



HAL
open science

Bayesian analysis for mediation and moderation using g-priors

Jean-Michel Galharret, Anne Philippe

► **To cite this version:**

Jean-Michel Galharret, Anne Philippe. Bayesian analysis for mediation and moderation using g-priors. 2021. hal-03104707v1

HAL Id: hal-03104707

<https://hal.science/hal-03104707v1>

Preprint submitted on 9 Jan 2021 (v1), last revised 13 Nov 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bayesian analysis for mediation and moderation using g -priors.

J.M. Galharret and A. Philippe

January 9, 2021, Laboratoire de Mathématiques Jean Leray (LMJL- UMR 6629), F-44000
Nantes, France;

Abstract

We propose a Bayesian analysis of moderate mediation model. We show that g -priors can be extended to mediation model. For this choice of prior, we provide an explicit form of the marginal distribution. This result is applied to testing moderation using Bayes factor and indirect effects in mediation models. We apply our results to the study of the association between empowering leadership and organizational commitment.

Keywords: Bayes factor, direct and indirect effects, g -priors, interaction, linear model, model selection, structural equation, testing procedure.

1 Introduction

In human sciences, mediation refers to a phenomenon in which the effect of an exposure variable X on an outcome Y can be decomposed into a direct effect and an indirect effect via a third variable M . In this paper, we focus on the association between empowering leadership and organizational commitment mediated by well-being at work. This mediation analysis is performed on two populations corresponding to different companies. We are interested in the invariance of the associations between the previous variables. From a statistical point of view, such problems can be dealt with using mediation analysis and moderate mediation analysis.

Mediation analysis is one of the classic applications of the Structural Equation Modeling (SEM) (see for instance Judd and Kenny, 1981; Bollen, 2014; Muthén et al., 2016) and regression models (see for instance MacKinnon, 2008; Hayes, 2018). These effects can be interpreted in terms of correlation and association. The total effect of X on Y is the association between X and Y . The direct effect of X on Y is the partial correlation between X and Y controlling the M effect. The indirect effect of X on Y is the remaining of the total effect that will pass through M . More precisely, in the standard situation, where X, M, Y are continuous, the most widely model used to evaluate the effect of X on Y is the linear regression: $Y = i_0 + CX + \varepsilon$. In this case, C measures the total effect of X on Y . In the presence of the mediator M , the linear model is of the form:

$$\begin{cases} Y = i_2 + bM + cX + \varepsilon_2 \\ M = i_1 + aX + \varepsilon_1, \end{cases} \quad (1)$$

where ε_i have Gaussian distributions with 0 mean and σ_i^2 variance for $i = 1, 2$ (see for instance MacKinnon, 2008; Jose, 2013; Hayes, 2018). Combining both equations, the total effect is equal to $c + ab$, c is the direct effect, and the indirect effect is the product ab . In one of the most cited articles on mediation, Baron and Kenny (1986) describe the test procedure introduced by Sobel (1982) to test the indirect effect using the product ab . This is an asymptotic procedure based on the Delta method and on the central limit theorem. More generally, all these effects can be redefined within the context of causality (see for instance Pearl, 2001; Robins, 2003; VanderWeele, 2015). In this case, they measure the causal effects of X on Y instead of the association between of X and Y . The correspondance between the causal and SEM mediation analysis is discussed in MacKinnon et al. (2020). Consider a generalisation of the mediation model (1) defined as follows:

$$[\mathcal{M}_0] : \begin{cases} Y = i_2 + c^T X + bM + e_0W + \varepsilon_Y \\ M = i_1 + a^T X + d_0W + \varepsilon_M \end{cases} \quad (2)$$

where $Y \in \mathbb{R}^n$ is the continuous outcome, $M \in \mathbb{R}^n$ is the continuous mediator, $X \in \mathbb{R}^{n,p}$ is the matrix of exposition vectors and W is a covariate. The exposition variables X and the covariate W are supposed to be deterministic. We assume that the error terms are independent Gaussian random variables: $\varepsilon_M \sim \mathcal{N}_n(0, \sigma_M^2 I_n)$, $\varepsilon_Y \sim \mathcal{N}_n(0, \sigma_Y^2 I_n)$. The unknown parameters satisfy $a \in \mathbb{R}^p$, $c \in \mathbb{R}^p$; $b, d_0, e_0 \in \mathbb{R}$. In this multivariate exposition model, the effects are defined for each component. The direct effect of component X_k on Y is c_k and its indirect effect is $a_k \times b$.

In our application, X corresponds to the four dimensions of the empowering leadership (enhancing the meaningfulness of work, fostering participation in decision making, expressing confidence in high performance, providing autonomy) M corresponds to well-being at work and W is the activity sector (industry or fire station, see details to follow). To evaluate the effect of the populations, we consider the binary variable W as a moderator.

Moderating the mediation model consists of adding interaction effects in equations of $[\mathcal{M}_0]$. Hereafter, the matrix $X : W$ designates the interactions between the variables X_1, \dots, X_p and W . Several moderations can be investigated: one path (or all paths) between X and Y is (are) moderated by W (i.e. a, c in Figure 1), the path b between M and Y can also be moderated by W . Preacher et al. (2007) discuss all these different cases and they compare normal-theory and bootstrapping standard errors for assessing conditional indirect effects.

The model including all paths is written as follows

$$[\mathcal{M}_1] : \begin{cases} Y = i_2 + c^T X + bM + e_0W + e_1^T X : W + f_1 M : W + \varepsilon_Y, \\ M = i_1 + a^T X + d_0W + d_1^T X : W + \varepsilon_M \end{cases} \quad (3)$$

where $d_1 \in \mathbb{R}^p$, $e_1 \in \mathbb{R}^p$, $f_1 \in \mathbb{R}$ correspond respectively to the moderation of the effect of X on M , the moderation of the effect of X on Y and the moderation of the effect of M on Y . Figure 1 gives a graphical summary of both models $[\mathcal{M}_0]$ and $[\mathcal{M}_1]$.

