Econometrics and Statistics 27 (2023) 161–172

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Bayesian analysis for mediation and moderation using g−priors.

Jean-Michel Galharret, Anne Philippe*

Nantes Université, CNRS, Laboratoire de Mathématiques Jean Leray, LMJL, Nantes F-44000, France

ARTICLE INFO

Article history:
Received 9 January 2021
Revised 22 December 2021
Accepted 29 December 2021
Available online 12 January 2022

Keywords:
Bayes factor
direct and indirect effects
g−priors

ABSTRACT

A Bayesian analysis is proposed using an extension of g-priors for moderated mediation models. For this choice of priors, an explicit form of the marginal distribution is obtained. Testing procedure on the existence of direct, indirect and moderated effects are constructed using Bayes factor approach. This methodology is applied to analyze the association between empowering leadership and organisational commitment in two companies.

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 work place well-being. This mediation analysis is performed on two populations corresponding to different companies. We are interested in the effect of the company factor on the associations between the previous variables. From a statistical point of view, such problems can be addressed by mediation analysis and moderated mediation analysis.

Mediation analysis is one of the classical 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 M, Y are continuous, the most widely used model 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:

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

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

* Corresponding author.
E-mail addresses: jean-michel.galharret@univ-nantes.fr (J.-M. Galharret), anne.philippe@univ-nantes.fr (A. Philippe).

https://doi.org/10.1016/j.ecosta.2021.12.009

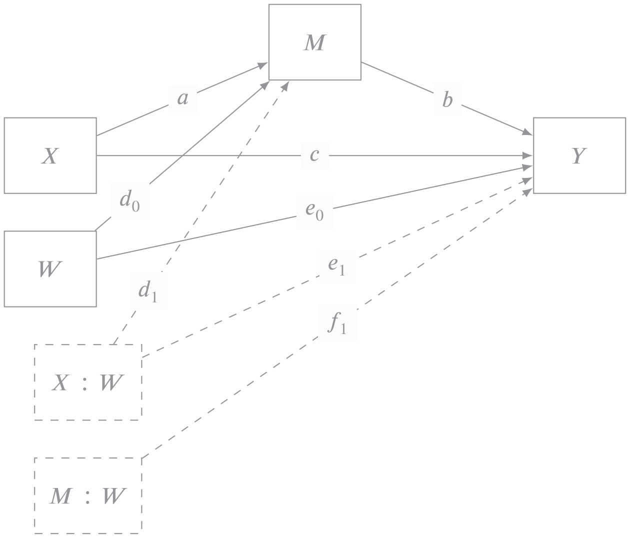

Fig. 1. Mediation model (solid lines) and moderated mediation model (with dashed lines added).

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 X and Y. The correspondance between the causal and SEM mediation analyses is discussed in MacKinnon et al. (2020). Consider a generalisation of the mediation model (1) defined as follows:

$$[\mathcal{M}_0]: \begin{cases} Y = i_2 + Xc + bM + e_0W + \varepsilon_Y \\ M = i_1 + Xa + d_0W + \varepsilon_M, \end{cases} \tag{2}$$

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 exposure vectors and W is a covariate. The exposure 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 exposure 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 work place well-being and W is the activity sector (see Section 4 for details). To evaluate the effect of the populations, we consider a 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 $X_1, \ldots, X_p$ and W. Severals 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

$$[\mathcal{M}_1]: \begin{cases} Y = i_2 + Xc + bM + e_0W + X : We_1 + f_1M : W + \varepsilon_Y, \\ M = i_1 + Xa + d_0W + X : Wd_1 + \varepsilon_M, \end{cases} \tag{3}$$

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, X on Y and 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

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

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

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

where $\widehat{\theta}_1$ and $\widehat{\theta}_0$ 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.

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 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. For each linear regression model, Nuijten et al. (2015) use the $g-$priors distribution introduced by Zellner (1984). They test the indirect effect by combining Bayes factors calculated independently on these two regression equations that define (1). More precisely, for testing that the product term $\alpha\beta$ is equal to zero, they calculate two Bayes factor, one for testing $\alpha = 0$ and another for testing $\beta = 0$. Then, the final decision rule is obtained as the product of these two Bayes factors. This choice is not well justified as it is based on an assumption of independence between the two regression equations. This approach does not take into account the fact that mediation is defined by a system of linear equations. This aspect is however fundamental since the mediator M is an explanatory variable in one equation and a response variable in the other. In a more general context than Nuijten et al. (2015), including a covariate W and interaction terms $X : W$ and $M : W$, we show that it is possible to address this problem from the joint distribution of (Y, M) given by both equations in (2). This approach makes it possible to obtain the joint posterior distribution of all parameters from which one can deduce posterior distribution of direct and indirect effect. Unlike Nuijten et al. (2015), the Bayes factor for testing the absence of effects is well defined in a standard way. The Bayesian method is also studied for moderation. For instance, Wang and Preacher (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 $g-$priors 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 $g-$priors 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, $Y \in \mathbb{R}^n$, ε is a zero mean Gaussian vector with independent coordinates. We denote σ^2 the variance of the Gaussian noise ε. The Zellner's $g-$priors on the parameters (β, σ^2) are of the form

$$\pi_g(\beta, \sigma^2) = \pi_g(\beta|\sigma^2)\sigma^{-2},$$

where the conditionnal distribution of β given σ^2 is a Gaussian distribution with covariance matrix proportional to the inverse Fisher information of β. Recall that the Fisher information for the coefficient β is

$$\mathcal{I}(\beta) = \sigma^{-2}\mathbb{X}^T\mathbb{X}, \tag{4}$$

and therefore we have

$$\beta|\sigma^2 \sim \mathcal{N}_p(\tilde{\beta}, g\sigma^2(\mathbb{X}^T\mathbb{X})^{-1}). \tag{5}$$

