Appendix A
Using PROCESS

This appendix describes how to install and execute PROCESS, how to set up a PROCESS command, and it documents its many features, some of which are not described elsewhere in this book. As PROCESS is modified and features are added, supplementary documentation will be released at *www.afhayes.com*. Check this web page regularly for updates. Also available at this page is a complete set of model templates identifying each model that PROCESS can estimate.

This documentation focuses on the SPSS version of PROCESS. All features and functions described below are available in the SAS version as well and work as described here, with minor modifications to the syntax. At the end of this documentation (see page 438), a special section devoted to SAS describes some of the differences in syntax structure for the SAS version compared to what is described below.

Overview

PROCESS is a computational tool for path analysis–based moderation and mediation analysis as well as their integration in the form of a conditional process model. In addition to estimating unstandardized model coefficients, standard errors, *t* and *p*-values, and confidence intervals using either OLS regression (for continuous outcomes) or maximum likelihood logistic regression (for dichotomous outcomes), PROCESS generates direct and indirect effects in mediation models, conditional effects (i.e., "simple slopes") in moderation models, and conditional indirect effects in conditional process models with a single or multiple mediators. PROCESS offers various methods for probing two- and three-way interactions and can construct percentile bootstrap, bias-corrected bootstrap, and Monte Carlo confidence intervals for indirect effects. In mediation models, multiple mediator variables can be specified to operate in parallel or in serial. Heteroscedasticity-consistent standard errors are available for inference about model coefficients, in the Sobel test for indirect effects, and when probing interactions in moderation analysis. Various measures of effect size for indirect effects are generated in simple and parallel multiple mediation models, along with bootstrap confidence intervals for effect size inference. An option is available for partialing out contextual-level variation when individual data are nested under a higher-level organizational structure. Individual paths in moderated mediation models can be estimated as moderated by one or two variables either additively or multiplicatively. Some models estimated by PROCESS allow up to four moderators simultaneously.

Preparing for Use

PROCESS can be used as either a command-driven macro or installed as a custom dialog for setting up the model using SPSS's point-and-click user interface. When executed as a macro, the PROCESS.sps file (available from *www.afhayes.com*) should first be opened as a syntax file. Once it has been opened, execute the entire file exactly as is. ***Do not modify the code at all***. Once the PROCESS.sps program has been executed, it can be closed and the PROCESS command is available for use in any SPSS program. Running PROCESS.sps activates the macro, and it will remain active so long as SPSS remains open. The PROCESS file must be loaded and reexecuted each time SPSS is opened in order to use the features of the PROCESS command. See the "Examples" section starting on page 422 for how to set up a PROCESS command in a syntax window. Please also read "Model Designation and Estimation" (page 426) and the "Notes" section (page 439) for important details pertinent to execution, including calling PROCESS with the SPSS `INSERT` command.

To install PROCESS as a custom dialog into the SPSS menus, execute PROCESS.spd (available from *www.afhayes.com*) by double-clicking it on the desktop or opening and installing it from within SPSS under the Utilities menu. Administrative access to the machine on which PROCESS is being installed is required when using a Windows operating system, and you must execute SPSS as an administrator. Once successfully installed, PROCESS will appear as a new menu item in SPSS nested under Analyze → Regression. If you do not have administrative access, contact your local information technology specialist for assistance in setting up administrative access to the machine on which you wish to install PROCESS.

Although the dialog box offers a "Paste" button, its use is not recommended. Users interested in embedding PROCESS commands in their own syntax should use the syntax-driven macro (PROCESS.sps) rather than the custom dialog. Execution of PROCESS.sps as described earlier is not necessary when the model is set up using a dialog box. Some options available in the macro cannot be accessed through the dialog box.

Syntax Structure

The first line of syntax below is required for all PROCESS commands. The remaining commands in brackets are optional or model dependent. Brackets, parentheses, and asterisks should not be included in the PROCESS command. "**" Denotes the default argument when the option is omitted.


```
process vars=varlist/y=yvar/x=xvar/m=mvlist/model=num
        [/w=wvar] [/z=zvar]
        [/v=vvar] [/q=qvar]
        [/wmodval=wval]
        [/zmodval=zval]
        [/vmodval=vval]
        [/qmodval=qval]
        [/mmodval=mval]
        [/xmodval=xval]
        [/cluster=clvar]
        [/contrast=(0**)(1)]
        [/boot=z(1000**)]
        [/mc=g(0**)]
        [/conf=ci(95**)]
        [/effsize=(0**)(1)]
        [/normal=(0**)(1)]
        [/jn=(0**)(1)]
        [/coeffci=(0)(1**)]
        [/varorder=vord(2**)]
        [/hc3=(0**)(1)]
        [/covmy=cov(0**)]
        [/total=(0**)(1)]
        [/center=(0**)(1)]
        [/quantile=(0**)(1)]
        [/detail=(0)(1**)]
        [/plot=(0**)(1)]
        [/seed=sd]
        [/percent=(0**)(1)]
        [/iterate=it(10000**)]
        [/converge=cvg(.00000001)]
        [/save=(0**)(1)].
```

Examples

(1) Simple Moderation

```
process vars=newlaws alcohol concerns use age/y=newlaws/x=alcohol
    /m=concerns/model=1/quantile=1/center=1/plot=1/jn=1.
```

- Estimates a simple moderation model with the effect of alcohol on newlaws moderated by concerns.

- use and age are included in the model as covariates.
- alcohol and concerns are mean centered prior to analysis.
- Generates the conditional effects of alcohol on newlaws at values of concerns equal to the 10th , 25th, 50th, 75th, and 90th percentiles of the distribution in the sample.
- Produces a table of estimated values of newlaws for various values of alcohol and concerns.
- Implements the Johnson–Neyman technique to identify the values on the continuum of concerns at which point the effect of alcohol on newlaws transitions between statistically significant and nonsignificant at the .05 level.

(2) Moderated Moderation


```
process vars=mathprob gender explms treat/y=mathrpob/x=treat/m=explms
    /w=gender/model=3/mmodval=4.
```

- Estimates a moderated moderation model predicting mathprob from treat while including a three-way interaction between treat, explms, and gender in the model along with all required two-way interactions.
- Generates the conditional effect of treat on mathprob for both males and females when explms = 4.

(3) Simple Mediation

```
process vars=donate winner votes/y=votes/x=donate/m=winner/model=4
    /total=1/effsize=1/boot=10000.
```

- Estimates the total and direct effect of `donate` on `votes`, as well as the indirect effect of `donate` on `votes` through `winner`.
- Generates a bias-corrected 95% bootstrap confidence interval for the indirect effect using 10,000 bootstrap samples.
- Produces point estimates and bias-corrected 95% bootstrap confidence interval estimates of various indices of effect size for the indirect effect.

(4) Parallel Multiple Mediation

```
process vars=know educ attn elab sex age/y=know/x=educ/m=attn elab/model=4/
    /contrast=1/normal=1/conf=90/save=1.
```

- Estimates the direct effect of `educ` on `know`, as well as the total and specific indirect effects of `educ` on `know` through `attn` and `elab`, with `attn` and `elab` functioning as parallel mediators.
- `sex` and `age` are included in the model as covariates.
- Produces the Sobel test for the specific indirect effects.
- Generates 90% bias-corrected bootstrap confidence intervals for the indirect effects using 1,000 bootstrap samples.
- Calculates the difference between the two specific indirect effects and produces a bias-corrected bootstrap confidence interval for the difference.
- Creates a new data window containing the 1,000 bootstrap estimates of each of the regression coefficients.

(5) Serial Multiple Mediation

```
process vars=commit close desire happy nbhrhood/y=happy/x=commit
    /m=close desire/model=6/hc3=1/effsize=1/boot=10000/cluster=nbhrhood.
```

- Estimates the direct effect of `commit` on `happy`, as well as the total and all possible specific indirect effects of `commit` on `happy` through `close` and `desire`.

- `close` and `desire` function as mediators in serial, with `close` affecting `desire`.
- Standard errors for model coefficients are based on the HC3 heteroscedasticity-consistent standard error estimator.
- Generates 95% bias-corrected bootstrap confidence intervals for the indirect effects using 10,000 bootstrap samples.
- Produces point estimates and bias-corrected 95% bootstrap confidence intervals for various indices of effect size for the indirect effects.
- With cases nested within neighborhoods (coded with a variable named `nbhrhood`), partials out neighborhood-level effects from all estimates.

(6) Conditional Process Model Example 1


```
process vars=frame euskept peffic risk turnout/y=turnout/x=frame/m=risk
    /w=euskept/z=peffic/model=68/boot=20000/wmodval=2/center=1.
```

- Estimates the direct effect of `frame` on `turnout`, as well as the conditional indirect effects of `frame` on `turnout` through `risk`. The effect of `frame` on `risk` is modeled as multiplicatively moderated by both `peffic` and `euskept`, and the effect of `risk` on `turnout` is modeled as moderated by `euskept`.
- `euskept`, `peffic`, and `frame` are mean centered prior to analysis.
- Calculates the conditional indirect effects of `frame` on `turnout` through `risk` among cases 2 units above the sample mean on `euskept` and with values of `peffic` at the sample means, as well as with `peffic` one standard deviation above and below the sample mean.

- Generates bias-corrected 95% bootstrap confidence intervals for the conditional indirect effects using 20,000 bootstrap samples.

(7) Conditional Process Model Example 2

```
process vars=calling livecall carcomm workmean jobsat/y=jobsat/m=carcomm
    workmean/x=calling/w=livecall/model=7/boot=5000/seed=34421.
```

- Estimates the direct effect of `calling` on `jobsat`, as well as the conditional indirect effects of `calling` on `jobsat` through both `carcomm` and `workmean` operating in parallel. The effects of calling on both `carcomm` and `workmean` are modeled as moderated by `livecall`.

- Produces the conditional indirect effects of `calling` when `livecall` is equal to the sample mean as well as plus and minus one standard deviation from the mean.

- Generates bias-corrected 95% bootstrap confidence intervals for the conditional indirect effects using 5,000 bootstrap samples.

- Seeds the random number generator for bootstrap sampling with the value 34421.

(8) Conditional Process Model Example 3

```
process vars=protest sexism respappr anger liking age sex/y=liking
    /x=protest/m=respappr anger/w=sexism/model=8/boot=5000/quantile=1
    /percent=1.
```


- Estimates the effect of `protest` on `liking` directly as well as indirectly through `respappr` and `anger`, with both direct and indirect effects moderated by `sexism`. The effect of `protest` on `respappr` as well as the effect `protest` on `anger` is modeled as moderated by `sexism`.

- `age` and `sex` are included in the model as covariates.

