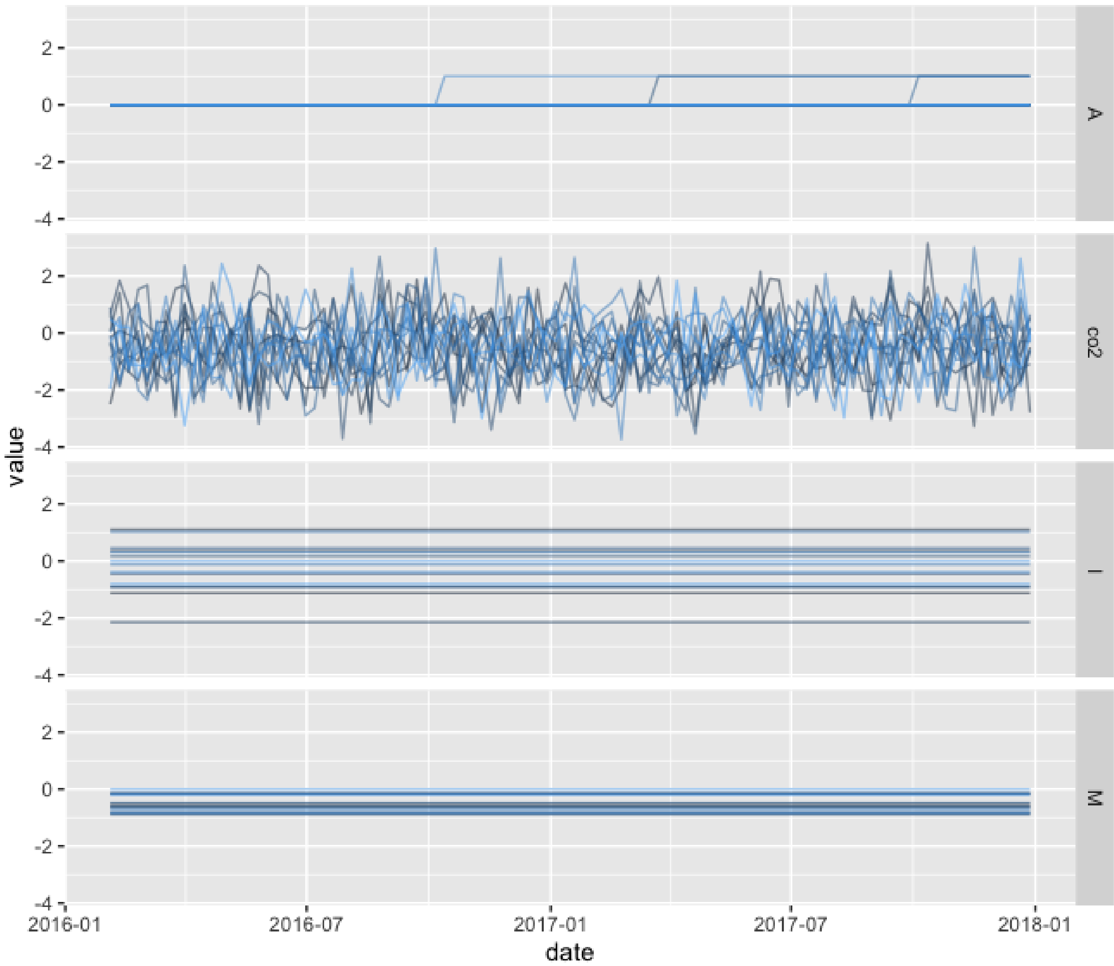



**Figure B.2:** Plot of simulated World 1 (described in 3.2.1) using parameter values obtained by using the real data to sample the parameters in statistical model 1. The parameter values are $\pi_1 \sim \mathcal{N}(-.11, 0.02)$, $\pi_2 \sim \mathcal{N}(-.02, 0.13)$, $\pi_3 \sim \mathcal{N}(0.22, 0.13)$, $\sigma_C \sim \mathcal{N}(1, 0.001)$, $\beta_M \sim \mathcal{N}(-5.47, 0.1)$. (According to summary table C.2 below.)