To designate the $g-$priors we use the notation

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

This prior depends on hyper-parameters $\widetilde{\beta} \in \mathbb{R}, g \in \mathbb{R}$ to be fixed according to the prior information.

Our objective is to adapt the $g-$priors to the context of mediation models. A first attempt was made by Nuijten et al. (2015) who construct the $g-$priors by considering 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.

The challenge is to define a joint $g-$priors on all the parameters of the mediation model gathering the parameters of the two equations defined in (2), that is $\theta := (\theta_M, \sigma_M^2, \theta_Y, \sigma_Y^2)$, where $\theta_M = (i_1, a, d_0)$ and $\theta_Y = (i_2, c, b, d_0)$.

Proposition 1. *The Fisher information for the regression coefficients (θ_Y, θ_M) of the mediation model defined in (2) is*

$$\mathcal{I}(\theta_Y, \theta_M) = \begin{pmatrix} \frac{1}{\sigma_Y^2}\ \mathbb{X}^T\mathbb{X} & \frac{1}{\sigma_Y^2}\ \mathbb{X}^T\mathbb{X}\theta_M & 0 \\ \frac{1}{\sigma_Y^2}\ \theta_M^T\mathbb{X}^T\mathbb{X} & \frac{1}{\sigma_Y^2}(\ \theta_M^T\ \mathbb{X}^T\mathbb{X}\theta_M + \sigma_M^2 I) & 0 \\ 0 & 0 & \frac{1}{\sigma_M^2}\ \mathbb{X}^T\mathbb{X} \end{pmatrix},$$

where $\mathbb{X}$ denotes the design matrix $[\mathbb{1}, X, W]$.

Proof. In the model (2), the joint distribution of Y, M is:

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

where f_Y is the Gaussian distribution with mean $i_2+Xc+bm+e_0W$ and variance σ_Y^2 and where f_M is the Gaussian distribution with mean i_1+Xa+d_0W and variance σ_M^2. Then the log likelihood function can be split as the sum of two functions $\mathcal{L}(\theta_Y,\sigma_Y^2)$ and $\mathcal{L}(\theta_M,\sigma_M^2)$,

$$\begin{aligned}\log V(\theta) &= \log f_Y(Y|M,\theta_Y,\sigma_Y^2,X,W)+\log f_M(M|\theta_M,\sigma_M^2,X,W)\\ &:= \mathcal{L}(\theta_Y,\sigma_Y^2)+\mathcal{L}(\theta_M,\sigma_M^2).\end{aligned} \tag{6}$$

Note that $\mathcal{L}(\theta_Y,\sigma_Y^2)$ and $\mathcal{L}(\theta_M,\sigma_M^2)$ are respectively the log likelihood functions of the two Gaussian regression models appearing in (2). According to (4) and (6), we deduce that

$$\mathcal{I}(\theta_Y,\theta_M)=\begin{pmatrix}\frac{1}{\sigma_Y^2}\mathbb{E}([\mathbb{X},M]^T[\mathbb{X},M]) & 0\\ 0 & \frac{1}{\sigma_M^2}\ \mathbb{X}^T\mathbb{X}\end{pmatrix}.$$

Since

$$\begin{aligned}\mathbb{E}(M)&=\mathbb{X}\theta_M\\ \mathbb{E}(\ M^TM)&=\mathbb{E}(\ (\mathbb{X}\theta_M+\varepsilon_M)^T(\mathbb{X}\theta_M+\varepsilon_M))=\ \theta_M^T\ \mathbb{X}^T\mathbb{X}\theta_M+\sigma_M^2I_n,\end{aligned}$$

we have

$$\mathbb{E}(\ [\mathbb{X},M]^T.[\mathbb{X},M])=\mathbb{E}\begin{pmatrix}\mathbb{X}^T\mathbb{X} & \mathbb{X}^TM\\ M^T\mathbb{X} & M^TM\end{pmatrix}=\begin{pmatrix}\mathbb{X}^T\mathbb{X} & \mathbb{X}^T\mathbb{X}\theta_M\\ \theta_M^T\ \mathbb{X}^T\mathbb{X} & \theta_M^T\ \mathbb{X}^T\mathbb{X}\theta_M+\sigma_M^2I_n\end{pmatrix}.$$

This concludes the proof. □

According to Proposition 1, the Fisher information depends on the parameter θ_M. Therefore, the g–prior cannot be defined as previously where the conditional distribution of the coefficients (θ_Y,θ_M) is a Gaussian distribution whose covariance matrix is proportional to the inverse of the Fisher information.