We want to test the existence of these effects, this amounts to testing the null hypothesis

$$\mathcal{H}_0 : d_1 = e_1 = 0_p, f_1 = 0.$$

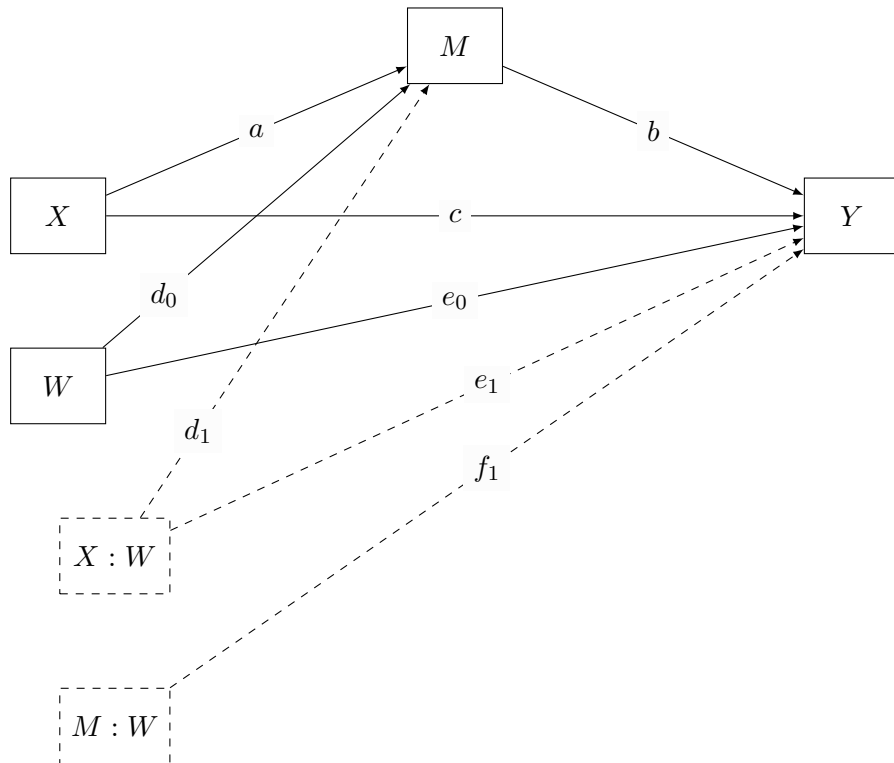


Figure 1: Mediation model (solid lines) and moderate mediation model (with dashed lines added).

The statistical problem in the SEM framework is standard (see for instance Bollen, 2014, P 292-300). The nested models $[\mathcal{M}_0]$ and $[\mathcal{M}_1]$ are compared using the likelihood ratio test (LR-test)

$$\lambda = 2 \log \left(\frac{\mathcal{L}(\hat{\theta}_1)}{\mathcal{L}(\hat{\theta}_0)} \right),$$

where $\hat{\theta}_1$ and $\hat{\theta}_0$ are respectively the maximum likelihood estimators of the parameters in models $[\mathcal{M}_0]$ and $[\mathcal{M}_1]$. Under the null hypothesis λ has a chi-square distribution with $2p + 1$ degrees of freedom.

In this paper, we address this problem from a Bayesian framework using the Bayes factor. The Bayesian parameter estimation for mediation models is studied (for instance in Yuan and MacKinnon, 2009; Miočević et al., 2018). They consider an independent Gaussian priors on mediation coefficients or non informative priors in the spirit of prior distributions used in linear regression models. Nuijten et al. (2015) and Biesanz et al. (2010) propose a Bayesian testing procedure for the indirect effect ab . More precisely, Nuijten et al. (2015) test the indirect effect by combining Bayes factors calculated independently on the two regression equations that define (1). For each linear regression model, they use the g -priors distribution introduced by Zellner (1984). To get one step estimation, we complete this approach by extending the g -prior construction to the mediation model defined in (2). One advantage is to obtain a

joint posterior distribution of all parameters from which one can deduce posterior distribution of direct and indirect effect. The Bayesian method is also studied for moderation. For instance, Wang (2015) extends the Yuan and MacKinnon (2009) model for testing whether mediation is moderated.

The paper is organized as follows. In Section 2, we construct an extension of g -prior adapted to models (2) and (3). In Section 3, we show that an explicit form of Bayes factor can be obtained for selecting the best model between $[\mathcal{M}_0]$ and $[\mathcal{M}_1]$ defined in (2) and (3). This Bayes factor gives a tool for testing the moderation of the effects (i.e. $\mathcal{H}_0 : d_1 = e_1 = f_1 = 0$). Finally, Section 4 contains our application on the effect of empowering leadership on organizational commitment.

2 Zellner's g -priors choice

The g -prior distributions were introduced by Zellner (see for instance Zellner, 1971, 1984), for the coefficients of multiple linear regression models

$$Y = \mathbb{X}\beta + \varepsilon,$$

where $\mathbb{X} \in \mathbb{R}^{n \times p}$ is the design matrix, $Y \in \mathbb{R}^n$ is a Gaussian vector. We denote σ^2 the variance of the Gaussian noise ε . The Zellner's g -priors distribution on the parameters (β, σ^2) is¹

$$\beta | \sigma^2 \sim \mathcal{N}_p(\tilde{\beta}, g\sigma^2(\mathbb{X}^T \mathbb{X})^{-1}) \text{ and } \pi(\sigma^2) \propto \sigma^{-2}. \quad (4)$$

To designate the g -priors we use the notation

$$(\beta, \sigma^2) \sim \mathcal{Z}_g(g, \tilde{\beta}, \mathbb{X}).$$

This prior depends on hyper-parameters $\tilde{\beta} \in \mathbb{R}, g \in \mathbb{R}$ to be fixed according to the prior information. The standard choice $g = N$ gives the prior the same weight as one observation (see Marin and Robert, 2014).

In the context of mediation models, Nuijten et al. (2015) use the g -priors by addressing the model (2) as two independent regression models. From this point of view, it is not possible to obtain the posterior distribution of the indirect effect ab since the parameters a, b are estimated in both independent models.

We show that it is possible to define an extension of g -prior to the parameters $\theta := (\theta_M, \sigma_M^2, \theta_Y, \sigma_Y^2)$ of the mediation model. Using the notations in (2) we have $\theta_M = (i_1, a, d_0)$ and $\theta_Y = (i_2, c, b, d_0)$. In the global model (2), the joint distribution of M, Y, θ is:

$$p(m, y, \theta | X, W) = f_Y(y | m, \theta, X, W) f_M(m | \theta, X, W) \pi(\theta),$$

where f_Y is the Gaussian distribution with mean $i_2 + c^T X + bm + e_0 W$ and variance σ_Y^2 and where f_M is the Gaussian distribution with mean $i_1 + a^T X + d_0 W$ and variance σ_M^2 .