- Generates 95% percentile-based bootstrap confidence intervals based on 5,000 bootstrap samples for the conditional indirect effect of `protest` at the 10th, 25th, 50th, 75th, and 90th percentile values of `sexism`.

- Produces the indirect effects of the product of `protest` and `sexism` on `liking` through `respappr` as well as through `anger`, along with a percentile-based 95% bootstrap confidence interval.

Model Designation and Estimation

PROCESS can estimate many different models, and which model is estimated is determined by the ***num*** argument in the required **`model`** specification. The more popular and frequently used models that PROCESS can estimate are depicted conceptually and in the form of a path diagram beginning on page 442, along with their corresponding model number as recognized by PROCESS in the **`model`** specification. PROCESS can estimate over 70 models. Additional templates containing the conceptual and statistical diagrams corresponding to models PROCESS can estimate can be found at *www.afhayes.com*.

Each model has certain minimum requirements as to which variables must be designated and provided in the PROCESS command. Any variable in the dataset that appears in the model must be listed in the ***varlist*** argument of the PROCESS command (e.g., **`vars=`*****xvar yvar mvlist wvar***). Furthermore, all models require

- a single outcome variable ***yvar*** listed in the **`y`** specification (i.e., **`y=`*****yvar***), where ***yvar*** is the name of the variable in your data functioning as Y in the model
- a single antecedent causal agent ***xvar*** listed in the **`x`** specification (i.e., **`x=`*****xvar***), where ***xvar*** is the name of the variable in your data functioning as X in the model

- either a single moderator (models 1, 2, and 3) or at least one mediator (models 4 and higher) specified in the ***mvlist*** in the **`m`** specification (i.e, **`m=`*****mvlist***), where ***mvlist*** is the name of the variable or variables in the data functioning as moderator (models 1, 2, and 3) or mediator(s) (models 4 and higher).

Other than **`x`**, **`y`**, **`m`**, **`model`**, and **`vars`**, the remaining required inputs to PROCESS will be model dependent. In general, any variable that is a part of the conceptual model in the model template must be provided as an input to PROCESS, and any variable that is not a part of the conceptual model must be left out unless such variables are to be treated as covariates by inclusion in ***varlist***. For instance, observe in the model templates section (see page 442) that model 21 has, in addition to X, M, and Y, two moderators W and V. Thus, PROCESS must also be told which two variables in the dataset correspond to W and V in the diagram. This would be done with the use of the **w** and **v** specifications (e.g., **w=*wvar*** and **v=*vvar***), where ***wvar*** and ***vvar*** are the names of the variables in the data file corresponding to W and V.

The **y**, **x**, **w**, **z**, **v**, and **q** specfications each allow only one variable, and a variable can be listed in one and only one of the ***yvar***, ***xvar***, ***mvlist***, ***wvar***, ***zvar***, ***vvar***, and ***qvar*** arguments. For instance, a variable cannot be listed as both **w** in ***wvar*** and **m** in ***mvlist***. However, both would have to appear in ***varlist***.

In the SPSS version of PROCESS, the variable names listed in the ***varlist***, ***yvar***, ***xvar***, ***mvlist***, ***wvar***, ***zvar***, ***vvar***, and ***qvar*** arguments must match the case (i.e., uppercase, lowercase, or combinations thereof) of the variables in the dataset. So `ATTITUDE`, `Attitude`, and `AttiTuDe` are different variables according to PROCESS. Thus, **y=ATTITUDE** will produce an error even if `Attitude` exists in your data file. In addition, the potential for errors at execution is increased when variable names are more than eight characters in length. Thus, the user is advised to *reduce all long variable names in the dataset and that are to be used in a PROCESS command down to eight characters at maximum.*

Although PROCESS has a number of error-trapping routines built in, it will not catch all errors produced by improper formatting of a PROCESS command, improper listing of variables and variable names, and so forth. Any errors it has trapped will be displayed in an errors section of the PROCESS output. Errors it has not successfully trapped will appear as a long list of SPSS execution errors that will be largely unintelligible.

Multiple Mediators

All mediation models (models 4 and higher) can have up to 10 mediators operating in parallel, with the exception of model 6, which is restricted to between two and four and models the mediators as operating in serial. Mediators operating in parallel are all modeled as affected by ***xvar*** and, in turn, affect ***yvar***, but they are not modeled to transmit their effects to any other mediators in the model (see section 5.1). Mediators operating in serial are linked in a causal chain, with the first mediator affecting the second, the second the third, and so forth (see section 5.4). The order of the mediators in ***mvlist*** is not conquential to the estimation of the model except in model 6. In model 6, the first variable in ***mvlist*** is assumed to be causally prior to the second variable in ***mvlist***, which is causally prior to the third, and so forth.

Pairwise comparisons between specific indirect effects can be requested for models 4, 5, and 6 by setting the argument in the **contrast** option to 1 (i.e., **contrast=1**). These comparisons will appear in the output in the indirect effects section with labels (C1), (C2), and so forth. A table that maps the label to the specific indirect effects being compared is provided at the bottom of the output. Bootstrap or Monte Carlo confidence intervals are provided for inference for these pairwise comparisons when the contrast option is used in conjunction with bootstrapping or the Monte Carlo option. See section 5.3 for a discussion of contrasts between indirect effects.

Models 7 and higher include the moderation of an effect either to or from a mediator. The model templates toward the end of this appendix and at *www.afhayes.com* illustrate which path is moderated for a given model. In models with multiple mediators, the moderation applies to all corresponding paths for each mediator. For example, if model 7 is specified with two mediators, med1 and med2, then both the effect from ***xvar*** to med1 and the effect from ***xvar*** to med2 will be estimated as moderated by ***wvar***. There is no way of restricting the estimation of the moderation to only one of the paths using the PROCESS procedure. Doing so requires the use of a structural equation modeling program.

Covariates

Any variable that appears in ***varlist*** but that does not appear anywhere else in the PROCESS command will be treated as a covariate. By default, covariates are included as predictor variables in the models of all mediators (i.e., all models of the variables in ***mvlist***) as well as in the model of ***yvar***. This can be changed through the use of the **covmy** option. Setting the ***cov*** argument to 1 (i.e., **covmy=1**) includes the covariates in the model of all mediators in ***mvlist*** but not outcome variable ***yvar***. Using an argument of 2 (i.e., **covmy=2**) specifies estimation of a model that includes all covariates in the model of outcome ***yvar*** but not the mediator variables in ***mvlist***. It is not possible to specify some covariates only in the model of ***yvar*** and other covariates only in the model(s) of ***mvlist***.

Probing Interactions and Generating Conditional Effects

In any model that involves a moderated effect, PROCESS will produce estimates of conditional effects (direct and/or indirect) at various values of the moderator based on the equations at the bottom of each of the model templates beginning on page 442. By default, when a moderator is dichotomous, conditional effects at the two values of the moderator are generated. But when a moderator is quantitative, conditional effects are estimated by default at the sample mean of the moderator, as well as plus and minus one standard deviation from the moderator mean.

Three alternatives for probing interactions are available in PROCESS. For quantitative moderators, the **`quantile`** option generates conditional effects at the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of the moderator. This option is available by setting the argument in the **`quantile`** option to 1 (i.e., **`quantile=1`**). Unlike when the mean and ± one standard deviation from the mean is used, these quantile values are guaranteed to be within the range of the observed data. For a discrete quantitative moderator (i.e., a quantitative moderator with relatively few observed values), some of the quantile values of the moderator may be identical. For example, the 10th and 25th percentile of the moderator may be the same value. This will produce some redundancy in the output.

The second alternative is to request the conditional effect of interest at a specific value of the moderator or moderators. This is accomplished through the use of the **`wmodval`**, **`zmodval`**, **`vmodval`**, **`mmodval`**, **`qmodval`**, and **`xmodval`** options, setting the corresponding argument (***wval***, ***zval***, ***vval***, ***mval***, ***wval***, and/or ***xval***) to the value of the moderator at which you'd like the estimate of the conditional effect. For example, model 22 includes two moderators, W and V. To generate an estimate of the conditional indirect effect of ***xvar*** on ***yvar*** through the variables in ***mvlist*** when $W = 1$ and $V = 2$, append **`wmodval=1`** and **`vmodval=2`** to the PROCESS command. This will also generate an estimate of the conditional direct effect of X when $W = 1$. When used in conjunction with the **`center`** option, values of the moderator provided should be based on the mean-centered metric rather than the original metric of measurement. This option for probing an interaction is not available in the custom dialog version of PROCESS.

The third alternative is the Johnson–Neyman technique, requested by setting the argument in the **`jn`** option to 1 (i.e., **`jn=1`**). This is available for models 1 and 3. For model 1, this approach identifies the value(s) on the moderator variable (***mvlist***) continuum at which point (or points) the effect of ***xvar*** on ***yvar*** transitions between statistically significant and not, using the α-level of significance as the criterion. By default, $\alpha = 0.05$. This can be changed using the **`conf`** option, setting the desired confidence to $100(1 - \alpha)$. For example, for $\alpha = 0.01$, specify **`conf=99`**. For model 3, PROCESS computes the point or points along the continuum of moderator ***wvar*** at which the two-way interaction between ***xvar*** and ***mvlist*** transitions between statistically significant and not. In addition to identifying these points, PROCESS produces a table to aid in the identification of the regions of significance. See section 7.4 for a discussion of the Johnson–Neyman technique.

When a model includes more than one moderator, a table of conditional effects is generated for all combinations of the moderators based on the options or defaults used and described earlier. For example, in model 21, if **wmodval=1** is specified but the **vmodval** option is not used and V is a quantitative moderator, PROCESS will generate a table of conditional indirect effects when $W = 1$ and V is equal to the mean, as well as plus and minus one standard deviation from the mean.

Statistical Inference for Indirect Effects

PROCESS offers a normal theory approach (i.e., the Sobel test), bootstrap confidence intervals, and Monte Carlo confidence intervals for inference about the indirect effect in models with a mediation component.

By setting the argument in the **normal** option to 1 (i.e., **normal=1**), PROCESS generates the normal theory-based Sobel test for the indirect effects in simple and parallel multiple mediator models (models 4, 5). A p-value is derived using the standard normal distribution. This test is not available for the total indirect effect or conditional indirect effects in mediation models with moderated paths. The standard error estimator for an indirect effect used in the normal theory test is determined by the ***vord*** argument in the **varorder** option. When set to 1 (i.e., **varorder=1**), the first-order standard error estimator is used, whereas when is set to 2 (i.e., **varorder=2**, the default), the second-order standard error is used.