It is also not possible to assume that the parameters of both equations in (2) are independent

$$\pi(\theta)=\pi_1(\theta_M,\sigma_M^2)\pi_2(\theta_Y,\sigma_Y^2),$$

and to choose for π_1,π_2 the g–priors defined in (5) with design matrix $[\mathbb{1},X,W]$ and $[\mathbb{1},X,W,M]$ respectively. Indeed, the random variable M appears in the design matrix of π_2.

To avoid this problem we propose a new strategy whose main idea is to the decompose the joint distribution of M,Y,θ as follows,

$$\begin{aligned}p(m,y,\theta_Y,\sigma_Y^2,\theta_M,\sigma_M^2|X,W)\ =\ & f_Y(y|M=m,\theta_Y,\sigma_Y^2,\theta_M,\sigma_M^2,X,W)\pi_Y(\theta_Y,\sigma_Y^2|M=m,\theta_M,\sigma_M^2,X,W)\times\\ & f_M(m|\theta_M,\sigma_M^2,X,W)\pi_M(\theta_M,\sigma_M^2|X,W).\end{aligned}$$

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

$$\begin{aligned}p(m,y,\theta_Y,\sigma_Y^2,\theta_M,\sigma_M^2|X,W)\ =\ & f_Y(y|M=m,\theta_Y,\sigma_Y^2,X,W)\pi_Y(\theta_Y,\sigma_Y^2|M=m,X,W)\\ & \times\ f_M(m|\theta_M,\sigma_M^2,X,W)\pi_M(\theta_M,\sigma_M^2|X,W),\end{aligned} \tag{7}$$

where f_Y is the Gaussian distribution with mean $i_2+Xc+bm+e_0W$ and variance σ_Y^2 and where f_M is the Gaussian distribution with mean i_1+Xa+d_0W and variance σ_M^2.

The g–priors defined in (5) can be chosen for π_M and π_Y with a design matrix $[\mathbb{1},X,W]$ and $[\mathbb{1},X,W,M]$ respectively. More precisely, we have

$$\pi_M(\theta_M,\sigma_M^2|X,W)\propto\frac{1}{\sigma_M^{p+4}}\exp\Big(-\frac{1}{2g_M\sigma_M^2}(\theta_M-\widetilde{\theta}_M)^T[\mathbb{1},X,W]^T[\mathbb{1},X,W](\theta_M-\widetilde{\theta}_M)\Big) \tag{8}$$

$$\pi_Y(\theta_Y,\sigma_Y^2|X,W,M)\propto\frac{1}{\sigma_Y^{p+5}}\exp\Big(-\frac{1}{2g_Y\sigma_Y^2}(\theta_Y-\widetilde{\theta}_Y)^T[\mathbb{1},X,W,M]^T[\mathbb{1},X,W,M](\theta_Y-\widetilde{\theta}_Y)\Big), \tag{9}$$

where $g_M,\widetilde{\theta}_M,g_Y,\widetilde{\theta}_Y$ are the hyperparameters of the g–priors. Equations (7), (8), (9) fully define the Bayesian mediation model with the g–priors. The posterior distribution satisfies

$$\pi(\theta_M,\sigma_M^2,\theta_Y,\sigma_Y^2|X,W,M,Y)\propto p(M,Y,\theta_Y,\sigma_Y^2,\theta_M,\sigma_M^2|X,W), \tag{10}$$

where p is defined in (7).

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.