A simple way to construct the prior distribution of θ is to assume that the parameters of both equations in (2) are independent, i.e.

$$\pi(\theta) = \pi_M(\theta_M, \sigma_M^2) \pi_Y(\theta_Y, \sigma_Y^2).$$

¹the notation $Z_1 | Z_2 \sim \mathbb{P}$ means that the conditional distribution of Z_1 given Z_2 is \mathbb{P} .

From this decomposition, it is not possible to take the g -priors for π_M, π_Y as defined in (4), because the observation M will appear in the design matrix of π_Y .

We propose the following approach to adapt g -priors to model (2). The main idea is to decompose the joint distribution of M, Y, θ as follows

$$p(m, y, \theta_Y, \sigma_Y^2, \theta_M, \sigma_M^2 | X, W) = f_Y(y | M = m, \theta_Y, \sigma_Y^2, \theta_M, \sigma_M^2, X, W) \pi(\theta_Y, \sigma_Y^2 | M = m, \theta_M, \sigma_M^2, X, W) \times f_M(m | \theta_M, \sigma_M^2, X, W) \pi(\theta_M, \sigma_M^2 | X, W).$$

According to the dependencies defined by the directed acyclic graph represented in Figure 2, we can simplify p ,

$$p(m, y, \theta_Y, \sigma_Y^2, \theta_M, \sigma_M^2 | X, W) = f_Y(y | M = m, \theta_Y, \sigma_Y^2, X, W) \pi_Y(\theta_Y, \sigma_Y^2 | M = m, X, W) \times f_M(m | \theta_M, \sigma_M^2, X, W) \pi_M(\theta_M, \sigma_M^2 | X, W), \quad (5)$$

where f_Y is the Gaussian distribution with mean $i_2 + c^T X + bm + e_0 W$ and variance σ_Y^2 and where f_M is the Gaussian distribution with mean $i_1 + a^T X + d_0 W$ and variance σ_M^2 . The g -priors defined in (4) can be chosen for π_M and π_Y with a design matrix $[\mathbf{1}, X, W]$ and $[\mathbf{1}, X, W, M]$ respectively. More precisely, we have

$$\pi_M(\theta_M, \sigma_M^2 | X, W) \propto \frac{1}{\sigma_M^{p+4}} \exp\left(-\frac{1}{2g_M \sigma_M^2} (\theta_M - \tilde{\theta}_M)^T [\mathbf{1}, X, W]^T [\mathbf{1}, X, W] (\theta_M - \tilde{\theta}_M)\right) \quad (6)$$

$$\pi_Y(\theta_Y, \sigma_Y^2 | X, W, M) \propto \frac{1}{\sigma_Y^{p+5}} \exp\left(-\frac{1}{2g_Y \sigma_Y^2} (\theta_Y - \tilde{\theta}_Y)^T [\mathbf{1}, X, W, M]^T [\mathbf{1}, X, W, M] (\theta_Y - \tilde{\theta}_Y)\right), \quad (7)$$

where $g_M, \tilde{\theta}_M, g_Y, \tilde{\theta}_Y$ are the hyperparameters of the g -priors.

Remark 1. Like Zellner (1984), we include the intercept of the model in the parameter β . Indeed, in many applications of mediation, the problem of location scale is not relevant since the outcome is a score without units.

Equations (5), (6), (7) fully defined the Bayesian mediation model with the g -priors. The posterior distribution satisfies

$$\pi(\theta_M, \sigma_M^2, \theta_Y, \sigma_Y^2 | X, W, M, Y) \propto p(M, Y, \theta_Y, \sigma_Y^2, \theta_M, \sigma_M^2 | X, W). \quad (8)$$

These prior distributions cover a wide range of contexts from non-informative to informative. Indeed we can bring information on θ from the hyperparameters $\tilde{\theta}_M, \tilde{\theta}_Y$, and g_M, g_Y quantify the weight of this information. In the absence of information, $\tilde{\theta}_M = \tilde{\theta}_Y = 0$ and $g_M = g_Y = N$ are the standard choices.

Remark 2. This model can be easily adapted to moderate mediation model defined in (3). Indeed, it is enough to add $X : W$ and $M : W$ in the design matrix.

Inference on the effects: In applications, one of the main issues is to test the existence of the mediated effect and if this effect is moderated by W . The posterior distribution $\pi_I(\bullet|X, W, M, Y)$ of the indirect effect $I = ab$ can be deduced from (8). As suggested by Biesanz et al. (2010), a decision rule for testing the absence of this effect (i.e. $ab = 0$) is obtained from the HPD region of level $1 - \alpha$. More precisely we accept the absence of effect if

$$0 \in H_\alpha := \{t : \pi_I(t|X, W, M, Y) > k_\alpha\},$$

where k_α is chosen such that

$$\int_{H_\alpha} \pi_I(t|X, W, M, Y) dt = 1 - \alpha.$$

This procedure can be applied for all the other effects in particular for the moderation model. Indeed, we can easily calculate the posterior distribution of both moderate effects: the moderate direct effect $c + e_1$ and the moderate indirect effect $(a + d_1)(b + f_1)$.

An alternative is to treat moderation as a problem of model selection. It is this approach that we discuss in the next section.

3 Bayes factor for testing moderate mediation

Our aim is to test if the variable W moderates the associations between X, M, Y . This problem can be expressed as a model selection problem between the two models $[\mathcal{M}_0]$ and model $[\mathcal{M}_1]$. For both models $[\mathcal{M}_0]$ and $[\mathcal{M}_1]$, we consider as prior distribution the g -priors defined in Section 2. The Bayes factor provides a classical solution for comparing two models. Let us recall the definition of Bayes factor for comparing two Bayesian models. Let \mathfrak{M}^i be the marginal distribution of the model $[\mathcal{M}_i]$,

$$\mathfrak{M}^i(y) = \int_{\Theta_i} f_i(y|\theta)\pi_i(\theta) d\theta,$$

where f_i, π_i are respectively the likelihood and the prior distribution of model $[\mathcal{M}_i]$. The Bayes factor is the ratio of the marginal densities for the two models

$$BF_{10} = \frac{\mathfrak{M}^1(y)}{\mathfrak{M}^0(y)}. \quad (9)$$

The evidence of $[\mathcal{M}_1]$ over $[\mathcal{M}_0]$ is obtained for $BF_{10} > 1$.