Bootstrap confidence intervals are the default and preferred over the normal theory-based Sobel test for inference about indirect effects because of the unrealistic assumption the Sobel tests makes about the shape of the sampling distribution of the indirect effect. By default, PROCESS generates 95% bias-corrected bootstrap confidence intervals for all indirect effects in any model that involves a mediation component (models 4 and higher). Percentile-based bootstrap confidence intervals will be generated rather than bias-corrected confidence intervals by setting the argument in the **percent** option to 1 (i.e., **percent=1**). The number of bootstrap samples can be set with the z argument in the **boot** option to any desired number (e.g., **boot=5000**; the default number of bootstrap samples is 1,000, but at least 5,000 is recommended for scientific publications). Set z to 0 to turn off bootstrapping. The level of confidence for confidence intervals can be changed by setting ***ci*** to the desired number anywhere between 50 and 99.9999 (90, 99, etc.) in the **conf** option (e.g., **conf=99**). If the number of bootstrap samples requested is too small given the requested level of confidence desired, PROCESS will automatically increase the number of bootstrap samples as required. A note will be produced at the bottom of the output to this effect when it occurs.

Monte Carlo confidence intervals can be requested instead of bootstrap confidence intervals for simple and multiple mediator models (models 4 and 5) through the use of the **mc** option, setting the ***g*** argument to the number of samples desired. For example, **mc=5000** requests Monte Carlo confidence intervals for the indirect effect based on 5,000 samples. The Monte Carlo option takes precedence over the bootstrapping option, so if **mc** is used in conjunction with **boot**, Monte Carlo confidence intervals will result. The **mc** option is ignored by PROCESS when estimating models 7 and above. In that case, bootstrap confidence intervals are generated and the number of bootstrap samples will default to 1,000 unless a larger number is requested.

In addition to the point estimate of the indirect effect and the endpoints of a confidence interval, PROCESS will also generate a bootstrap or Monte Carlo estimate of the standard error of the indirect effect. The standard error of the indirect effect is defined as the standard deviation of the *z* bootstrap or ***g*** Monte Carlo estimates.

Because bootstrapping and Monte Carlo methods are based on random sampling from the data (for bootstrapping) or from theoretical distributions (for Monte Carlo confidence intervals), confidence intervals and standard errors will differ slightly each time PROCESS is run as a result of the random sampling process. The more bootstrap or Monte Carlo samples that are requested, the less this variation between runs. It is possible to replicate a set of random samples by seeding the random number generator, setting ***sd*** in the **seed** option to any integer between 1 and 2,000,000,000 prior to running PROCESS (e.g., **seed=23543**). By default, the random number generator is seeded with a random number.

Saving Bootstrap Estimates

When estimating models 4 and above in conjunction with bootstrap confidence intervals for indirect effects, the bootstrap estimates of all regression coefficients can be saved for examination or additional analysis by setting the argument in the **save** option to 1 (i.e., **save=1**). This will produce a new data file in the SPSS session with as many rows as bootstrap samples requested, and as many columns as parameter estimates in the model being estimated. The columns of the data file containing the bootstrap samples will be in the order the parameter estimates first appear in PROCESS output from top to bottom. Parameter estimates for the total effect of *X* will not be included in this file when the **save** option is used in conjunction with the **total** option in models 4, 5, and 6.

The resulting data file must be saved in order to store it permanently, as subsequent runs of the PROCESS command with the **save=1** option will overwrite the prior file if it is not first saved permanently.

Confidence Intervals for Model Coefficients

By default, PROCESS generates ordinary least squares (for continuous ***mvlist*** or ***yvar***) or maximum-likelihood-based (for dichotomous ***yvar***) confidence intervals for all regression coefficients, as well as direct, total, and conditional effects. To suppress their printing, set the argument in the **coeffci** option to 0 (i.e., **coeffci=0**). The confidence level is set using the **conf** option (e.g., **conf=90** for 90% confidence intervals. The default is 95%).

Effect Size Indices for Indirect Effects

When estimating an unmoderated mediation model (models 4, 5, or 6) setting the argument in the **effsize** option to 1 (i.e., **effsize=1**) generates various estimates of the size of the indirect effect. Effect sizes available include the partially and completely standardized indirect effect, κ^2 (Preacher & Kelley, 2011), R^2 (Fairchild et al., 2009), the ratio of the indirect to total effect, and the ratio of the indirect to the direct effect. When used in conjunction with the bootstrapping option, bootstrap confidence intervals for these effect size measures are generated. The R^2 and κ^2 measures are available only in models with a single mediator, no covariates, and no clustering. Effect size measures are not generated for models with a dichotomous ***yvar*** or models with covariates.

The Total Effect in Unmoderated Mediation Models

In mediation models with no moderated effects (models 4, 5, and 6), PROCESS generates the direct and indirect effects of ***xvar*** on ***yvar*** by default, along with corresponding linear models used to estimate these effects. In models 4, 5, and 6 with continuous mediator(s) and outcome, the total effect of ***xvar*** on ***yvar*** is the sum of the direct and indirect effects of ***xvar***. By setting the argument in the **total** option to 1 (i.e., **total=1**), PROCESS will produce the total effect of ***xvar*** on ***yvar*** with a test of significance, along with the corresponding model of ***yvar*** without the proposed mediators in the model. Use of the **total** option with a dichotomous outcome will generate similar output, but the total effect will not be equal to the sum of the direct and indirect effects due to the arbitrary scaling of the error in estimation in the logistic regression model. Thus, the difference between the total and direct effects is not equivalent to the indirect effect.

Missing Data

PROCESS assumes complete data and will exclude cases from the analysis that are missing on any of the variables in ***varlist***. Any missing data substitution or imputation desired by the user should be conducted prior to the execution of PROCESS.

Mean Centering in Models with Interactions

In models that include parameters for estimating interaction effects (i.e., all models PROCESS estimates except 4 and 6), moderation is assumed to be linear, with products of variables serving to represent the moderation. The user has the option of requesting PROCESS to mean center all variables used in the construction of products of predictors prior to model estimation by setting the argument in the ***center*** option to 1 (i.e., `center=1`). All output for conditional effects will be based on moderator values using the mean-centered metric (e.g., the conditional effect of ***xvar*** on ***yvar*** at values of ***wvar*** will be based on values of ***wvar*** after mean centering).

By default, variables used to form products are not mean centered. When mean centering is requested, arguments of options used for estimating conditional effects at specific values of the moderator(s) should be values based on a mean-centered metric. For example, the SPSS command


```
PROCESS vars=smoking surgery anxiety addict/y=smoking/x=surgery/
   m=anxiety/w=addict/model=7/wmodval=1.5.
```

will produce the conditional indirect effect of `surgery` on `smoking` through `anxiety` when `addict` = 1.5, whereas the SPSS command

```
PROCESS vars=smoking surgery anxiety addict/y=smoking/x=surgery/
   m=anxiety/w=addict/model=7/center=1/wmodval=1.5.
```

produces the conditional indirect effect of `surgery` on `smoking` through `anxiety` when `addict` is 1.5 measurement units above the sample mean of `addict`.

Visualizing Interactions in Moderation-Only Models

To help visualize and interpret the nature of the moderation of ***xvar***'s effect on ***yvar*** in models 1, 2, and 3, the **`plot`** option generates a table of predicted values of ***yvar*** from the model using various values of ***xvar*** and the moderator or moderators. This table is generated by setting the argument in the **`plot`** option to 1 (i.e., **`plot=1`**). Any covariates in the model are set to their sample mean when deriving the predicted values in the table generated.

In the table, the estimated value of ***yvar*** is listed as "yhat." For OLS regression, this is simply the estimate of ***yvar*** from the regression model for various values of ***xvar*** and the moderator(s), with covariates (if any) set to their sample means. For logistic regression, "yhat" is the estimated log odds of the event coded with ***yvar*** (with the higher code arbitrarily treated as the event modeled). The **`plot`** option for logistic regression models will also produce a column labeled "prob," which is the estimated log odds converted to a probability using the formula

$$\text{prob} = \frac{e^{yhat}}{1 + e^{yhat}}$$

Nonindependence and Spuriousness Due to Cluster Effects

Subsets of cases in an analysis sometimes are nested under a common organizational unit or "cluster," such as patients in hospitals, kids in schools, or households within neighborhoods. When cases are derived from several organizational units, some of the relationships observed may be attributable to unmodeled effects of organizational units or clusters. When there are many cases in many organizational units, multilevel modeling is the best strategy for dealing with the nonindependence such clustering can produce. An alternative approach when the number of cluster units is small and one is willing to assume fixed effects of the variables in the model is to remove any effect due to organizational unit or cluster by using dummy variables to partial out effects due to cluster from estimates of the coefficients and standard errors in the model.

PROCESS has an option that implements the latter procedure, sometimes called the "fixed effects approach to clustering" (see, e.g., Cohen et al., 2003, pp. 539–544). By specifying the variable that codes organizational unit as the ***clvar*** argument in the **`cluster`** option, PROCESS automatically produces $k - 1$ dummy variables coding which of the k clusters a case is nested under. These $k - 1$ dummy variables are then included as additional predictors in all linear models generated as part of the analysis. PROCESS allows a maximum of 20 cluster units when this option is requested. This option is not available in the custom dialog interface.

For example, the SPSS command

```
PROCESS vars=smoking surgery anxiety hsptl/y=smoking/x=surgery/m=anxiety
   /model=4/cluster=hsptl.
```

estimates the direct effect of `surgery` on `smoking`, as well as the indirect effect of `surgery` through `anxiety` while partialing out differences between cases in `anxiety`, `surgery`, and `smoking` due to which hospital a person attended, with hospital coded in a variable in the data file named `hsptl`. Notice that the clustering variable must also be provided in ***varlist***.

The model coefficients for the dummy variables are not displayed in the output, nor are they added to the data file. However, model summary information (e.g., R^2) will include the effects of cluster as well as the other variables in the model.

An important limitation of the cluster option is that no variables in the model can be measured at the level of the cluster (i.e., in multilevel terms, none of the variables can be measured at level-2). That is, there must be variation in each of the variables in the model (including the covariates) within cluster. When this condition is not satisfied, a matrix inversion error will result and PROCESS will terminate.

Heteroscedasticity-Consistent Standard Errors

By default, PROCESS uses an estimator for the standard errors of the regression coefficients that assumes homoscedasticity of the errors in estimation of the outcome variable. PROCESS can also generate standard errors using the HC3 estimator, described in Long and Ervin (2000) and Hayes and Cai (2007). This heteroscedasticity-consistent standard error estimator is requested by setting the argument in the **`hc3`** option to 1 (i.e., **`hc3=1`**). Any computation that uses the standard error of a regression coefficient will automatically employ the HC3 estimator when this option is requested, including the Sobel test, the Johnson–Neyman method, tests of conditional effects in moderation analysis, and the test of the significance of R^2 for models of ***yvar***, as well as ***mvlist*** in mediation analysis.