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

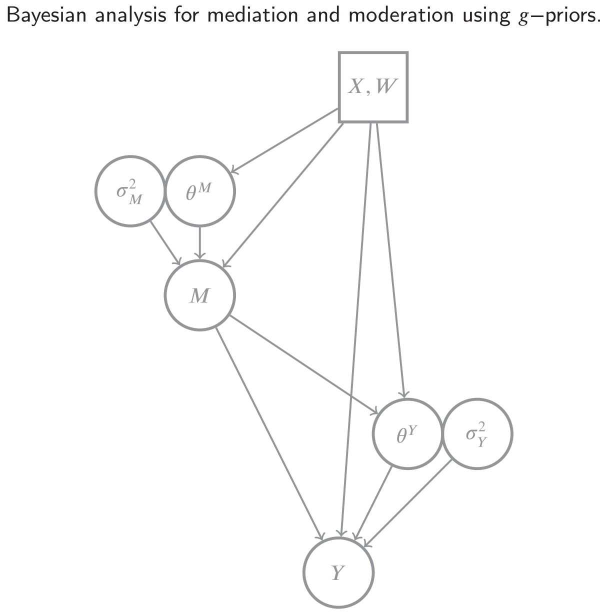

Fig. 2. Directed acyclic graph for the mediation model $[\mathcal{M}_0]$

Discussion of the choice of g**:** These prior distributions cover a wide range of contexts from noninformative to informative. Indeed we can bring information on θ from the hyperparameters $\widetilde{\theta}_M, \widetilde{\theta}_Y$, and g_M, g_Y quantify the weight of this information. In absence of information, $\widetilde{\theta}_M = \widetilde{\theta}_Y = 0$ and $g_M = g_Y = n$ are the standard choices recommended for instance by Kass and Wasserman (1995). Such choices are motivated by the fact that this $g-$prior provides the same amount of information on the regression parameters as one observation. Many articles deal with the choice of g parameter in the particular case of Gaussian linear regression model. In Berger and Pericchi (2001) and George and Foster (2000), the empirical Bayes approach is considered in choosing g. The parameter g can be seen as an unknown parameter. In this case, we have to choose prior on g, for instance, the Zellner-Siow prior (see Zellner and Siow, 1980) and the hyper $g-$priors (see Cui and George, 2008). In Remark 4, we also comment on the impact of the choice of g on the behavior of the Bayes factor.

Inference on the effects: In applications, one of the main issues of statistical analysis 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 (10). 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 highest posterior density (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 moderated effects: the moderate direct effect $c + e_1$ and the moderate indirect effect $(a + d_1)(b + f_1)$.

An alternative approach 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 moderated 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 $[\mathcal{M}_1]$. For both models, 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 $[\mathcal{M}_i]$, $i \in \{0, 1\}$,

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

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

$$BF_{10} = \frac{\mathfrak{M}^1(y)}{\mathfrak{M}^0(y)}. \tag{11}$$

The evidence of $[\mathcal{M}_1]$ over $[\mathcal{M}_0]$ corresponds to $BF_{10} > 1$.

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

- Model $\mathcal{M}_1$:

$$\begin{aligned} Y &= i_2 + Xc + bM + e_0W + X:We_1 + f_1M:W + \varepsilon_Y, \\ (\theta_Y, \sigma_Y^2)|X,W,M &\sim \mathcal{Z}_g(g_2, \widetilde{\beta}_2, \mathbb{X}_{Y_1}) \text{ with } \mathbb{X}_{Y_1} = \left[\mathbb{1}, W, X, M, X:W, M:W\right], \\ M &= i_1 + Xa + d_0W + X:Wd_1 + \varepsilon_M, \\ (\theta_M, \sigma_M^2)|X,W &\sim \mathcal{Z}_g(g_1, \widetilde{\beta}_1, \left[\mathbb{1}, W, X, X:W\right]) \text{ with } \mathbb{X}_{M_1} = \left[\mathbb{1}, W, X, X:W\right]. \end{aligned}$$

- Model $\mathcal{M}_0$:

$$\begin{aligned} Y &= i_2 + Xc + bM + e_0W + \varepsilon_Y, \\ (\theta_Y, \sigma_Y^2)|X,W,M &\sim \mathcal{Z}_g(g_2, \check{\beta}_2, \left[\mathbb{1}, W, X, M\right]) \text{ with } \check{\beta}_2 = (\widetilde{\beta}_2^1, \dots, \widetilde{\beta}_2^{p+3}) \text{ and } \mathbb{X}_{Y_0} = \left[\mathbb{1}, W, X, M\right] \\ M &= i_1 + Xa + d_0W + \varepsilon_M, \\ (\theta_M, \sigma_M^2)|X,W &\sim \mathcal{Z}_g(g_1, \check{\beta}_1, \left[\mathbb{1}, W, X\right]) \text{ with } \check{\beta}_1 = (\widetilde{\beta}_1^1, \dots, \widetilde{\beta}_1^{p+2}) \text{ and } \mathbb{X}_{M_0} = \left[\mathbb{1}, W, X\right]. \end{aligned}$$

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

Proposition 2. *The Bayes factor to compare the models $\mathcal{M}_1$ versus $\mathcal{M}_0$ defined above is given by*

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

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

$$BF_{10} = \frac{\mathfrak{M}^1(m, y|X, W)}{\mathfrak{M}^0(m, y|X, W)}. \tag{12}$$

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

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

where p is defined in (7). It can be rewritten

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

where f_Y, f_M, π_Y, π_M in (7), (8), (9). By Fubini theorem, we obtain

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

From the last equation, we can identify $\mathfrak{M}^0$ as the product of the marginal densities of both linear regression models defining $\mathcal{M}_0$,

$$\mathfrak{M}^0(m, y|X, W) = \mathfrak{M}_M^0(m|X, W) \times \mathfrak{M}_Y^0(y|X, W, M), \tag{13}$$

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 equation 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

$$BF_{10} = BF_{10}^{M|X} \times BF_{10}^{Y|M,X}, \tag{14}$$

where

- $BF_{10}^{M|X}$ is the Bayes factor of competing models $M = i_1 + Xa + d_0W + X : Wd_1 + \varepsilon_M$ versus $M = i_1 + Xa + d_0W + \varepsilon_M$,
- $BF_{10}^{Y|M,X}$ is the Bayes factor of competing models $Y = i_2 + Xc + bM + e_0W + X : We_1 + f_1M : W + \varepsilon_Y$ versus $Y = i_2 + Xc + bM + e_0W + \varepsilon_Y$.