For the moderation model, we show that an explicit form of the Bayes factor is available. Our competing Bayesian models are as follows

- Model \mathcal{M}_1 :

$$\begin{aligned} Y &= i_2 + c^T X + bM + e_0 W + e_1^T X : W + f_1 M : W + \varepsilon_Y, \\ (\theta_Y, \sigma_Y^2)|X, W, M &\sim \mathcal{Z}_g(g_2, \tilde{\beta}_2, \mathbb{X}_{Y_1}) \text{ with } \mathbb{X}_{Y_1} = [\mathbf{1}, W, X, M, X : W, M : W] \\ M &= i_1 + a^T X + d_0 W + d_1^T X : W + \varepsilon_M \\ (\theta_M, \sigma_M^2)|X, W &\sim \mathcal{Z}_g(g_1, \tilde{\beta}_1, [\mathbf{1}, W, X, X : W]) \text{ with } \mathbb{X}_{M_1} = [\mathbf{1}, W, X, X : W] \end{aligned}$$

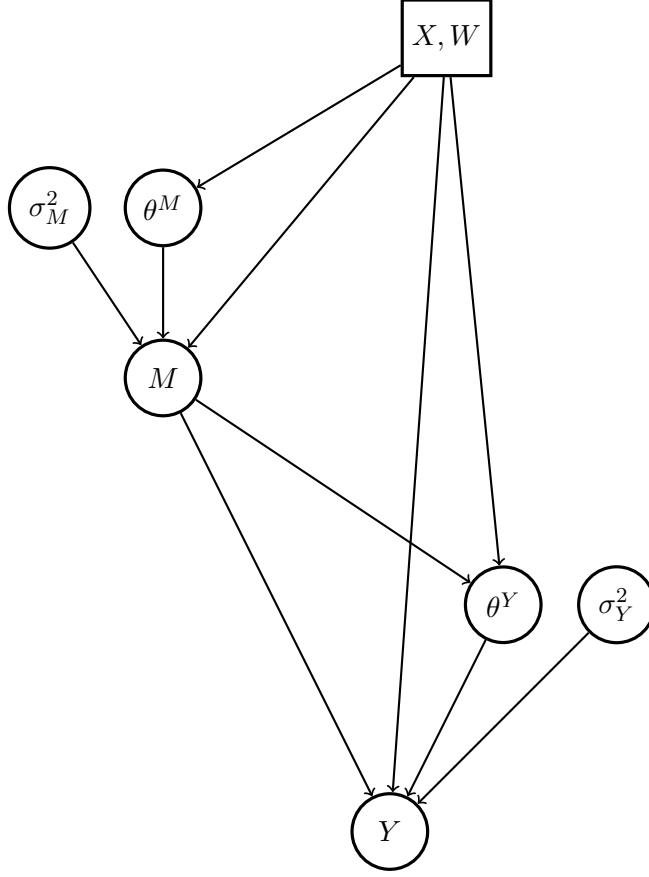


Figure 2: Directed acyclic graph for the mediation model $[\mathcal{M}_0]$

- Model \mathcal{M}_0 :

$$\begin{aligned}
 Y &= i_2 + c^T X + bM + e_0 W + \varepsilon_Y, \\
 (\theta_Y, \sigma_Y^2) | X, W, M &\sim \mathcal{Z}_g(g_2, \check{\beta}_2, [\mathbf{1}, W, X, M]) \text{ with } \check{\beta}_2 = (\tilde{\beta}_2^1, \dots, \tilde{\beta}_2^{p+3}) \text{ and } \mathbb{X}_{Y_0} = [\mathbf{1}, W, X, M] \\
 M &= i_1 + a^T X + d_0 W + \varepsilon_M \\
 (\theta_M, \sigma_M^2) | X, W &\sim \mathcal{Z}_g(g_1, \check{\beta}_1, [\mathbf{1}, W, X]) \text{ with } \check{\beta}_1 = (\tilde{\beta}_1^1, \dots, \tilde{\beta}_1^{p+2}) \text{ and } \mathbb{X}_{M_0} = [\mathbf{1}, W, X]
 \end{aligned}$$

The form of the hyperparameters ensures that the information on common parameters are the same.

Proposition 1. *The Bayes factor to compare the models \mathcal{M}_1 versus \mathcal{M}_0 defined above is given by*

$$BF_{10} = \frac{(g_1 + 1)^{p/2}}{(g_2 + 1)^{(p+1)/2}} \left(\frac{Y^T Y - \frac{g_2}{g_2+1} Y^T \mathbb{X}_{Y_0} (\mathbb{X}_{Y_0}^T \mathbb{X}_{Y_0})^{-1} \mathbb{X}_{Y_0}^T Y - \frac{\|\mathbb{X}_{Y_0} \tilde{\beta}_2\|^2}{g_2+1}}{Y^T Y - \frac{g_2}{g_2+1} Y^T \mathbb{X}_{Y_1} (\mathbb{X}_{Y_1}^T \mathbb{X}_{Y_1})^{-1} \mathbb{X}_{Y_1}^T Y - \frac{\|\mathbb{X}_{Y_1} \tilde{\beta}_2\|^2}{g_2+1}} \right)^{n/2} \times \left(\frac{M^T M - \frac{g_1}{g_1+1} M^T \mathbb{X}_{M_0} (\mathbb{X}_{M_0}^T \mathbb{X}_{M_0})^{-1} \mathbb{X}_{M_0}^T M - \frac{\|\mathbb{X}_{M_0} \tilde{\beta}_1\|^2}{g_1+1}}{M^T M - \frac{g_1}{g_1+1} M^T \mathbb{X}_{M_1} (\mathbb{X}_{M_1}^T \mathbb{X}_{M_1})^{-1} \mathbb{X}_{M_1}^T M - \frac{\|\mathbb{X}_{M_1} \tilde{\beta}_1\|^2}{g_1+1}} \right)^{n/2}.$$

Proof. By definition, the Bayes factor is the ratio of the marginal densities

$$BF_{10} = \frac{\mathfrak{M}^1(m, y|X, W)}{\mathfrak{M}^0(m, y|X, W)}. \quad (10)$$