When heteroscedasticity-consistent standard errors are requested for models 1, 2, and 3, neither the change in R^2 due to interaction(s) nor a test of significance for the change is provided in the output.

Abbreviated Output

By default, PROCESS produces all regression coefficients, standard errors, *t* and *p*-values, and model summary information for the models of ***yvar*** and, in all but models 1, 2, and 3, the proposed mediators in ***mvlist***. In complex models this output can be quite lengthy. If desired, the user can suppress the printing of model information with use of the **`detail`** option, setting the argument to 0 (i.e., **`detail=0`**).

Binary Outcome Variables

PROCESS can estimate models with either a continuous or a binary ***yvar*** and will automatically detect whether or not ***yvar*** is binary and estimate accordingly. If PROCESS detects only two distinct values on the outcome variable, the direct and indirect effects, as well as the path(s) from the proposed mediator(s) to the outcome, are estimated using logistic regression; otherwise OLS is used. Confidence intervals for indirect effects are estimated in the usual way as the product of the path from the independent variable to the proposed mediator and the path from the proposed mediator to the outcome. Measures of effect size are not available for models with a binary outcome. Note that with binary outcomes the indirect and total effects of ***xvar*** are scaled differently, and so the total effect will not typically be equal to the sum of the direct and indirect effects. Thus, the difference between the total and the direct effect of ***xvar*** on ***yvar*** cannot be used as a substitute for the indirect effect, nor can one use this difference in a metric of effect size, such as the proportion of the effect that is mediated.

Logistic regression coefficients are estimated using a Newton–Raphson iteration algorithm. The number of iterations and convergence criterion can be set using the **`iterate`** and **`converge`** options in the command syntax, which default to 10,000 and 0.00000001, respectively.

Indirect Effects in Mediated Moderation

Mediated moderation is a term sometimes used to describe the phenomenon in which the moderation of an effect is carried to an outcome variable through a mediator. Of interest when testing a mediated moderation hypothesis is the estimation of the indirect effect of the product of ***xvar*** and the moderator(s). Along with estimates of the conditional indirect effect of ***xvar*** on ***yvar*** at values of the moderator, output from models 8 and 12 includes the indirect effect of the highest-order interaction in the model. In model 8, the highest-order interaction is the two-way interaction between ***xvar*** and ***wvar***, and in model 12, the highest-order interaction is the three-way interaction between ***xvar***, ***wvar***, and ***zvar***. A bootstrap confidence interval for the indirect effect of the highest-order product term can be used for inference as to whether "moderation is mediated." As described in section 12.3, an inference about the indirect effect of this highest-order interaction can also be interpreted as a test of whether the indirect effect of ***xvar*** on ***yvar*** through the variable(s) in ***mvlist*** is moderated by ***xvar***. See section 11.4 for a discussion of the meaningfulness and substantive interpretability of mediated moderation relative to moderated mediation.

Multiple Independent and Dependent Variables

In some cases the user might like to estimate a model that includes multiple independent variables, each linked to the same mediator or set of mediators. PROCESS does not allow more than one variable to be listed in ***xvar***. Nevertheless, as described in section 6.4, PROCESS can be used to estimate the coefficients in such a model. By default, covariates are mathematically treated exactly like independent variables in the estimation, with paths to all mediators and the outcome, so if the desired model has k independent variables, PROCESS can be run k times, each time listing one variable as the independent variable in ***xvar*** and treating remaining $k-1$ independent variables as covariates. Each run of PROCESS will generate the effects for the variable currently listed in ***xvar***. It is recommended that a common random number seed be used for each run of PROCESS when bootstrapping or using the Monte Carlo option, so that the same set of samples will be used when confidence intervals for indirect effects are estimated for different independent variables.

Even though only a single variable can be provided in ***yvar***, PROCESS can be used to estimate the direct and indirect effects of X on k dependent variables (Y) when the indirect effect passes through the same mediator or set of mediators and no causal path between the k Y variables is assumed. This is accomplished by running PROCESS k times, once for each dependent variable. By setting the random number seed to the same value for each run, the bootstrap or Monte Carlo samples will be the same at each run, and the results obtained will be as if all the paths were estimated in one model with k dependent variables. See section 6.4 for a discussion.

Mapping PROCESS Models onto MODMED and Edwards and Lambert (2007) Models

PROCESS can estimate conditional indirect effects for all models described in Preacher et al. (2007) and implemented in MODMED for SPSS. However, the model numbers are different. Edwards and Lambert (2007) describe various models that combine moderation and mediation using names rather than numbers, all of which can also be estimated by PROCESS. The table below maps model numbers in MODMED to corresponding model numbers in PROCESS and the model names used by Edwards and Lambert (2007).

MODMED	Edwards and Lambert (2007)	PROCESS
1	—	74
2	Direct effect and first-stage moderation	8
3	Second-stage moderation	14
4	—	22
5	Total effect moderation	59
—	First-stage moderation	7
—	First- and second-stage moderation	58
—	Direct effect moderation	5
—	Direct effect and second-stage moderation	15

Installation, Execution, and Syntax Modifications for SAS Users

The SAS version of PROCESS functions similarly to the SPSS version, and most of the instructions described in this appendix apply to the SAS version, with only the minor modifications described below. Like the SPSS version, the SAS version is a program file (PROCESS.sas), which when executed creates a new command that SAS understands called **`%process`**. Once PROCESS.sas is executed (without changing the file whatsoever), then the **`%process`** command is available for use and the program can be closed. Once you close SAS, you have to define the **`%process`** command by executing PROCESS.sas again. **PROCESS for SAS requires the PROC IML module.** To determine whether you have the PROC IML module installed, run the following commands in SAS:

```
proc iml;
print "PROC IML is installed";
quit;
```

When this code is executed, check the log for any errors, as well as your output window for the text "PROC IML is installed." Any errors in the log or a failure to see this text suggests that PROC IML is not installed on your version of SAS.

The syntax structure for PROCESS for SAS is almost identical to the SPSS version, with five important exceptions:

- The command name is **%process** rather than **process**.
- All parts of the command between **%process** and the ending semicolon (**;**) must be in parentheses.
- The data file being analyzed must be specified in the command as **data=*file*** where ***file*** is the name of a SAS data file.
- Options and specifications must be delimited with a comma (**,**) rather than a slash (**/**). For example, suppose the data corresponding to example 7 on page 425 were stored in a SAS work file named "jobs." The SAS version of the PROCESS command corresponding to example 7 would be

```
%process (data=jobs,vars=calling livecall carcomm workmean jobsat,y=jobsat,
   m=carcomm workmean,x=calling,w=livecall,model=7,boot=5000,seed=34421);
```


- The **save** option requires a file name for the resulting file of bootstrap estimates. For example, **save=mod14bt** tells SAS to save the bootstrap estimates of the regression coefficients to a temporary work file named `mod14bt`.

Notes

- In the SPSS version of PROCESS variable names are case sensitive and must match the case in the SPSS data file.
- "`xxx`" is a reserved variable name. Do not include any variable in your data set named "`xxx`" in the PROCESS procedure.

- PROCESS does not recognize variable names beyond the eighth character. Longer variable names will sometimes confuse PROCESS and are best avoided unless precautions are taken to rename variables that have the same first eight characters, thereby making them distinct to PROCESS.

- For all models, ***xvar***, ***yvar***, ***wvar***, ***zvar***, *qvar*, and ***var*** can be either dichotomous or quantitative with at least interval-level properties. The SPSS version of PROCESS ignores the properties of the data specified in the "Measure" column of the Variable View section of the data file.

- For models 4 and higher, variables in ***mvlist*** must be quantitative variables and are assumed to have at least interval-level measurement properties. In models 1, 2, and 3, the variable in ***mvlist*** can be either quantitative or dichotomous.

- A case will be deleted from the analysis if user- or system-missing on any of the variables in ***varlist***.

- Do not use STRING formatted variables in any of your models. Doing so will produce errors. All variables should be NUMERIC format.

- All covariates in the ***vars*** list are always assigned to the model of ***yvar***, ***mvlist***, or both (the default). It is not possible to assign some covariates to the model of ***yvar*** and others to the model(s) of ***mvlist***.

- The ***yvar***, ***xvar***, ***wvar***, ***zvar***, ***qvar***, ***vvar***, and ***clvar*** arguments are limited to one variable each. Up to 10 variables can be listed in ***mvlist***, except when estimating model 6, in which case ***mvlist*** is limited to four variables. Each variable should be specified in only one of these arguments.

- PROCESS does not offer any models that combine moderation with serial multiple mediation.

- All regression coefficients in the output are unstandardized. For continuous mediators or outcomes, all paths are estimated using OLS regression. For a dichotomous outcome, the path coefficients are maximum-likelihood-based logistic regression coefficients. PROCESS does not have an option for generating standardized regression coefficients. Users interested in standardized coefficients should standardize the variables in the model prior to execution of PROCESS. In mediation models based on standardized variables as input, a bootstrap confidence interval for an indirect effect produced by PROCESS should not be interpreted as a confidence interval for the standardized indirect effect. Standardization of dichotomous predictor variables or reporting of standardized effects for dichotomous predictors generally are not meaningful and should be avoided.

- In PROCESS, the bootstrapping routine is used only for the construction of bootstrap confidence intervals and bootstrap standard errors of indirect effects (conditional or unconditional) in models with a mediation component. Neither model coefficients, their standard errors, nor any other inferential tests are based on bootstrap methods.

- Bootstrapping takes time. The larger the sample, and the more complex the model, the longer the computations take. If you see the message "Running matrix" in the bottom right-hand corner of one of the SPSS windows, this means PROCESS is working on your data. Please be patient. Logistic regression with a large dataset combined with bootstrapping can take a very long time.

- The custom dialog version of PROCESS will construct a few variable names in your data file as it is working, and then delete these variables when it completes the computations. As a result, SPSS will ask you if you want to save your data when you quit SPSS. If you have made no modifications to the data file yourself and the data were already saved before executing, there is no need to save the data file again.