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

$$p(y) = (g+1)^{-(p+1)/2}\pi^{-n/2}\Gamma(n/2)\left(y^Ty - \frac{g}{g+1}y^T\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^Ty - \frac{\|\mathbb{X}\widetilde{\beta}\|^2}{g+1}\right)^{-n/2}. \tag{15}$$

Using (14), we deduce the expression announced in the statement of the Proposition and this concludes the proof. □

Remark 3. The Bayes factor is not defined when the prior distributions 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, the two competing models have the variance of the noise as a common parameter. If this parameter has the same prior distribution $\frac{1}{\sigma^2}$ in both models then the normalization issue disappears (see Marin and Robert, 2014, Section 3.4.3).

Remark 4. The choice of g parameter has a strong impact on the behavior of the Bayes factor. When g is fixed, the Bayes factor suffers from a consistency issues, for instance, when the coefficient determination R^2 tends to 1 (see Berger and Pericchi, 2001). The use of Bayes' empirical approach or of a hierarchical prior on g (already mentioned in Section 2) brings a solution to this problem (see Liang et al., 2008, for Gaussian regression model). For $g = n$ in regression, the Bayes factor behaves like the well-known model selection criterion, the BIC (see Kass and Wasserman, 1995). An alternative approach is to use a hierarchical prior on g such as the Zellner-Siow prior (see Zellner and Siow, 1980) or the hyper $g-$priors (see Cui and George, 2008). Proposition 2 provides the explicit form of the Bayes factor for $g_1 = g_2 = n$ or the empirical Bayes approach. For hierarchical priors on g, this is generally no longer the case. For Zellner-Siow prior and hyper-g prior, the Bayes factor can be expressed from one-dimensional integral or special functions, that can be easily approximated using standard numerical integration techniques (see Liang et al., 2008, for these expressions). For more complex prior, the bridge sampling method can be used to approximate Bayes factor (see Gronau et al., 2017, for a tutorial).

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 rationale validation to justify their use in any model. We propose two alternative approaches to these scales to try to quantify the evidence of $\mathcal{M}_1$ over $\mathcal{M}_0$ taking into account the specific form of the model. The two proposed approaches use the fact that we compare nested models $\mathcal{M}_0 \subset \mathcal{M}_1$. The first is based on the Bayes factor BF_{10} distribution under the predictive distribution. The second consists in constructing a frequentist test where the test statistic is the Bayes factor. The construction of test statistics from Bayesian inference is also discussed, for instance in Chen and Walker (2021). For two-sided tests they consider as a test statistic, the Kullbach-Leibler divergence between the expected posterior under the null and the observed posterior.

Approach 1. To quantify the evidence of $\mathcal{M}_1$ over $\mathcal{M}_0$, we would like to evaluate the probability that $BF_{10} > 1$. García-Donato and Chen (2005) propose a criterium based on the prior predictive distribution $\mathfrak{M}$. This approach cannot be extended to our model comparison. Indeed $\mathfrak{M}$ is not well-defined since the prior distribution on common variance is improper as noted in Remark 3. We suggest in the context of the pseudo Bayes factor to replace the prior predictive distribution by the predictive distribution $p_1(z^\star|Z)$ under the model $\mathcal{M}_1$ defined by

$$p_1(z^\star|Z) = \int f_1(z^\star|Z, \theta)\pi_1(\theta|Z)\,\mathrm{d}\theta,$$

where f_1, π_1 are respectively the conditional likelihood and the posterior distribution under model $\mathcal{M}_1$, and Z is the observed sample. This predictive distribution provides data similar to the observations by taking into account the regions of the parameter space that are most expected a posteriori. The sampling $Z^\star$ according to the predictive distribution $p_1(z^\star|Z)$ proceeds as follows

1. Simulate $\theta \sim \pi_1(\bullet|Z)$,
2. Simulate $Z^* \sim f_1(\bullet|Z, \theta) = f_1(\bullet|\theta)$.

Note that for our moderation model the conditional likelihood does not depend on Z. As the models are nested, the distribution of $BF_{10}(Z^\star)$ should behave similarly to the Bayes factor of the original sample $BF_{10}(Z)$. Therefore, we propose quantifying the evidence of $\mathcal{M}_1$ over $\mathcal{M}_0$ by the probability $\mathbb{P}(BF_{10}(Z^\star) > 1|Z)$. Indeed, this corresponds to the probability of deciding $\mathcal{M}_1$ against $\mathcal{M}_0$ for predictive samples that look like the observed sample Z.