For model \mathcal{M}_0 , the marginal density is

$$\mathfrak{M}^0(m, y|X, W) = \int_{\Theta_0} p(m, y, \theta_Y, \sigma_Y^2, \theta_M, \sigma_M^2 | X, W) d\theta_Y d\sigma_Y^2 d\theta_M d\sigma_M^2,$$

where p is defined in (5). It can be rewritten

$$\mathfrak{M}^0(m, y|X, W) = \int_{\Theta_0} f_Y(y|X, W, M = m, \theta_Y, \sigma_Y^2) \pi_Y(\theta_Y, \sigma_Y^2 | X, W, M = m) \times f_M(M|X, W, \theta_M, \sigma_M) \pi_M(\theta_M, \sigma_M^2 | X, W) d\theta_Y d\sigma_Y^2 d\theta_M d\sigma_M^2,$$

where f_Y, f_M, π_Y, π_M in (5), (6), (7). By the Fubini theorem, we obtain

$$\mathfrak{M}^0(m, y|X, W) = \int_{\Theta_{0,Y}} f_Y(y|X, W, M = m, \theta_Y, \sigma_Y^2) \pi_Y(\theta_Y, \sigma_Y^2 | X, W, M = m) d\theta_Y d\sigma_Y^2 \times \int_{\Theta_{0,M}} f_M(m|X, W, \theta_M, \sigma_M^2) \pi_M(\theta_M, \sigma_M^2 | X, W) d\theta_M d\sigma_M^2.$$

From the last equation, we can identify that \mathfrak{M}^0 is the product of the marginal densities of both linear regression models which define \mathcal{M}_0 ,

$$\mathfrak{M}^0(m, y|X, W) = \mathfrak{M}_M^0(m|X, W) \times \mathfrak{M}_Y^0(y|X, W, M), \quad (11)$$

where $\mathfrak{M}_M^0(m|X, W)$ and $\mathfrak{M}_Y^0(y|X, W, M)$ correspond respectively to the marginal densities of M and Y in each linear equations in model $[\mathcal{M}_0]$. The same result is valid for the model \mathcal{M}_1 after adaptation of the design matrix. Therefore, the Bayes factor is expressed in the form

$$BF_{10} = BF_{10}^{M|X} \times BF_{10}^{Y|M,X}, \quad (12)$$

where

- $BF_{10}^{M|X}$ is the Bayes factor of competing models $M = i_1 + a^T X + d_0 W + d_1^T X : W + \varepsilon_M$ versus $M = i_1 + a^T X + d_0 W + \varepsilon_M$,
- $BF_{10}^{Y|M,X}$ is the Bayes factor of competing models $Y = i_2 + c^T X + bM + e_0 W + e_1^T X : W + f_1 M : W + \varepsilon_Y$ versus $Y = i_2 + c^T X + bM + e_0 W + \varepsilon_Y$.

For the linear regression model with g -priors the form of the marginal densities are well-known (see for instance Zellner, 1984). Indeed, if we consider the Gaussian linear model

$$y|\beta, \sigma^2 \sim \mathcal{N}_n(\mathbb{X}\beta, \sigma^2 I_N),$$

with $(\beta, \sigma^2) \sim \mathcal{Z}_g(g, \tilde{\beta}, \mathbb{X})$, then the marginal density of y is

$$p(y) = (g+1)^{-(p+1)/2} \pi^{-n/2} \Gamma(n/2) \left(y^T y - \frac{g}{g+1} y^T \mathbb{X} (\mathbb{X}^T \mathbb{X})^{-1} \mathbb{X}^T y - \frac{\|\mathbb{X} \tilde{\beta}\|^2}{g+1} \right)^{-n/2}. \quad (13)$$

Using (12), we deduce the expression stated and this concludes the proof. \square

Remark 3. *The Bayes factor is not defined when the priors are improper. However there is a major exception to this ban on improper priors that we apply to the g -priors for the parameters σ_M^2, σ_Y^2 . Indeed for both models under comparison have a common parameter, the variance of the noise. If this parameter shares the same prior distribution $\frac{1}{\sigma^2}$ for both models then the normalization issue vanishes (see Marin and Robert, 2014, Section 3.4.3).*

Decision rule from Bayes factor The $BF_{10} > 1$ threshold gives evidence of $[\mathcal{M}_1]$ over $[\mathcal{M}_0]$. Different scales have been proposed to quantify the evidence of \mathcal{M}_1 over \mathcal{M}_0 from the value of the Bayes factor (see Jeffreys, 1961; Kass and Raftery, 1995). However, these scales have no validation to justify their use to any model. We propose two ways to adapt the quantification of the evidence to a specific model selection problem with nested models $\mathcal{M}_1 \subset \mathcal{M}_0$.

First, we can consider the predictive distribution of BF_{10} under \mathcal{M}_1 . More precisely, this is the distribution of $BF_{10}(Z^*)$ where Z^* is distributed according to the predictive distribution under the model \mathcal{M}_1

$$p_1(z^*|z) = \int f_1(z^*|z, \theta) \pi_1(\theta|z) d\theta,$$

where f_1, π_1 are respectively the conditional likelihood and the posterior distribution under model \mathcal{M}_1 and z the observation. This distribution provides the behavior of the Bayes factor for predicted samples Z^* under \mathcal{M}_1 , which behave similarly to the original sample Z . The probability $\mathbb{P}(BF_{10}(Z^*) > 1|Z)$ quantifies the evidence of \mathcal{M}_1 over \mathcal{M}_0 .

A frequentist alternative to this approach is to compute a p-value associated to the Bayes factor. For linear regression, Zhou and Guan (2018) calculated this p-value from the asymptotic distribution of Bayes factor. This distribution can be also estimated by parametric bootstrap since the models are

nested. Indeed, by estimating θ in \mathcal{M}_1 , we get a consistent estimate of all components of θ , independently of the true model.

In the context of moderation, the explicit form of the Bayes obtained in Proposition 1 makes approximation of these distributions easier.

4 Application to empowering leadership study in both companies

The aim is to explain the organizational commitment (Y) as a function of the empowering leadership. In particular, we want to understand if a part of the effect passes through well-being (M). We are also interested in the variations of these effects between two activity sectors (represented by W). For this study, the data come from a statistical survey carried out on 255 employees in the aeronautical industry and 211 firemen in the area of Nantes, France. The empowering leadership is evaluated through four dimensions: meaning at work (X_1), participation in decision-making (X_2), confidence (X_3), and autonomy (X_4). The well-being, the organizational commitment and the empowering leadership are not directly observed, but they are measured by survey questionnaires (see Table 1). This dataset has been studied in Caillé et al. (2020), where details are given on the sample collection and issues from a psychology perspective.

empowering leadership Ahearne et al. (2005)				well-being Gilbert et al. (2011)	organizational commitment Meyer et al. (1993)
X_1	X_2	X_3	X_4	M	Y
3	3	3	3	22	6

Table 1: Number of survey questions for each variable of interest.