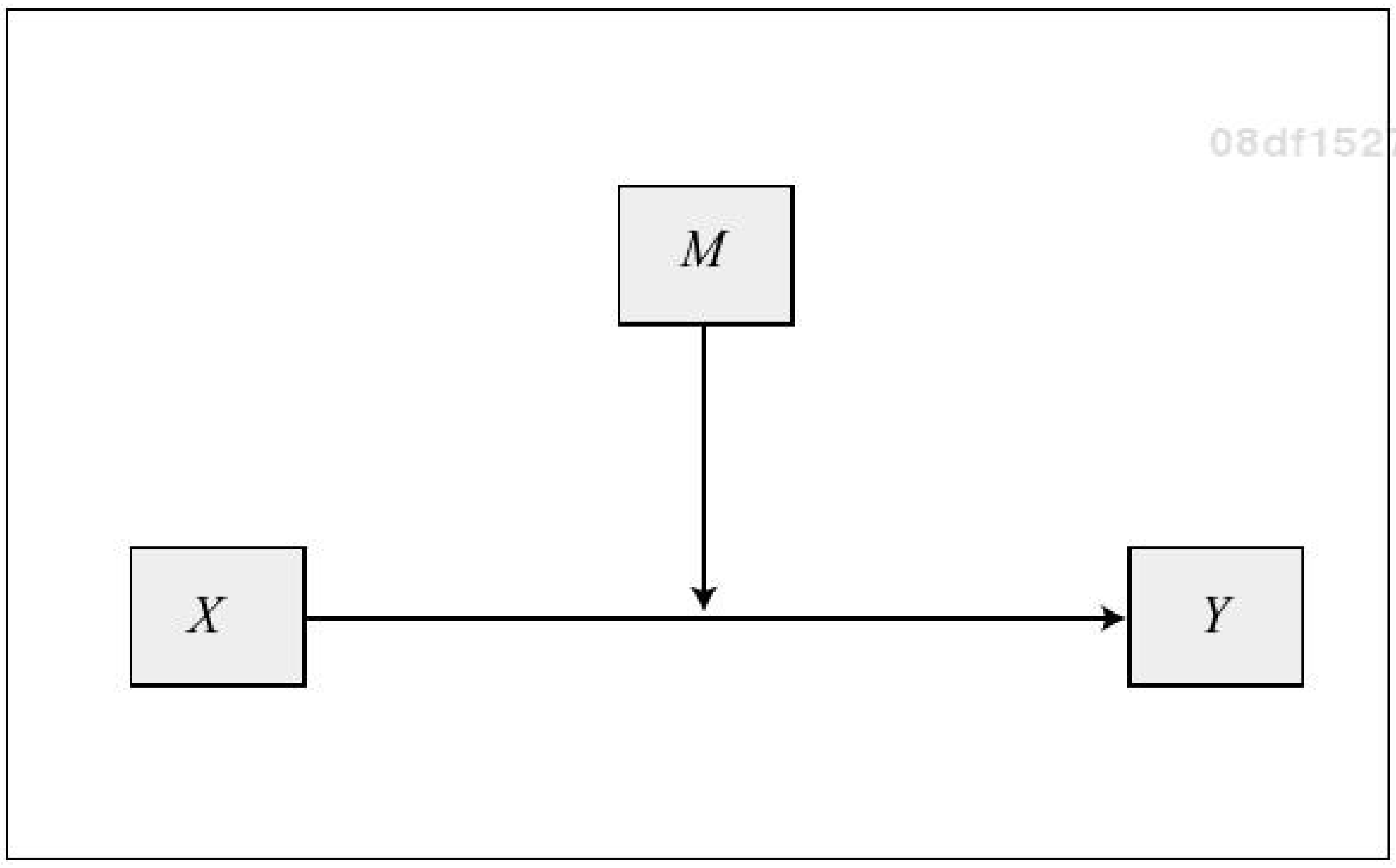
- The PROCESS procedure code cannot be imbedded in a syntax file with an `INCLUDE` command in SPSS, but it can be called with an `INSERT` command. This eliminates the need to manually load and run PROCESS.sps prior to execution of a set of commands which call the PROCESS macro. See the *Command Syntax Reference* available through the Help menu in SPSS for details on the use of the `INSERT` command.

PROCESS MODEL TEMPLATES

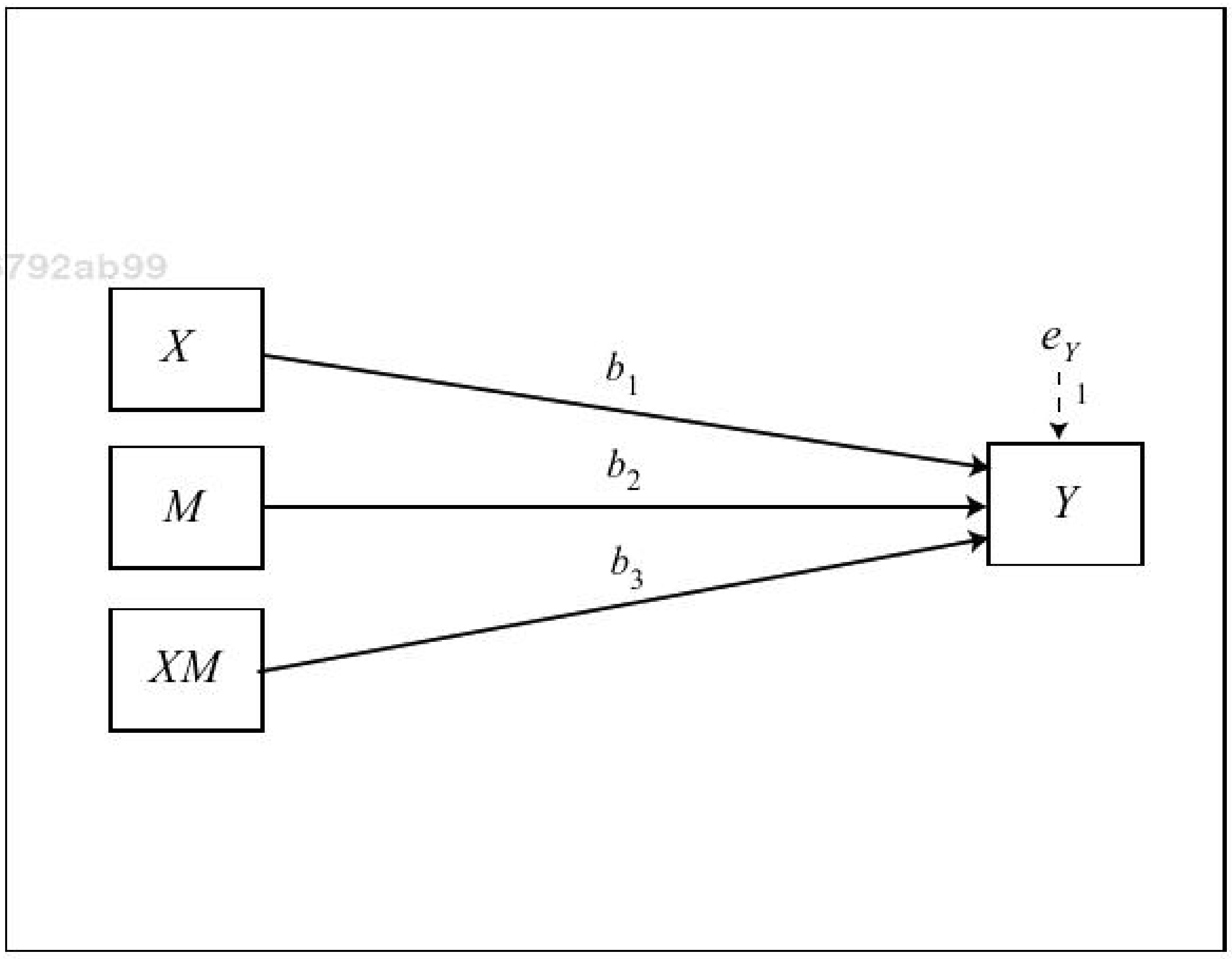
Additional templates are available at *www.afhayes.com.*