Approach 2. The Bayes factor can also be used as a test statistic, this frequentist approach requires the knowledge of the Bayes factor distribution under the null hypothesis defined by $\mathcal{M}_0$. This approach is studied by Zhou and Guan (2018) in the particular case of linear regression. They build the critical region, and calculate the associated p-value using the asymptotic

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

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

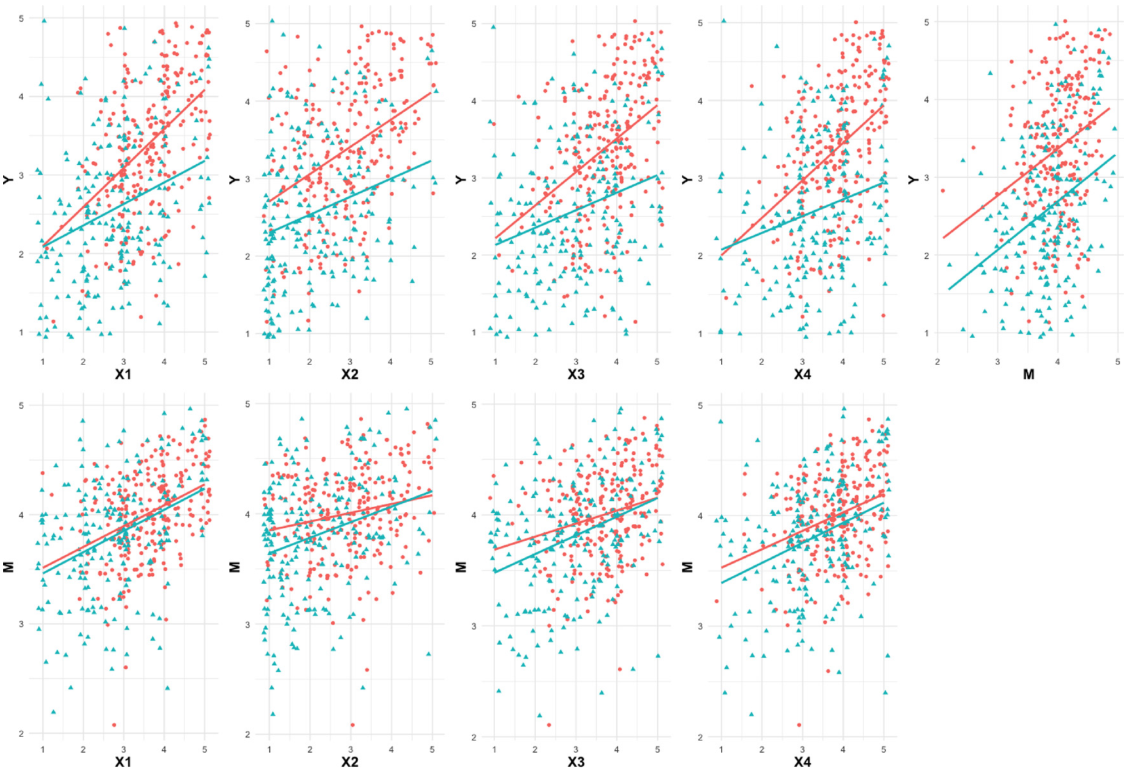

Fig. 3. Representation of scatterplots of exposure variables, mediator and outcome $(X_1, X_2, X_3, X_4, M, Y)$.

distribution of the Bayes factor. The Bayes factor distribution under the null hypothesis can also be estimated by parametric bootstrap since the models are nested. Indeed, let $\mathcal{M}_0 = \{f_{\theta_0},\ \theta_0 \in \Theta_0\}$ and $\mathcal{M}_1 = \{f_{\theta_0,\lambda},\ (\theta_0,\lambda) \in \Theta_0 \times \Lambda\}$ be the two nested parametric families. A consistent parametric estimator $(\widehat{\theta}_0, \widehat{\lambda})$ of (θ_0, λ) for the model $\mathcal{M}_1$ will also be a consistent one for θ_0 if $\lambda = 0$. Therefore, we can approximate the Bayes factor distribution under the null hypothesis using bootstrapping samples from $f_{\widehat{\theta}_0} \in \mathcal{M}_0$. This frequentist context does not allow us to speak about the evidence of $\mathcal{M}_1$ against $\mathcal{M}_0$, but the usual Type I error is used to evaluate the significance of λ.

In the context of moderation, the explicit form of the Bayes obtained in Proposition 2 makes approximation of these distributions easier.

4. Application to the case study of empowering leadership in two different 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. 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). Well-being, organizational commitment and empowering leadership are not directly observed, but they are measured by survey questionaries (see Table 1). This dataset has been studied in Caillé et al. (2020).

The scatterplots in Figure 3 clear 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 noninformative. So, we take $\widetilde{\beta} = 0$ and $g = n$ for both the mediation model and the moderate mediation model.

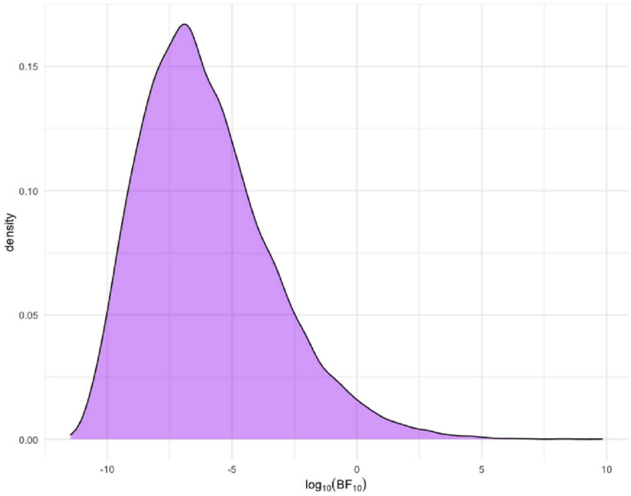

Fig. 4. Predictive distribution of $\log(BF_{10})$ under the moderation model $[\mathcal{M}_1]$.