The scatterplots in Figure 3 show the associations between the variables X, M, Y . We assume that these variables satisfy the mediation model (2).

In the absence of prior information on the parameters, we choose hyperparameters of g -prior so that they will be non informative. So, we take $\tilde{\beta} = 0$ and $g = n$ for both the mediation model and the moderate mediation model.

Moderation analysis: Using the moderation model defined in (3), we analyse the effect of activity sector W . We assume that all the paths between empowering leadership, well-being at work and organizational commitment are moderated by the activity sector. We compare the model without interaction $[\mathcal{M}_0]$ and the overall model $[\mathcal{M}_1]$ using the Bayes factor. Thanks to Proposition 1, we have an explicit form of the Bayes factor and so we easily get its value $\Lambda_n = \log_{10}(BF_{10}) = -8.635$. The evidence is in favor of $[\mathcal{M}_0]$ i.e. the activity sector does not moderate the associations between the variables X, M, Y . To confirm this conclusion, we evaluate the predictive distribution of $\log(BF_{10})$ as described in Section 3 (see Figure 4). It is negative with a high probability

$$\mathbb{P}(\log(BF_{10}(M^*, Y^*)) < 0 | M, Y) = 0.96.$$

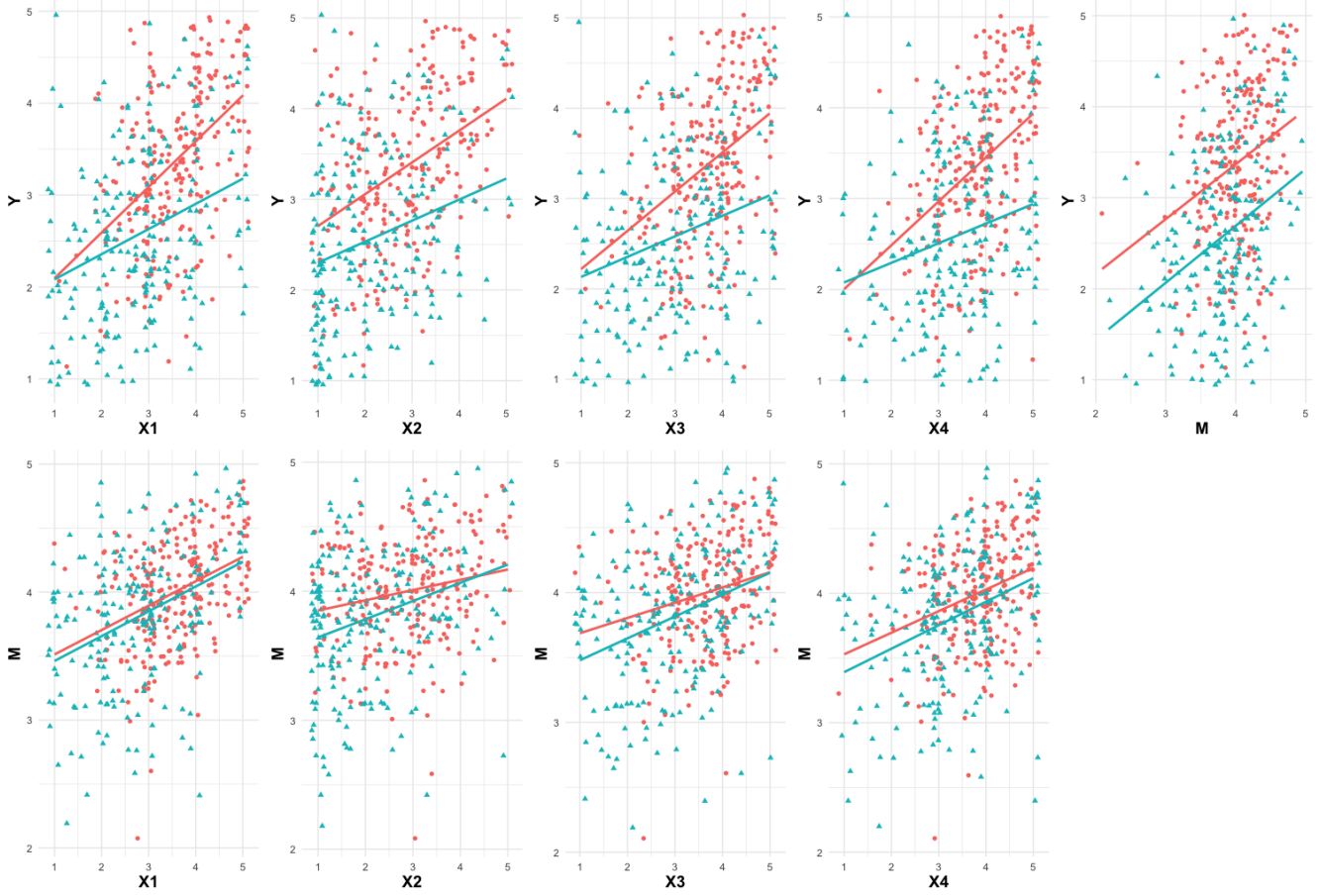


Figure 3: Representation of scatterplots of exposition variables, mediator and outcome (X_1, X_2, X_3, X_4, M, Y).

This leads to a strong evidence of 96% in favor of \mathcal{M}_0 .

We can compare this approach with the frequentist test based on the Bayes factor. Its distribution under the null hypothesis [\mathcal{M}_0] is represented in Figure 5. The p-value is equal to $p = .057$. This test at significant level 5% leads to the same conclusion on the absence of moderate effects.

Caillé et al. (2020) apply a likelihood ratio test, and they obtained a p-value equal to $p = .058$. The two frequentist tests give very similar results, and lead to the same conclusion as the Bayes factor.

Table 2 gives the Bayes estimate and HPD interval for all parameters of the moderation model. For all interaction parameters, the HPD intervals contain the value 0, and so these coefficients are not significant at level 5%. This confirms again the absence of moderation.