# C

# Appendix C

## Summary of samples on real data

| | mean | se_mean | sd | %2.5. | %25. | %50. | %75. | %97.5. | n_eff | Rhat |
|---|---|---|---|---|---|---|---|---|---|---|
| pi_1 | -0.11 | 0.00 | 0.02 | -0.15 | -0.13 | -0.11 | -0.10 | -0.08 | 7221.84 | 1.00 |
| pi_2 | 0.02 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 27984.74 | 1.00 |
| pi_3 | 0.22 | 0.01 | 0.13 | -0.01 | 0.13 | 0.21 | 0.30 | 0.48 | 251.82 | 1.00 |
| sigma_co2 | 1.00 | 0.00 | 0.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.01 | 29095.46 | 1.00 |
| beta_M | -5.47 | 0.00 | 0.10 | -5.68 | -5.54 | -5.47 | -5.41 | -5.28 | 4861.69 | 1.00 |

**Table C.2:** Results using real data described in 4.2 and statistical model described in 3.3.1. Results are from 2 chains, each with 2 chains, each 20000 iterations and 100000 warmup iterations. So total post-warmup draws was 20000.

# Real data and statistical model 1

| variable | mean | sd | Rhat | n_eff |
|---|---|---|---|---|
| M[1] | 0.11 | 0.11 | 1.00 | 710.07 |
| M[2] | 0.15 | 0.12 | 1.00 | 480.05 |
| M[3] | -0.12 | 0.11 | 1.00 | 669.27 |
| M[4] | 0.13 | 0.11 | 1.00 | 649.94 |
| M[5] | 0.18 | 0.10 | 1.00 | 1252.87 |
| M[6] | 0.18 | 0.09 | 1.00 | 1330.12 |
| M[7] | 0.21 | 0.11 | 1.00 | 558.53 |
| M[8] | 0.14 | 0.09 | 1.00 | 8220.83 |
| M[9] | -0.32 | 0.25 | 1.00 | 281.97 |
| M[10] | -0.12 | 0.11 | 1.00 | 645.34 |
| M[11] | 0.11 | 0.10 | 1.00 | 707.21 |
| M[12] | 0.08 | 0.10 | 1.00 | 1048.35 |
| M[13] | -0.85 | 0.53 | 1.00 | 258.08 |
| M[14] | 0.02 | 0.09 | 1.00 | 6068.72 |
| M[15] | 0.05 | 0.09 | 1.00 | 2353.88 |
| M[16] | 0.08 | 0.10 | 1.00 | 1099.98 |
| M[17] | 0.03 | 0.09 | 1.00 | 3756.06 |
| M[18] | 0.20 | 0.14 | 1.00 | 377.83 |
| M[19] | -0.09 | 0.10 | 1.00 | 842.65 |
| M[20] | 0.07 | 0.09 | 1.00 | 1457.06 |
| M[21] | -0.11 | 0.11 | 1.00 | 663.90 |
| M[22] | 0.03 | 0.09 | 1.00 | 2756.11 |
| M[23] | 0.10 | 0.10 | 1.00 | 863.66 |
| M[24] | -0.02 | 0.11 | 1.00 | 549.42 |
| M[25] | 0.28 | 0.14 | 1.00 | 368.84 |
| M[26] | 0.23 | 0.11 | 1.00 | 639.92 |
| M[27] | 0.26 | 0.13 | 1.00 | 473.91 |
| M[28] | 0.26 | 0.12 | 1.00 | 481.82 |
| M[29] | -0.64 | 0.44 | 1.00 | 263.25 |
| M[30] | 0.22 | 0.15 | 1.00 | 379.21 |
| M[31] | 0.33 | 0.15 | 1.00 | 364.44 |
| M[32] | 0.04 | 0.08 | 1.00 | 27809.92 |
| M[33] | -0.06 | 0.09 | 1.00 | 1661.89 |
| M[34] | 0.16 | 0.12 | 1.00 | 449.91 |
| M[35] | 0.04 | 0.09 | 1.00 | 21879.01 |