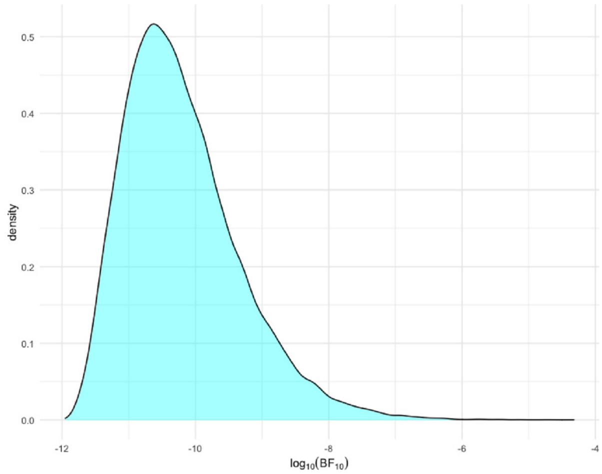

Fig. 5. Approximation of the null distibution of $\log(BF_{10})$ by parametric bootstrap. For the value -8.635 of the Bayes factor calculated on the observed dataset, the p-value is equal to $p = .057$.

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, work place well-being 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 2, 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

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

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

We can compare this approach to 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 $p = .057$. At significance level 5%, this test leads to the same conclusion of 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 on the full population corresponding to the employees of the two companies. From this model, we are testing the existence of the direct

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

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

		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
	$X3$	0.050	-0.066	0.158	0.051	-0.070	0.170
	$X4$	0.049	-0.068	0.167	0.049	-0.075	0.177
indirect effects of	$X1$	0.055	0.023	0.090	0.055	0.028	0.095
	$X2$	-0.012	-0.035	0.011	-0.012	-0.034	0.005
	$X3$	0.008	-0.018	0.034	0.008	-0.015	0.036
	$X4$	0.035	0.006	0.065	0.035	0.009	0.071

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 a 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 boostrap (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 contributes differently to the prediction of organizational commitment.

The indirect testing problem can also be 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

$$[\mathcal{M}ed_4^{(j)}] : \begin{cases} Y = i_2 + Xc + bM + \varepsilon_Y \\ M = i_1 + Xa + \varepsilon_M, \end{cases}$$

that corresponds to the existence of the indirect effect and the three nested models of $[\mathcal{M}ed_4^{(i)}]$

$$[\mathcal{M}ed_1^{(j)}] : \begin{cases} b = 0 \\ a_i = 0 \end{cases},$$

$$[\mathcal{M}ed_2^{(j)}] : \quad a_i = 0,$$

$$[\mathcal{M}ed_3^{(j)}] : \quad b = 0.$$

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

		$\mathbb{P}(\mathfrak{m}^{(j)} = i \mid X, M, Y)$			
		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**

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

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

where $\mathfrak{M}_i^{(j)}$ is the marginal distribution of model $[\mathcal{M}ed_i^{(j)}]$ and $\mathbb{P}(\mathfrak{m}^{(j)} = i)$ is its prior probability. Using (13) and (15), 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

$$\mathfrak{m}_*^{(j)} = \mathrm{argmax}_{i=1,\ldots,4} \mathbb{P}(\mathfrak{m}^{(j)} = i \mid X, M, Y).$$

If the selected model is $[\mathcal{M}ed_4^{(j)}]$ (i.e. $\mathfrak{m}_*^{(j)} = 4$), then there is evidence for the existence of the indirect effect of X_j on Y.

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

Acknowledgment

We thank the Associate Editor and the reviewers for their valuable comments which helped to improve this article.