Model 1

Conceptual Diagram



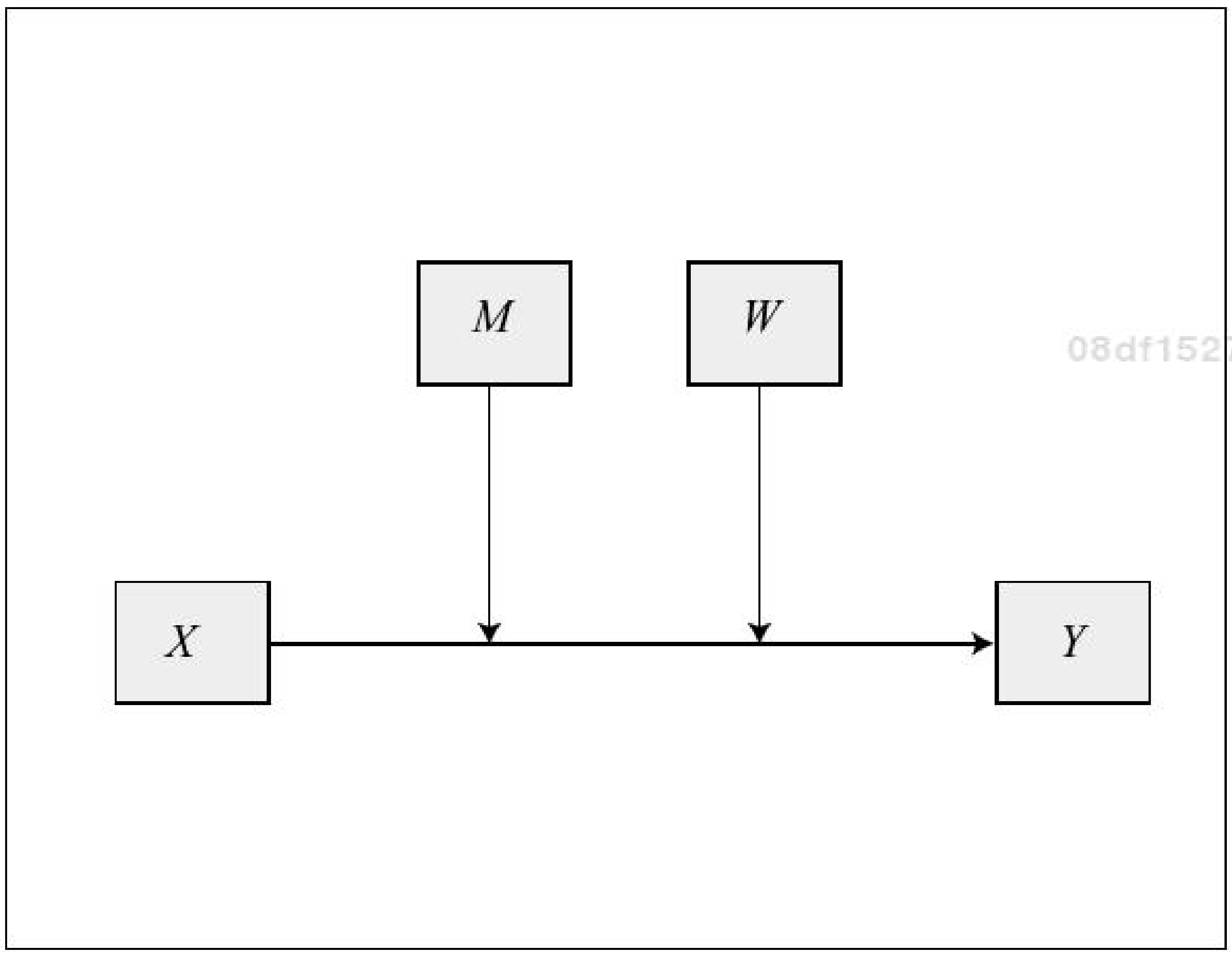
Statistical Diagram



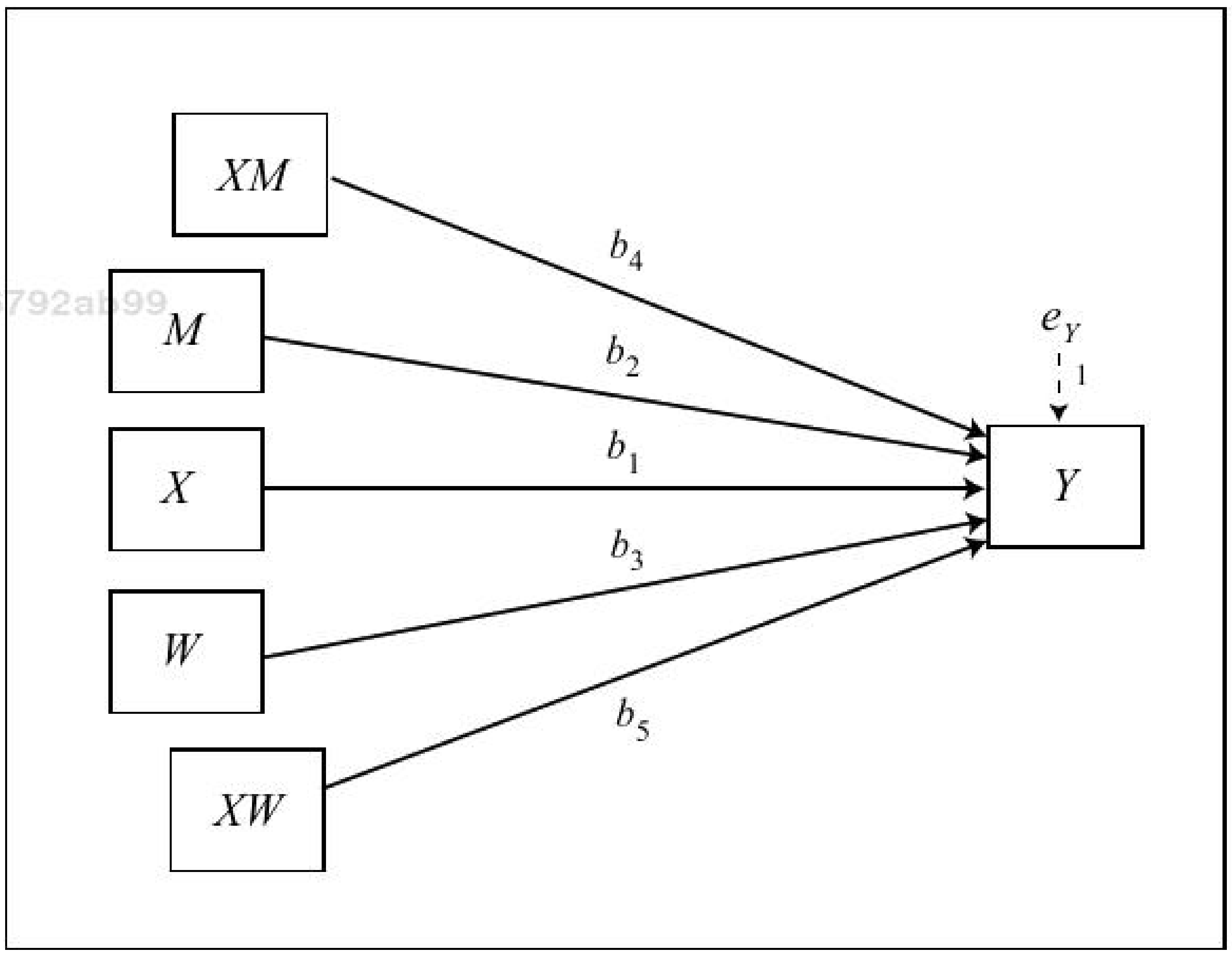
Conditional effect of X on $Y = b_1 + b_3M$

Model 2

Conceptual Diagram



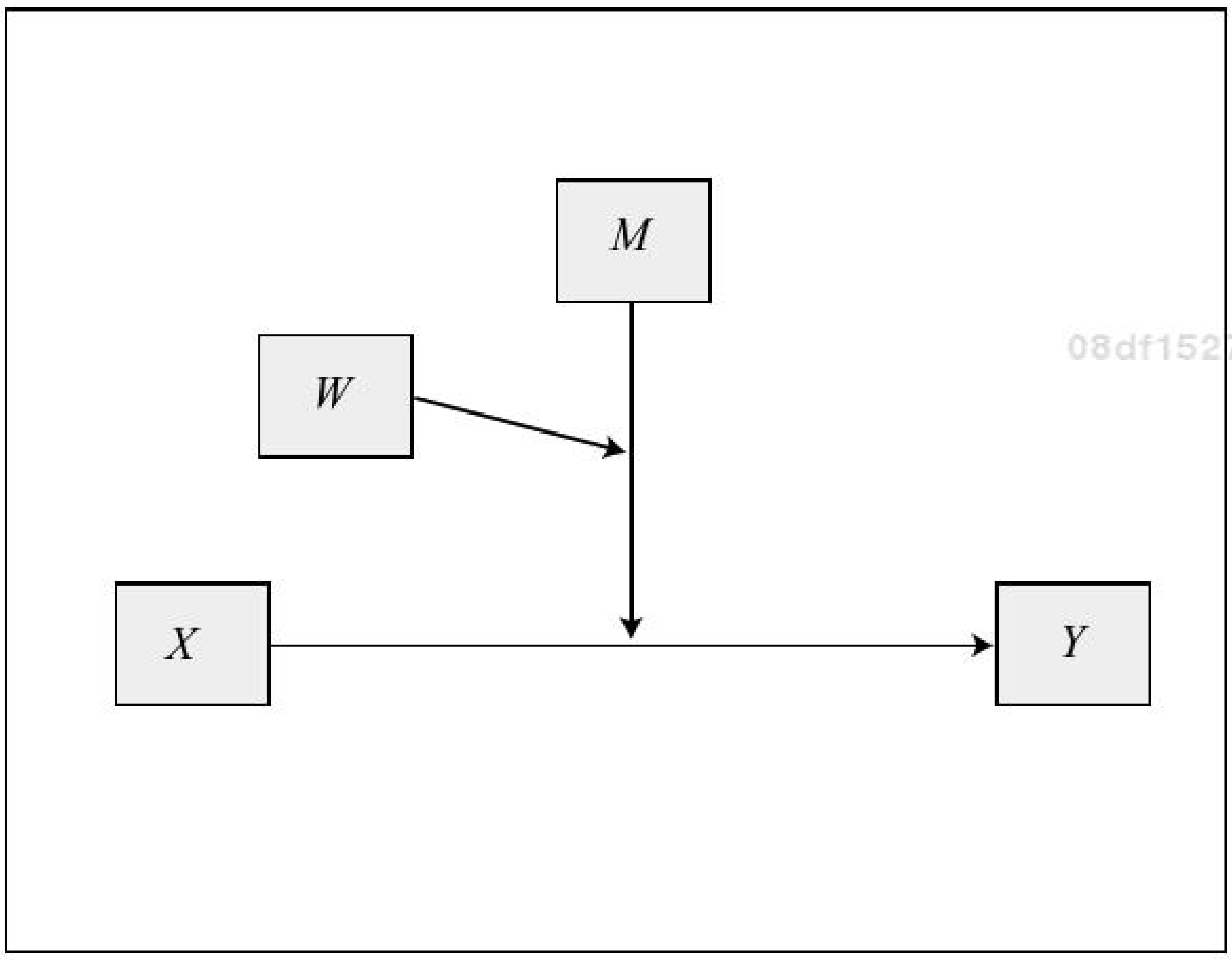
Statistical Diagram



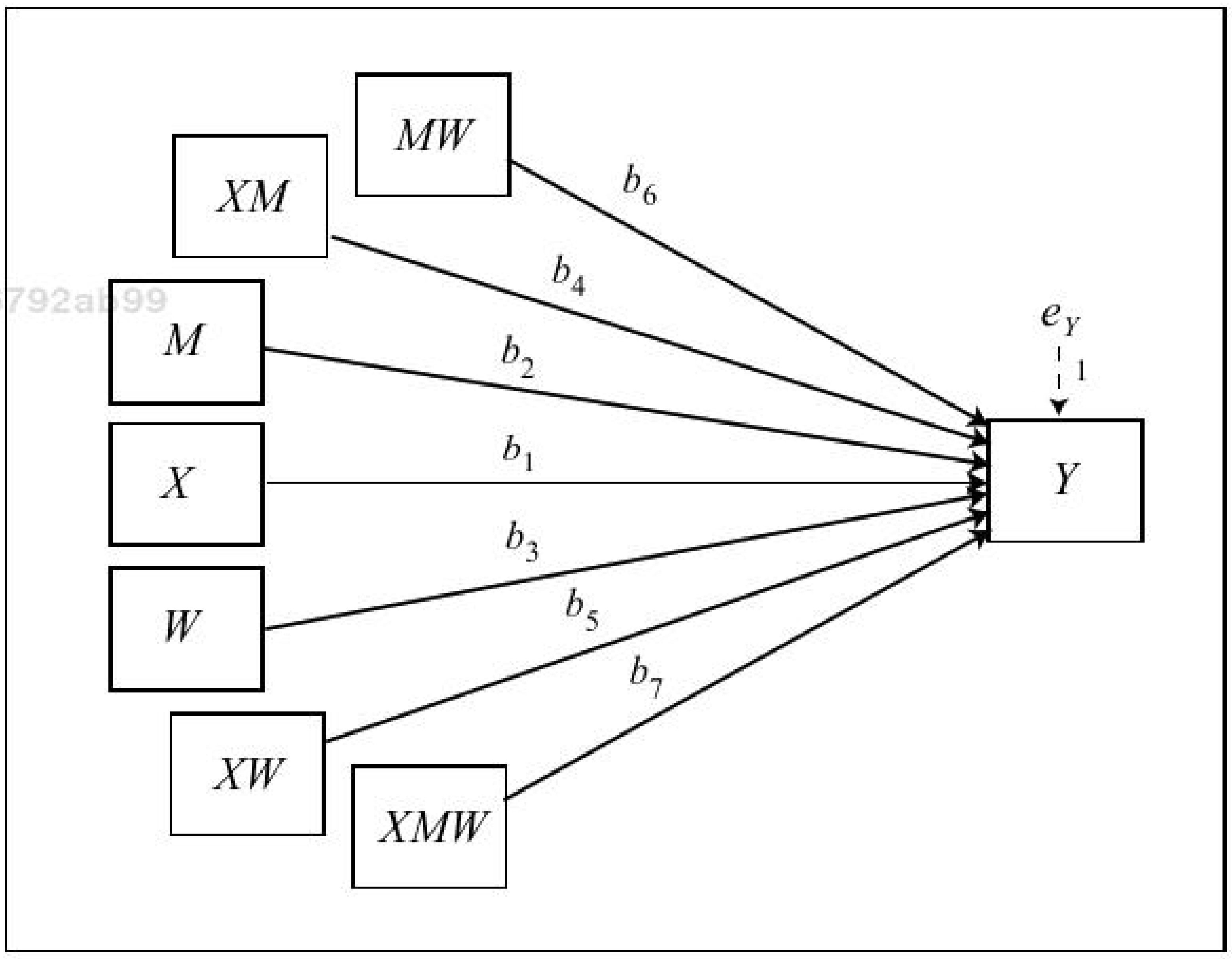
Conditional effect of X on $Y = b_1 + b_4M + b_5W$