Mediation model Given the absence of moderation, we focus on the mediation model (2) estimated

	parameter	Estimate	Lower 95%	Upper 95%	
θ_M	Intercept	3.165	2.844	3.485	
	W	-0.104	-0.514	0.267	
	X_1	0.167	0.074	0.260	
	X_2	-0.030	-0.107	0.044	
	X_3	-0.017	-0.115	0.083	
	X_4	0.100	-0.007	0.205	
	$X_1 : W$	-0.033	-0.156	0.094	
	$X_2 : W$	-0.001	-0.112	0.117	
	$X_3 : W$	0.067	-0.065	0.198	
	$X_4 : W$	-0.013	-0.156	0.120	
	θ_Y	Intercept	0.352	-0.612	1.310
		W	-0.139	-1.331	1.128
X_1		0.235	0.076	0.400	
X_2		0.121	-0.006	0.245	
X_3		0.095	-0.071	0.263	
X_4		0.145	-0.037	0.325	
M		0.231	-0.025	0.484	
$X_1 : W$		-0.118	-0.332	0.097	
$X_2 : W$		-0.048	-0.244	0.135	
$X_3 : W$		-0.059	-0.282	0.165	
$X_4 : W$		-0.157	-0.392	0.076	
$M : W$		0.244	-0.092	0.572	

Table 2: Bayes estimates and HPD interval under the moderate model $[\mathcal{M}_1]$ defined in (3).

on the full population corresponding to the employees of the two companies. From this model, we are testing the existence of the direct and indirect effects of each component. Recall that the direct effect of component X_k on Y is c_k and its indirect effect is the product $a_k b$.

The interest of the Bayesian analysis is to easily infer the indirect effect from the posterior distribution of the product $a_k b$. In particular, for testing the absence of indirect effect at level α , a possible decision rule is to accept the absence of effect if 0 belongs in the $100(1-\alpha)\%$ HPD interval of $a_k b$ (see Biesanz et al., 2010, for Gaussian priors). The Bernstein-von Mises theorem implies that this critical region provides test of asymptotic level α .

The classical approach to test the indirect effect is to estimate the distribution of the product $a_k b$ using bootstrap (see Preacher and Hayes, 2004). Table 3 presents the results of these two approaches. The conclusions are the same: there are direct effects of X_1, X_2 and indirect effects of X_1, X_4 . Therefore, each dimension of empowering leadership contribute differently to the prediction of organizational commitment.

The indirect testing problem can be also interpreted as a problem of model selection. We have to take

into account four competing models for testing the indirect effect of X_j on Y . They are defined as follows, the full model

$$[\mathcal{Med}_4^{(j)}] : \begin{cases} Y = i_2 + c^T X + bM + \varepsilon_Y \\ M = i_1 + a^T X + \varepsilon_M, \end{cases}$$

that corresponds to the existence of the indirect effect and the three nested models of $[\mathcal{Med}_4^{(i)}]$

$$\begin{aligned} [\mathcal{Med}_1^{(j)}] &: \begin{cases} b = 0 \\ a_i = 0 \end{cases}, \\ [\mathcal{Med}_2^{(j)}] &: a_i = 0, \\ [\mathcal{Med}_3^{(j)}] &: b = 0. \end{aligned}$$

Previously, we used the Bayes factor to compare two competing models. To extend to more than two models, the model index $\mathbf{m}^{(j)}$ is added to the unknown parameters (see for instance Marin and Robert, 2014, section 2.3.1). The posterior distribution of $\mathbf{m}^{(j)}$ is

$$\mathbb{P}(\mathbf{m}^{(j)} = i | X, M, Y) = \frac{\mathbb{P}(\mathbf{m}^{(j)} = i) \mathfrak{M}_i^{(j)}(m, y | X)}{\sum_{k=1}^4 \mathbb{P}(\mathbf{m}^{(j)} = k) \mathfrak{M}_k^{(j)}(m, y | X)}, \quad i = 1, \dots, 4,$$

where $\mathfrak{M}_i^{(j)}$ is the marginal distribution of model $[\mathcal{Med}_i^{(j)}]$ and $\mathbb{P}(\mathbf{m}^{(j)} = i)$ its prior probability. Using (11) and (13), we easily obtain an explicit form of the marginal distribution $\mathfrak{M}_i^{(j)}$. The decision rule is to select the model with the higher posterior probability

$$\mathbf{m}_*^{(j)} = \operatorname{argmin}_{i=1, \dots, 4} \mathbb{P}(\mathbf{m}^{(j)} = i | X, M, Y).$$

If the selected model is $[\mathcal{Med}_4^{(j)}]$ (i.e. $\mathbf{m}_*^{(j)} = 4$), then there is evidence for the existence of the indirect effect of X_j on Y .

Remark 4. Note that Nuijten et al. (2015) obtained the same decision rule by writing the evidence in favour of the mediation model as a comparison between two independent regression models. The concordance between the two approaches is explained by the product form of the marginal distribution given in (11).

In our application, we choose the uniform prior for $\mathbf{m}^{(j)}$ since there is no objective argument in favour of one of the models. Table 4 presents the posterior distribution of $\mathbf{m}^{(j)}$. This approach leads to the same decisions as the approaches based on HPD and bootstrap intervals.

In conclusion, the dimensions of empowerment leadership X_1, X_4 have an indirect effect on organizational commitment (Y) through the well-being (M). In contrast, there is absence of indirect effects of the variables X_2 (resp. X_3) on Y due to the absence of association between X_2 (resp. X_3) and the mediator M .

		Bayesian estimation			Bootstrap estimation		
		Estimate	Lower95%	Upper95%	Estimate	Lower95%	Upper95%
direct effects of	X_1	0.161	0.055	0.270	0.161	0.036	0.280
	X_2	0.111	0.015	0.207	0.112	0.017	0.211
	X_3	0.050	-0.066	0.158	0.051	-0.070	0.170
	X_4	0.049	-0.068	0.167	0.049	-0.075	0.177
indirect effects of	X_1	0.055	0.023	0.090	0.055	0.028	0.095
	X_2	-0.012	-0.035	0.011	-0.012	-0.034	0.005
	X_3	0.008	-0.018	0.034	0.008	-0.015	0.036
	X_4	0.035	0.006	0.065	0.035	0.009	0.071

Table 3: Comparison of Bayesian estimation (Bayes estimate and HPD interval) with the bootstrapping estimation under mediation model $[\mathcal{M}_0]$.