| M[36] | 0.14 | 0.09 | 1.00 | 6799.85 |
| M[37] | 0.25 | 0.12 | 1.00 | 559.38 |
| M[38] | 0.23 | 0.11 | 1.00 | 704.49 |
| M[39] | 0.07 | 0.09 | 1.00 | 1466.40 |
| M[40] | -0.46 | 0.28 | 1.00 | 273.25 |
| M[41] | 0.18 | 0.11 | 1.00 | 597.76 |
| M[42] | 0.24 | 0.12 | 1.00 | 520.50 |
| M[43] | 0.11 | 0.09 | 1.00 | 9280.17 |
| M[44] | 0.25 | 0.11 | 1.00 | 557.88 |
| M[45] | 0.27 | 0.13 | 1.00 | 405.57 |
| M[46] | 0.10 | 0.10 | 1.00 | 780.93 |
| M[47] | 0.13 | 0.08 | 1.00 | 19084.09 |
| M[48] | -0.46 | 0.27 | 1.00 | 278.81 |
| M[49] | 0.12 | 0.10 | 1.00 | 630.07 |
| M[50] | -0.07 | 0.10 | 1.00 | 794.78 |
| M[51] | 0.08 | 0.10 | 1.00 | 976.57 |
| M[52] | -0.09 | 0.13 | 1.00 | 415.47 |
| M[53] | 0.20 | 0.10 | 1.00 | 934.68 |
| M[54] | 0.23 | 0.12 | 1.00 | 527.78 |
| M[55] | 0.26 | 0.12 | 1.00 | 501.93 |
| M[56] | 0.26 | 0.12 | 1.00 | 501.22 |
| M[57] | 0.05 | 0.09 | 1.00 | 1552.97 |
| M[58] | 0.20 | 0.14 | 1.00 | 376.47 |
| M[59] | 0.05 | 0.09 | 1.00 | 21127.71 |
| M[60] | -0.32 | 0.26 | 1.00 | 280.38 |
| M[61] | 0.27 | 0.15 | 1.00 | 369.01 |
| M[62] | 0.20 | 0.11 | 1.00 | 536.11 |
| M[63] | -0.03 | 0.11 | 1.00 | 654.93 |
| M[64] | -0.10 | 0.15 | 1.00 | 378.22 |
| M[65] | 0.14 | 0.09 | 1.00 | 4188.47 |
| M[66] | -0.16 | 0.14 | 1.00 | 404.61 |
| M[67] | -0.09 | 0.12 | 1.00 | 472.24 |
| M[68] | 0.14 | 0.09 | 1.00 | 2786.17 |
| M[69] | 0.22 | 0.10 | 1.00 | 772.16 |
| M[70] | -0.49 | 0.32 | 1.00 | 270.52 |
| M[71] | 0.02 | 0.08 | 1.00 | 27383.43 |
| M[72] | 0.23 | 0.12 | 1.00 | 532.80 |
| M[73] | -0.06 | 0.13 | 1.00 | 430.11 |
| M[74] | -0.13 | 0.17 | 1.00 | 335.25 |
| M[75] | 0.30 | 0.14 | 1.00 | 408.72 |
| M[76] | 0.25 | 0.12 | 1.00 | 525.95 |
| M[77] | 0.14 | 0.11 | 1.00 | 645.39 |
| M[78] | 0.26 | 0.12 | 1.00 | 508.55 |
| M[79] | -0.44 | 0.28 | 1.00 | 275.83 |
| M[80] | -0.49 | 0.30 | 1.00 | 279.35 |
| M[81] | 0.02 | 0.08 | 1.00 | 9725.04 |
| M[82] | 0.12 | 0.10 | 1.00 | 785.34 |
| M[83] | 0.20 | 0.11 | 1.00 | 684.28 |
| M[84] | -0.42 | 0.26 | 1.00 | 282.07 |
| M[85] | 0.13 | 0.12 | 1.00 | 524.72 |
| M[86] | -0.04 | 0.09 | 1.00 | 2377.85 |
| M[87] | 0.32 | 0.14 | 1.00 | 379.61 |
| M[88] | 0.12 | 0.10 | 1.00 | 996.99 |
| M[89] | 0.13 | 0.09 | 1.00 | 10271.57 |
| M[90] | 0.10 | 0.09 | 1.00 | 3238.24 |
| M[91] | -0.17 | 0.13 | 1.00 | 430.58 |
| M[92] | -0.03 | 0.09 | 1.00 | 8180.13 |
| M[93] | 0.14 | 0.12 | 1.00 | 526.13 |