Model 3

Conceptual Diagram



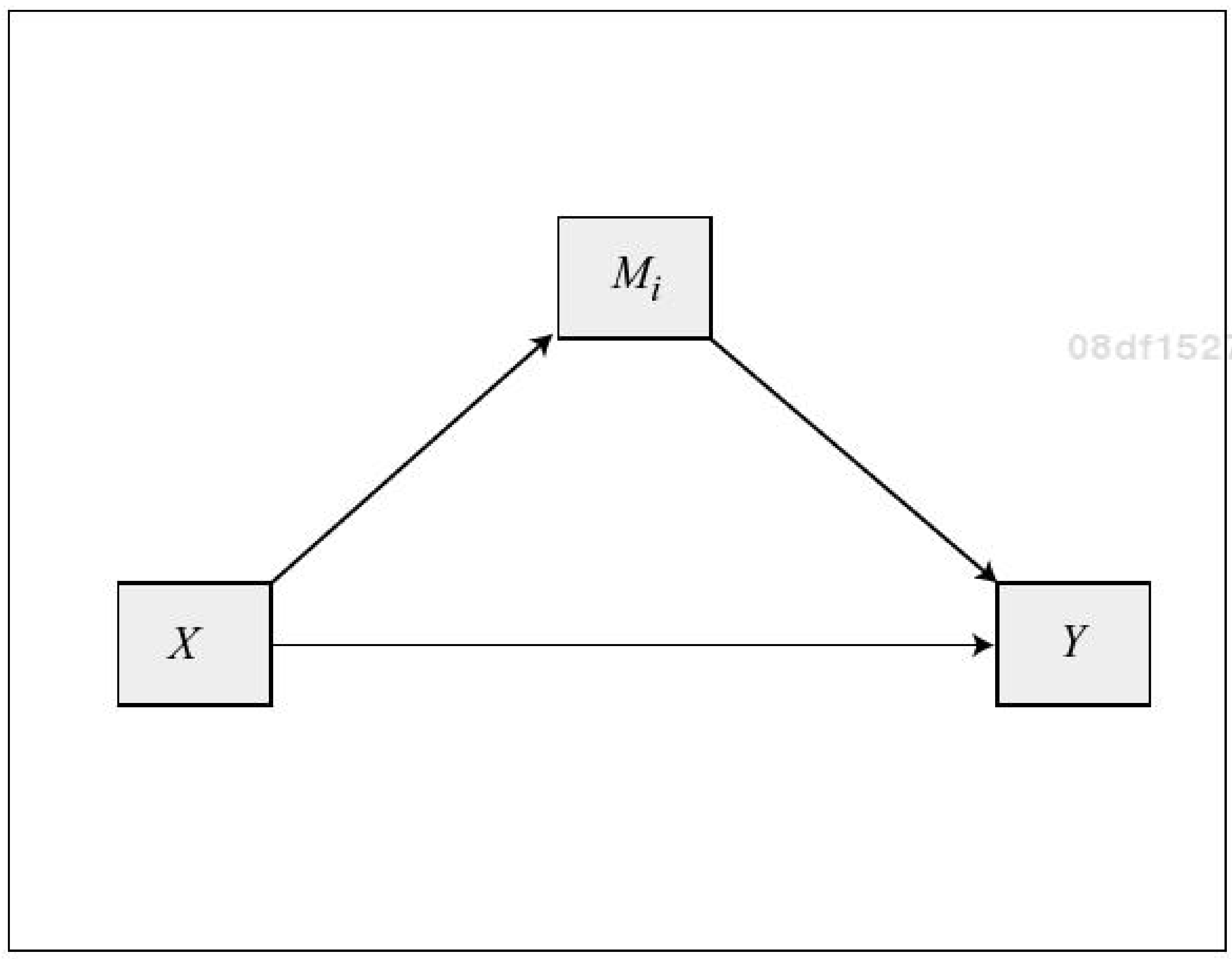
Statistical Diagram



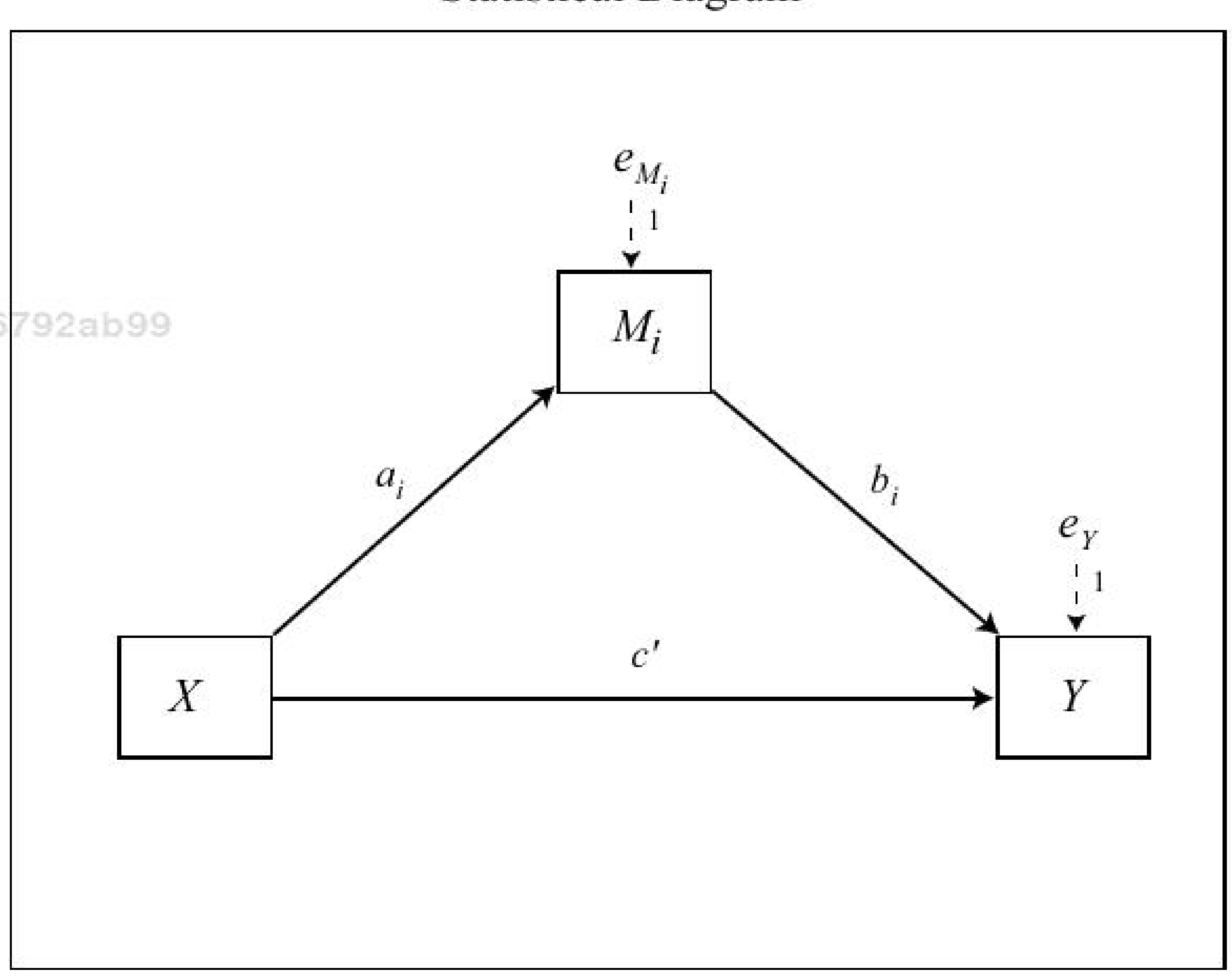
Conditional effect of X on $Y = b_1 + b_4M + b_5W + b_7MW$