		$\mathbb{P}(\mathbf{m}^{(j)} = i X, M, Y)$			
		1	2	3	4
indirect effects of	X_1	0.000	0.000	0.002	0.997
	X_2	0.002	0.920	0.000	0.078
	X_3	0.002	0.945	0.000	0.053
	X_4	0.001	0.362	0.001	0.636

Table 4: Posterior distribution of $\mathbf{m}^{(j)}$ for testing the existence of the indirect effect of each exposure X_j on Y , $j = 1, \dots, 4$. The probabilities in bold indicate $\mathbf{m}_*^{(j)}$.

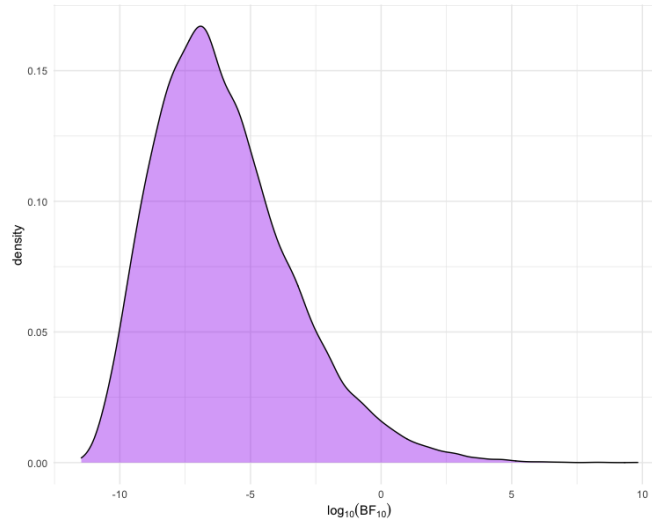


Figure 4: Predictive distribution of $\log(BF_{10})$ under the moderation model $[\mathcal{M}_1]$.

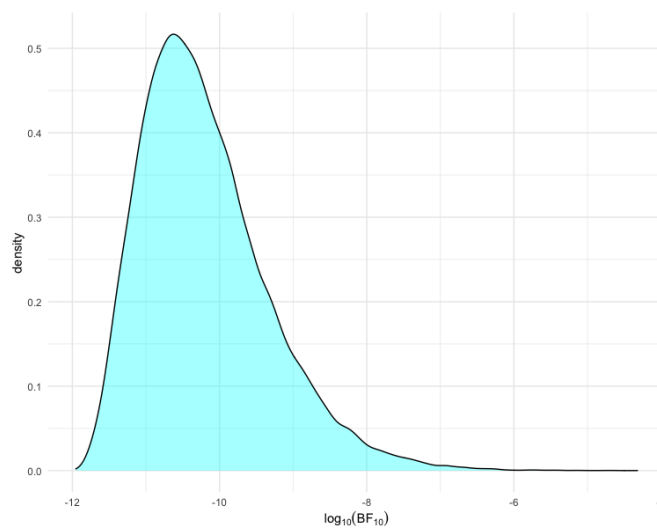


Figure 5: Approximation of the null distribution of $\log(BF_{10})$ by parametric bootstrap. For the value -8.635 of the Bayes factor calculated on the observed dataset, the p-value is equal to $p = .057$.

References

- Ahearne, M., Mathieu, J., and Rapp, A. (2005). To empower or not to empower your sales force? an empirical examination of the influence of leadership empowerment behavior on customer satisfaction and performance. *Journal of Applied Psychology*, 90:945–955.
- Baron, R. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):11–73.
- Biesanz, J. C., Falk, C. F., and Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects. *Multivariate Behavioral Research*, 45:661–701.
- Bollen, K. A. (2014). *Structural Equations with Latent Variables*. Wiley.
- Caillé, A., Courtois, N., Galharret, J.-M., and Jeoffrion, C. (2020). Influence du leadership habilitant sur le bien-être au travail et l’engagement organisationnel : étude comparative entre une organisation habilitante et une organisation classique. *Psychologie du Travail et des Organisations*, 26(3):247 – 261.
- Gilbert, M.-H., Dagenais-Desmarais, V., and Savoie, A. (2011). Validation d’une mesure de santé psychologique au travail. *European Review of Applied Psychology*, 61:195–203.
- Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-based Approach*. Methodology in the Social Sciences. Guilford Press.
- Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press.
- Jose, P. E. (2013). *Doing Statistical Mediation and Moderation*. Guilford Press, 1st edition.
- Judd, C. M. and Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5:602–619.
- Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90:773–795.
- MacKinnon, D. (2008). *Introduction to Statistical Mediation Analysis*. Multivariate Applications Series. Routledge Academic.
- MacKinnon, D. P., Valente, M. J., and Gonzalez, O. (2020). The correspondence between causal and traditional mediation analysis: the link is the mediator by treatment interaction. *Prevention Science*, 21(2):147–157.
- Marin, J.-M. and Robert, C. (2014). *Bayesian essentials with R*. Springer Textbooks in Statistics. Springer Verlag, New York.

- Meyer, J. P., Allen, N. J., and Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*, 78:538–551.
- Miočević, M., Gonzalez, O., Valente, M. J., and MacKinnon, D. P. (2018). A tutorial in bayesian potential outcomes mediation analysis. *Struct Equ Modeling*, 25:121–136.
- Muthén, B., Muthén, L., and Asparouhov, T. (2016). *Regression and Mediation Analysis Using Mplus*. Muthén & Muthén.
- Nuijten, Michèle B. and Wetzels, R., Matzke, D., Dolan, C. V., and Wagenmakers, E.-J. (2015). A default bayesian hypothesis test for mediation. *Behavior Research Methods*, 47:85–97.
- Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, UAI'01, pages 411–420. Morgan Kaufmann Publishers Inc.
- Preacher, K. J. and Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods*, 36:717–731.
- Preacher, K. J., Rucker, D., D., and Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42:185–227.
- Robins, J. (2003). Semantics of causal dag models and the identification of direct and indirect effects. In *Highly Structured Stochastic Systems*, Eds P.J. Green, N.L. Hjort, S. Richardson. Oxford University Press.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13:290–312.
- VanderWeele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press.
- Wang, Lijuan (Peggy) and Preacher, K. J. (2015). Moderated mediation analysis using bayesian methods. *Structural Equation Modeling A Multidisciplinary Journal*, 22:249–263.
- Yuan, Y. and MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14:301–322.
- Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. Wiley.
- Zellner, A. (1984). *Basic Issues in Econometrics*. University of Chicago Press.
- Zhou, Q. and Guan, Y. (2018). On the null distribution of bayes factor in linear regression. *Journal of the American Statistical Association*, 113(523):1362–1371.