| M[94] | 0.13 | 0.09 | 1.00 | 1293.57 |
| M[95] | -0.13 | 0.14 | 1.00 | 389.10 |
| M[96] | 0.15 | 0.09 | 1.00 | 6532.00 |
| M[97] | -0.24 | 0.17 | 1.00 | 336.25 |
| M[98] | 0.08 | 0.09 | 1.00 | 1403.70 |
| M[99] | -0.66 | 0.42 | 1.00 | 264.39 |
| M[100] | 0.01 | 0.09 | 1.00 | 1426.04 |
| M[101] | 0.11 | 0.09 | 1.00 | 9170.10 |
| M[102] | 0.14 | 0.11 | 1.00 | 518.66 |
| M[103] | -0.14 | 0.17 | 1.00 | 332.08 |
| M[104] | 0.16 | 0.12 | 1.00 | 464.08 |
| M[105] | 0.11 | 0.09 | 1.00 | 4761.84 |
| M[106] | -0.22 | 0.16 | 1.00 | 339.29 |
| M[107] | -0.02 | 0.12 | 1.00 | 530.22 |
| M[108] | 0.00 | 0.09 | 1.00 | 2817.19 |
| M[109] | -0.27 | 0.18 | 1.00 | 320.13 |
| M[110] | 0.16 | 0.13 | 1.00 | 435.34 |
| M[111] | 0.07 | 0.09 | 1.00 | 1176.19 |
| M[112] | 0.17 | 0.13 | 1.00 | 438.76 |
| M[113] | 0.05 | 0.09 | 1.00 | 8991.37 |
| M[114] | -0.60 | 0.42 | 1.00 | 261.45 |
| M[115] | 0.14 | 0.11 | 1.00 | 577.97 |
| M[116] | 0.09 | 0.09 | 1.00 | 1253.04 |
| M[117] | -0.67 | 0.42 | 1.00 | 262.37 |
| M[118] | 0.20 | 0.12 | 1.00 | 459.20 |
| M[119] | 0.20 | 0.11 | 1.00 | 619.09 |
| M[120] | 0.16 | 0.09 | 1.00 | 3032.96 |
| M[121] | 0.10 | 0.09 | 1.00 | 24230.64 |
| M[122] | 0.20 | 0.12 | 1.00 | 455.49 |
| M[123] | 0.22 | 0.10 | 1.00 | 802.97 |
| M[124] | -0.09 | 0.10 | 1.00 | 891.92 |
| M[125] | 0.23 | 0.11 | 1.00 | 598.70 |
| M[126] | 0.14 | 0.09 | 1.00 | 1292.03 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| M[127] | 0.15 | 0.09 | 1.00 | 5083.21 | M[195] | 0.12 | 0.09 | 1.00 | 9578.41 |
| M[128] | -0.20 | 0.17 | 1.00 | 325.75 | M[196] | 0.31 | 0.15 | 1.00 | 383.47 |
| M[129] | 0.13 | 0.11 | 1.00 | 570.92 | M[197] | 0.29 | 0.13 | 1.00 | 438.43 |
| M[130] | 0.14 | 0.11 | 1.00 | 529.29 | M[198] | 0.21 | 0.10 | 1.00 | 847.25 |
| M[131] | 0.03 | 0.09 | 1.00 | 5937.40 | M[199] | 0.16 | 0.09 | 1.00 | 2892.88 |
| M[132] | 0.06 | 0.09 | 1.00 | 1776.27 | M[200] | -0.03 | 0.09 | 1.00 | 2745.61 |
| M[133] | 0.14 | 0.11 | 1.00 | 547.30 | M[201] | -0.00 | 0.08 | 1.00 | 28398.18 |
| M[134] | 0.26 | 0.12 | 1.00 | 527.88 | M[202] | -0.59 | 0.35 | 1.00 | 264.60 |
| M[135] | -0.05 | 0.09 | 1.00 | 1167.45 | M[203] | -0.13 | 0.11 | 1.00 | 539.11 |
| M[136] | 0.04 | 0.08 | 1.00 | 13293.77 | M[204] | -0.09 | 0.10 | 1.00 | 872.12 |
| M[137] | 0.17 | 0.11 | 1.00 | 628.18 | M[205] | 0.17 | 0.09 | 1.00 | 2153.71 |
| M[138] | 0.04 | 0.09 | 1.00 | 7557.98 | M[206] | 0.11 | 0.09 | 1.00 | 23556.31 |