References

Ahearne, M., Mathieu, J., 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. doi:10.1037/0021-9010.90.5.945.
Baron, R., 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, 11–73.
Berger, J.O., Pericchi, L.R., 2001. Objective Bayesian methods for model selection: Introduction and comparison. In: Institute of mathematical statistics lecture notes - monograph series.
Biesanz, J.C., Falk, C.F., Savalei, V., 2010. Assessing mediational models: Testing and interval estimation for indirect effects. Multivariate Behavioral Research 45, 661–701. doi:10.1080/00273171.2010.498292.
Bollen, K.A., 2014. Structural Equations with Latent Variables. Wiley.
Caillé, A., Courtois, N., Galharret, J.M., 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, 247–261. doi:10.1016/j.pto.2020.05.001.
Chen, S., Walker, S.G., 2021. A new statistic for Bayesian Hypothesis Testing. Econometrics and Statistics doi:10.1016/j.ecosta.2021.10.009. URL https://www.sciencedirect.com/science/article/pii/S2452306221001246
Cui, W., George, E.I., 2008. Empirical Bayes vs. fully Bayes variable selection. Journal of Statistical Planning and Inference 138, 888–900.
García-Donato, G., Chen, M.H., 2005. Calibrating Bayes factor under prior predictive distributions. Statistica Sinica 15, 359–380.
George, E.I., Foster, D.P., 2000. Calibration and empirical Bayes variable selection. Biometrika 87, 731–747.
Gilbert, M.H., Dagenais-Desmarais, V., Savoie, A., 2011. Validation d'une mesure de santé psychologique au travail. European Review of Applied Psychology 61, 195–203. doi:10.1016/j.erap.2011.09.001.
Gronau, Q.F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D.S., Forster, J.J., Wagenmakers, E.J., Steingroever, H., 2017. A tutorial on bridge sampling. Journal of Mathematical Psychology 81, 80–97. doi:10.1016/j.jmp.2017.09.005.
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 Probbility. Oxford University Press.
Jose, P.E., 2013. Doing Statistical Mediation and Moderation, 1st Guilford Press.
Judd, C.M., Kenny, D.A., 1981. Process analysis: Estimating mediation in treatment evaluations. Evaluation Review 5, 602–619. doi:10.1177/0193841x8100500502.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795. doi:10.2307/2291091.
Kass, R.E., Wasserman, L., 1995. A reference Bayesian test for nested hypotheses and its relationship to the schwarz criterion. Journal of the american statistical association 90, 928–934.
Liang, F., Paulo, R., Molina, G., Clyde, M.A., Berger, J.O., 2008. Mixtures of g priors for Bayesian variable selection. Journal of the American Statistical Association 103, 410–423. doi:10.1198/016214507000001337.

MacKinnon, D., 2008. Introduction to statistical mediation analysis. Multivariate applications series. Routledge Academic.
MacKinnon, D.P., Valente, M.J., Gonzalez, O., 2020. The correspondence between causal and traditional mediation analysis: the link is the mediator by treatment interaction. Prevention Science 21, 147–157. doi:10.1007/s11121-019-01076-4.
Marin, J.M., Robert, C., 2014. Bayesian essentials with r. In: Springer textbooks in statistics. Springer Verlag, New York. URL https://hal.archives-ouvertes.fr/hal-01337395
Meyer, J.P., Allen, N.J., Smith, C.A., 1993. Commitment to organizations and occupations: Extension and test of a three-component conceptualization. Journal of Applied Psychology 78, 538–551. doi:10.1037/0021-9010.78.4.538.
Miočević, M., Gonzalez, O., Valente, M.J., MacKinnon, D.P., 2018. A tutorial in Bayesian potential outcomes mediation analysis. Struct Equ Modeling 25, 121–136. doi:10.1080/10705511.2017.1342541.
Muthén, B., Muthén, L., Asparouhov, T., 2016. Regression and Mediation Analysis Using Mplus. Muthén & Muthén.
Nuijten, M.B., Wetzels, R., Matzke, D., Dolan, C.V., Wagenmakers, E.J., 2015. A default Bayesian hypothesis test for mediation. Behavior Research Methods 47, 85–97. doi:10.3758/s13428-014-0470-2.
Pearl, J., 2001. Direct and indirect effects. In: Proceedings of the seventeenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., pp. 411–420. URL http://dl.acm.org/citation.cfm?id=2074022.2074073
Preacher, K.J., Hayes, A.F., 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behavior Research Methods 36, 717–731. doi:10.3758/bf03206553.
Preacher, K.J., Rucker, D.D., Hayes, A.F., 2007. Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. Multivariate Behavioral Research 42, 185–227. doi:10.1080/00273170701341316.
Robins, J., 2003. Semantics of causal dag models and the identification of direct and indirect effects. In: Highly structured stochastic systems. Oxford University Press, pp. 70–81.
Sobel, M.E., 1982. Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology 13, 290–312. doi:10.2307/270723.
VanderWeele, T., 2015. Explanation in Causal Inference: Methods for Mediation and Interaction. Oxford University Press.
Wang, L.P., Preacher, K.J., 2015. Moderated mediation analysis using Bayesian methods. Structural Equation Modeling A Multidisciplinary Journal 22, 249–263. doi:10.1080/10705511.2014.935256.
Yuan, Y., MacKinnon, D.P., 2009. Bayesian mediation analysis. Psychological Methods 14, 301–322. doi:10.1037/a0016972.
Zellner, A., 1971. An Introduction to Bayesian Inference in Econometrics. Wiley.
Zellner, A., 1984. Basic Issues in Econometrics. University of Chicago Press.
Zellner, A., Siow, A., 1980. Posterior odds ratios for selected regression hypotheses. In: Bernardo, J.M., DeGroot, M.H., Lindley, D.V., Smith, A.F.M. (Eds.), Proceedings of the first international meeting held in valencia, vol. 31. Valencia, Spain doi:10.1007/bf02888369.
Zhou, Q., Guan, Y., 2018. On the null distribution of Bayes factor in linear regression. Journal of the American Statistical Association 113, 1362–1371. doi:10.1080/01621459.2017.1328361.