| M[139] | 0.18 | 0.10 | 1.00 | 1344.05 | M[207] | 0.12 | 0.11 | 1.00 | 590.57 |
| M[140] | -0.13 | 0.16 | 1.00 | 350.58 | M[208] | 0.08 | 0.09 | 1.00 | 1168.01 |
| M[141] | 0.26 | 0.12 | 1.00 | 521.52 | M[209] | -1.34 | 0.78 | 1.00 | 254.73 |
| M[142] | -0.00 | 0.09 | 1.00 | 11809.86 | M[210] | 0.09 | 0.09 | 1.00 | 7761.20 |
| M[143] | 0.19 | 0.09 | 1.00 | 1385.33 | M[211] | 0.13 | 0.10 | 1.00 | 928.82 |
| M[144] | 0.19 | 0.12 | 1.00 | 483.86 | M[212] | 0.04 | 0.09 | 1.00 | 1867.75 |
| M[145] | 0.28 | 0.12 | 1.00 | 477.88 | M[213] | 0.21 | 0.14 | 1.00 | 403.08 |
| M[146] | 0.19 | 0.10 | 1.00 | 778.12 | M[214] | 0.12 | 0.09 | 1.00 | 2354.28 |
| M[147] | 0.29 | 0.13 | 1.00 | 424.69 | M[215] | -0.84 | 0.55 | 1.00 | 259.73 |
| M[148] | 0.06 | 0.09 | 1.00 | 12673.95 | M[216] | 0.29 | 0.14 | 1.00 | 401.66 |
| M[149] | 0.10 | 0.10 | 1.00 | 931.22 | M[217] | 0.26 | 0.12 | 1.00 | 495.29 |
| M[150] | 0.32 | 0.15 | 1.00 | 374.50 | M[218] | 0.09 | 0.09 | 1.00 | 1932.48 |
| M[151] | 0.15 | 0.12 | 1.00 | 506.33 | M[219] | 0.18 | 0.09 | 1.00 | 1810.35 |
| M[152] | 0.09 | 0.09 | 1.00 | 5534.65 | M[220] | 0.02 | 0.08 | 1.00 | 17423.87 |
| M[153] | 0.12 | 0.10 | 1.00 | 765.12 | M[221] | 0.08 | 0.09 | 1.00 | 2606.16 |
| M[154] | -0.06 | 0.09 | 1.00 | 2119.53 | M[222] | 0.22 | 0.12 | 1.00 | 539.30 |
| M[155] | 0.06 | 0.09 | 1.00 | 6942.53 | M[223] | 0.16 | 0.09 | 1.00 | 3894.70 |
| M[156] | 0.02 | 0.10 | 1.00 | 915.31 | M[224] | 0.14 | 0.09 | 1.00 | 1905.57 |
| M[157] | 0.22 | 0.12 | 1.00 | 519.03 | M[225] | 0.11 | 0.10 | 1.00 | 665.79 |
| M[158] | 0.20 | 0.10 | 1.00 | 1101.52 | M[226] | 0.07 | 0.09 | 1.00 | 1515.35 |
| M[159] | -0.09 | 0.14 | 1.00 | 391.16 | M[227] | 0.04 | 0.09 | 1.00 | 2618.51 |
| M[160] | 0.04 | 0.09 | 1.00 | 5382.45 | M[228] | -0.13 | 0.13 | 1.00 | 428.75 |
| M[161] | -0.25 | 0.22 | 1.00 | 295.00 | M[229] | 0.13 | 0.10 | 1.00 | 803.59 |
| M[162] | 0.11 | 0.10 | 1.00 | 774.67 | M[230] | 0.18 | 0.13 | 1.00 | 421.37 |
| M[163] | -0.01 | 0.08 | 1.00 | 26317.35 | M[231] | -0.08 | 0.10 | 1.00 | 937.18 |
| M[164] | -0.22 | 0.15 | 1.00 | 364.21 | M[232] | -0.34 | 0.28 | 1.00 | 281.75 |
| M[165] | 0.09 | 0.10 | 1.00 | 911.99 | M[233] | 0.19 | 0.10 | 1.00 | 697.48 |
| M[166] | -0.14 | 0.12 | 1.00 | 507.95 | M[234] | 0.05 | 0.09 | 1.00 | 1700.81 |
| M[167] | 0.24 | 0.11 | 1.00 | 559.86 | M[235] | -0.13 | 0.11 | 1.00 | 576.87 |
| M[168] | 0.03 | 0.10 | 1.00 | 1037.69 | M[236] | 0.09 | 0.09 | 1.00 | 1689.50 |
| M[169] | 0.24 | 0.11 | 1.00 | 629.32 | M[237] | 0.18 | 0.09 | 1.00 | 1408.76 |
| M[170] | -0.32 | 0.26 | 1.00 | 280.44 | M[238] | 0.19 | 0.13 | 1.00 | 436.70 |
| M[171] | 0.11 | 0.10 | 1.00 | 830.07 | M[239] | 0.17 | 0.13 | 1.00 | 437.54 |
| M[172] | 0.01 | 0.11 | 1.00 | 664.73 | M[240] | -0.21 | 0.17 | 1.00 | 334.39 |
| M[173] | 0.00 | 0.09 | 1.00 | 2467.21 | M[241] | 0.19 | 0.14 | 1.00 | 382.35 |
| M[174] | 0.25 | 0.12 | 1.00 | 496.28 | M[242] | 0.19 | 0.14 | 1.00 | 380.05 |
| M[175] | 0.22 | 0.12 | 1.00 | 470.91 | M[243] | 0.13 | 0.11 | 1.00 | 565.48 |
| M[176] | 0.03 | 0.10 | 1.00 | 947.64 | M[244] | 0.09 | 0.09 | 1.00 | 24551.22 |
| M[177] | 0.18 | 0.10 | 1.00 | 869.07 | M[245] | 0.08 | 0.10 | 1.00 | 1094.22 |
| M[178] | 0.02 | 0.08 | 1.00 | 27844.71 | M[246] | -0.12 | 0.11 | 1.00 | 591.09 |
| M[179] | 0.13 | 0.09 | 1.00 | 17395.46 | M[247] | 0.05 | 0.08 | 1.00 | 27832.41 |
| M[180] | 0.09 | 0.10 | 1.00 | 953.35 | M[248] | 0.19 | 0.09 | 1.00 | 1373.09 |
| M[181] | 0.26 | 0.13 | 1.00 | 450.40 | M[249] | 0.16 | 0.11 | 1.00 | 598.22 |
| M[182] | -0.04 | 0.09 | 1.00 | 1703.32 | M[250] | 0.20 | 0.15 | 1.00 | 364.62 |
| M[183] | 0.03 | 0.09 | 1.00 | 3480.08 | M[251] | 0.10 | 0.10 | 1.00 | 734.56 |
| M[184] | 0.18 | 0.13 | 1.00 | 423.56 | M[252] | 0.17 | 0.13 | 1.00 | 452.61 |
| M[185] | 0.19 | 0.09 | 1.00 | 1384.80 | M[253] | 0.23 | 0.11 | 1.00 | 559.93 |
| M[186] | -0.29 | 0.23 | 1.00 | 292.85 | M[254] | 0.04 | 0.09 | 1.00 | 2556.67 |
| M[187] | 0.09 | 0.10 | 1.00 | 1096.64 | M[255] | 0.09 | 0.10 | 1.00 | 963.43 |
| M[188] | 0.31 | 0.14 | 1.00 | 392.99 | M[256] | 0.16 | 0.12 | 1.00 | 512.42 |
| M[189] | -0.25 | 0.17 | 1.00 | 334.49 | M[257] | 0.26 | 0.12 | 1.00 | 512.34 |
| M[190] | 0.19 | 0.14 | 1.00 | 383.34 | M[258] | 0.19 | 0.12 | 1.00 | 486.30 |
| M[191] | 0.18 | 0.13 | 1.00 | 449.83 | M[259] | 0.04 | 0.10 | 1.00 | 1083.06 |
| M[192] | 0.23 | 0.11 | 1.00 | 602.10 | M[260] | 0.24 | 0.12 | 1.00 | 532.05 |
| M[193] | 0.13 | 0.09 | 1.00 | 1223.55 | M[261] | 0.24 | 0.11 | 1.00 | 625.82 |
| M[194] | 0.10 | 0.10 | 1.00 | 996.79 | M[262] | 0.25 | 0.12 | 1.00 | 475.79 |

| | | | | |
|---|---|---|---|---|
| M[263] | 0.12 | 0.10 | 1.00 | 844.59 |
| M[264] | 0.18 | 0.09 | 1.00 | 1697.78 |
| M[265] | 0.18 | 0.13 | 1.00 | 420.38 |
| M[266] | 0.31 | 0.14 | 1.00 | 398.51 |
| M[267] | -0.43 | 0.28 | 1.00 | 272.57 |
| M[268] | 0.16 | 0.10 | 1.00 | 706.26 |
| M[269] | 0.17 | 0.10 | 1.00 | 736.25 |
| M[270] | 0.13 | 0.09 | 1.00 | 24732.78 |
| M[271] | 0.27 | 0.13 | 1.00 | 436.61 |
| M[272] | 0.27 | 0.12 | 1.00 | 480.88 |
| M[273] | -0.56 | 0.40 | 1.00 | 263.69 |
| M[274] | 0.05 | 0.09 | 1.00 | 1729.54 |
| M[275] | 0.10 | 0.10 | 1.00 | 936.16 |
| M[276] | 0.07 | 0.09 | 1.00 | 4261.66 |
| M[277] | 0.09 | 0.09 | 1.00 | 1140.48 |
| M[278] | 0.17 | 0.10 | 1.00 | 744.45 |
| M[279] | 0.10 | 0.09 | 1.00 | 19820.72 |
| M[280] | 0.28 | 0.14 | 1.00 | 387.66 |
| M[281] | 0.08 | 0.09 | 1.00 | 1321.87 |
| M[282] | 0.29 | 0.13 | 1.00 | 426.23 |
| M[283] | -0.03 | 0.09 | 1.00 | 9311.02 |
| M[284] | 0.04 | 0.09 | 1.00 | 1482.60 |
| M[285] | -0.10 | 0.11 | 1.00 | 514.26 |
| M[286] | 0.13 | 0.09 | 1.00 | 16594.92 |
| M[287] | 0.12 | 0.10 | 1.00 | 1065.93 |
| M[288] | 0.16 | 0.12 | 1.00 | 481.08 |
| M[289] | 0.24 | 0.13 | 1.00 | 449.17 |
| M[290] | 0.24 | 0.11 | 1.00 | 569.46 |
| M[291] | 0.24 | 0.12 | 1.00 | 536.61 |
| M[292] | 0.11 | 0.11 | 1.00 | 619.96 |
| M[293] | -0.03 | 0.09 | 1.00 | 5366.79 |
| M[294] | 0.25 | 0.14 | 1.00 | 374.24 |
| M[295] | -0.29 | 0.25 | 1.00 | 288.29 |
| M[296] | 0.12 | 0.09 | 1.00 | 1293.78 |
| M[297] | 0.02 | 0.08 | 1.00 | 17391.52 |
| M[298] | -0.62 | 0.36 | 1.00 | 267.10 |
| M[299] | 0.23 | 0.11 | 1.00 | 690.51 |
| M[300] | -0.73 | 0.48 | 1.00 | 260.24 |
| M[301] | 0.20 | 0.14 | 1.00 | 373.75 |
| M[302] | 0.12 | 0.10 | 1.00 | 855.45 |
| M[303] | 0.22 | 0.14 | 1.00 | 375.62 |
| M[304] | -0.21 | 0.21 | 1.00 | 308.78 |
| M[305] | 0.18 | 0.10 | 1.00 | 721.71 |
| M[306] | 0.23 | 0.11 | 1.00 | 725.93 |
| M[307] | 0.19 | 0.11 | 1.00 | 561.70 |
| M[308] | 0.25 | 0.12 | 1.00 | 532.79 |
| M[309] | -0.89 | 0.57 | 1.00 | 256.95 |
| M[310] | -0.04 | 0.09 | 1.00 | 4369.59 |
| M[311] | 0.05 | 0.09 | 1.00 | 1890.57 |
| M[312] | 0.10 | 0.10 | 1.00 | 828.15 |
| M[313] | 0.09 | 0.10 | 1.00 | 779.18 |
| M[314] | -0.12 | 0.11 | 1.00 | 606.26 |
| M[315] | 0.11 | 0.11 | 1.00 | 673.32 |
| M[316] | -0.07 | 0.09 | 1.00 | 1474.47 |
| M[317] | 0.14 | 0.12 | 1.00 | 497.87 |
| M[318] | 0.07 | 0.09 | 1.00 | 5626.91 |
| M[319] | 0.11 | 0.09 | 1.00 | 23654.13 |
| M[320] | 0.28 | 0.13 | 1.00 | 439.41 |

**Table C.3:** Summary of samples of $M^i$ when running the model on Svalna data.