Model 4

Conceptual Diagram



Statistical Diagram



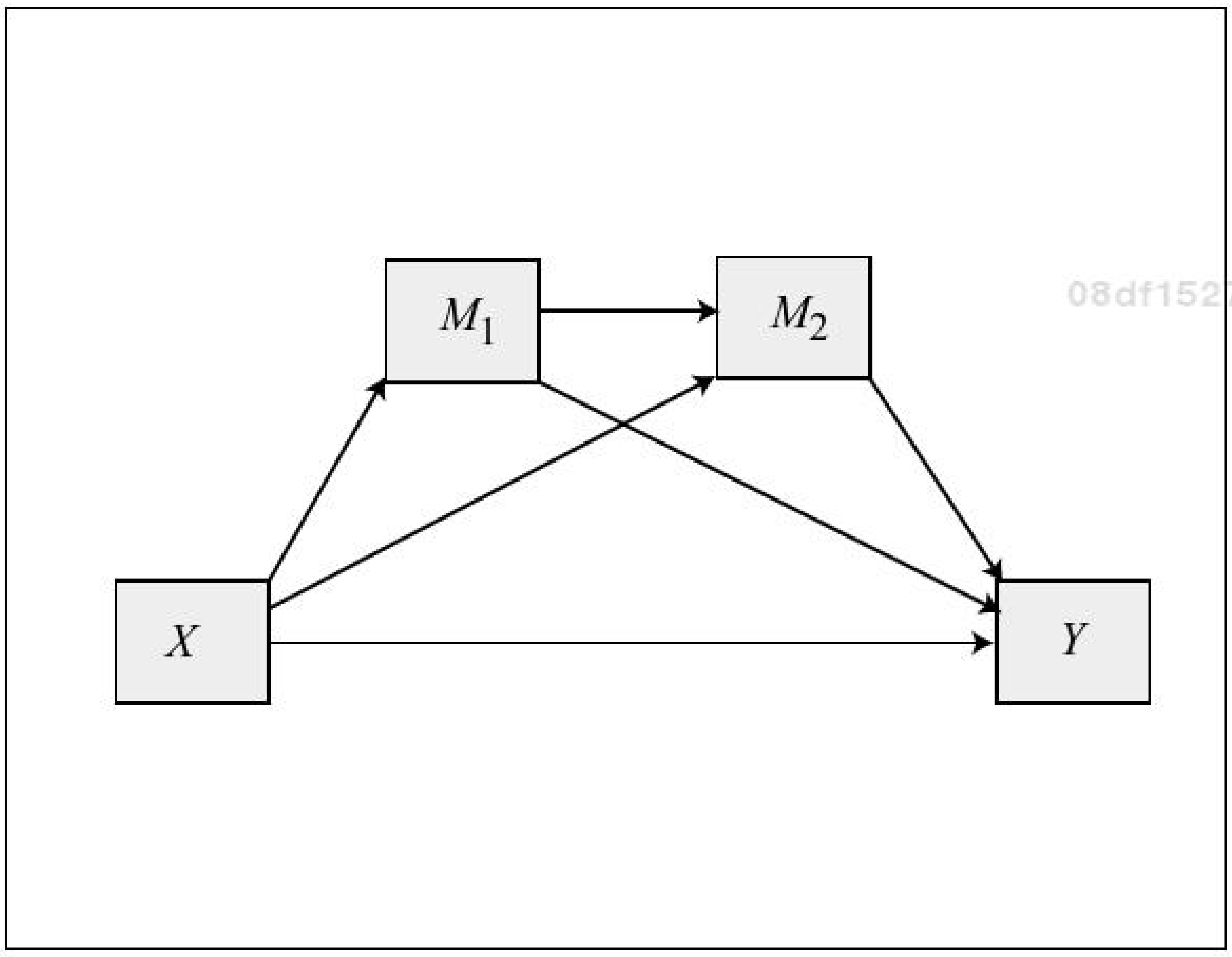
Indirect effect of X on Y through $M_i = a_i\, b_i$

Direct effect of X on $Y = c'$

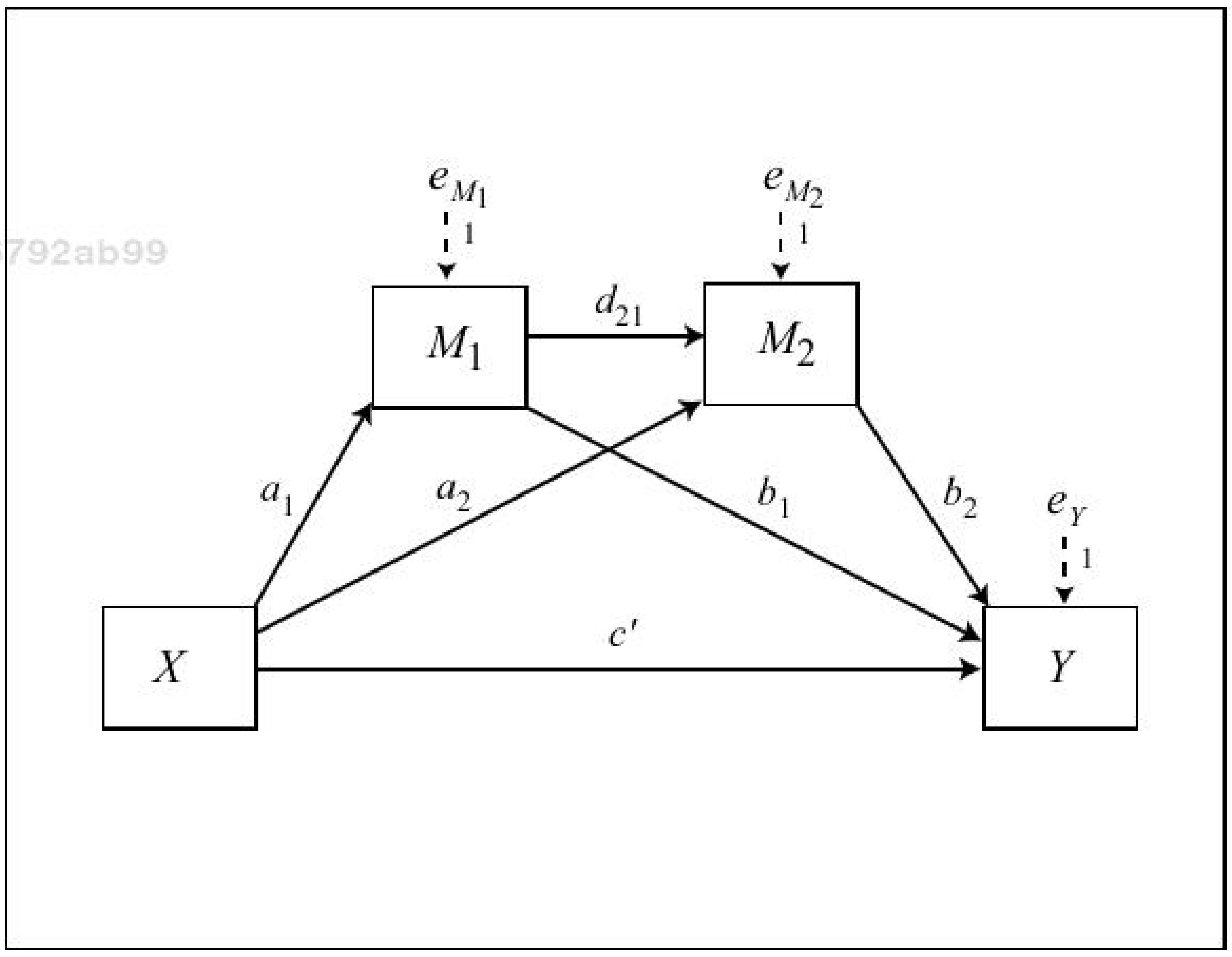
Note: Model 4 allows up to 10 mediators operating in parallel.

Model 6
(2 mediators)

Conceptual Diagram



Statistical Diagram

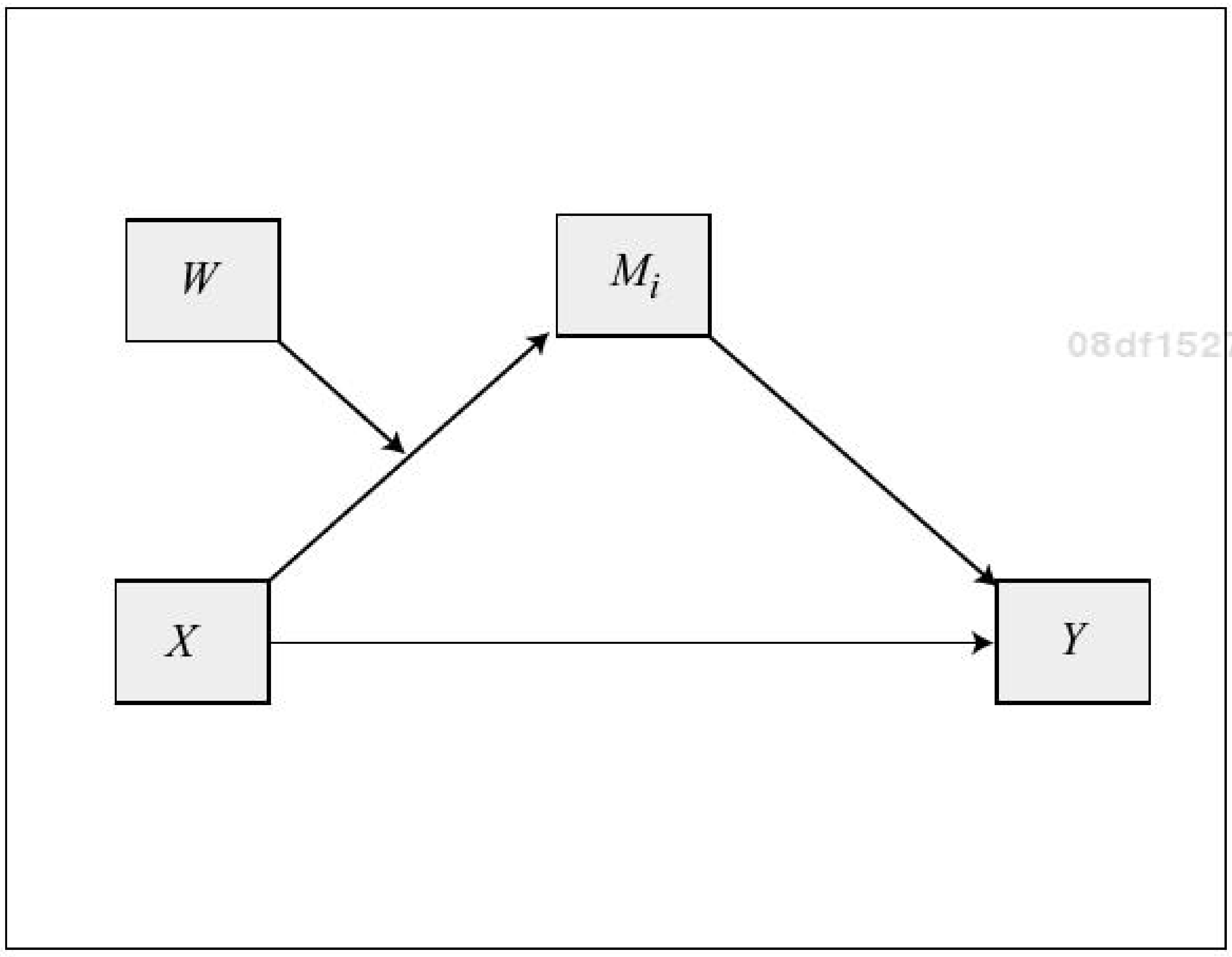


Indirect effect of X on Y through M_i only = $a_i\, b_i$
Indirect effect of X on Y through M_1 and M_2 in serial = $a_1\, d_{21}\, b_2$
Direct effect of X on $Y = c'$

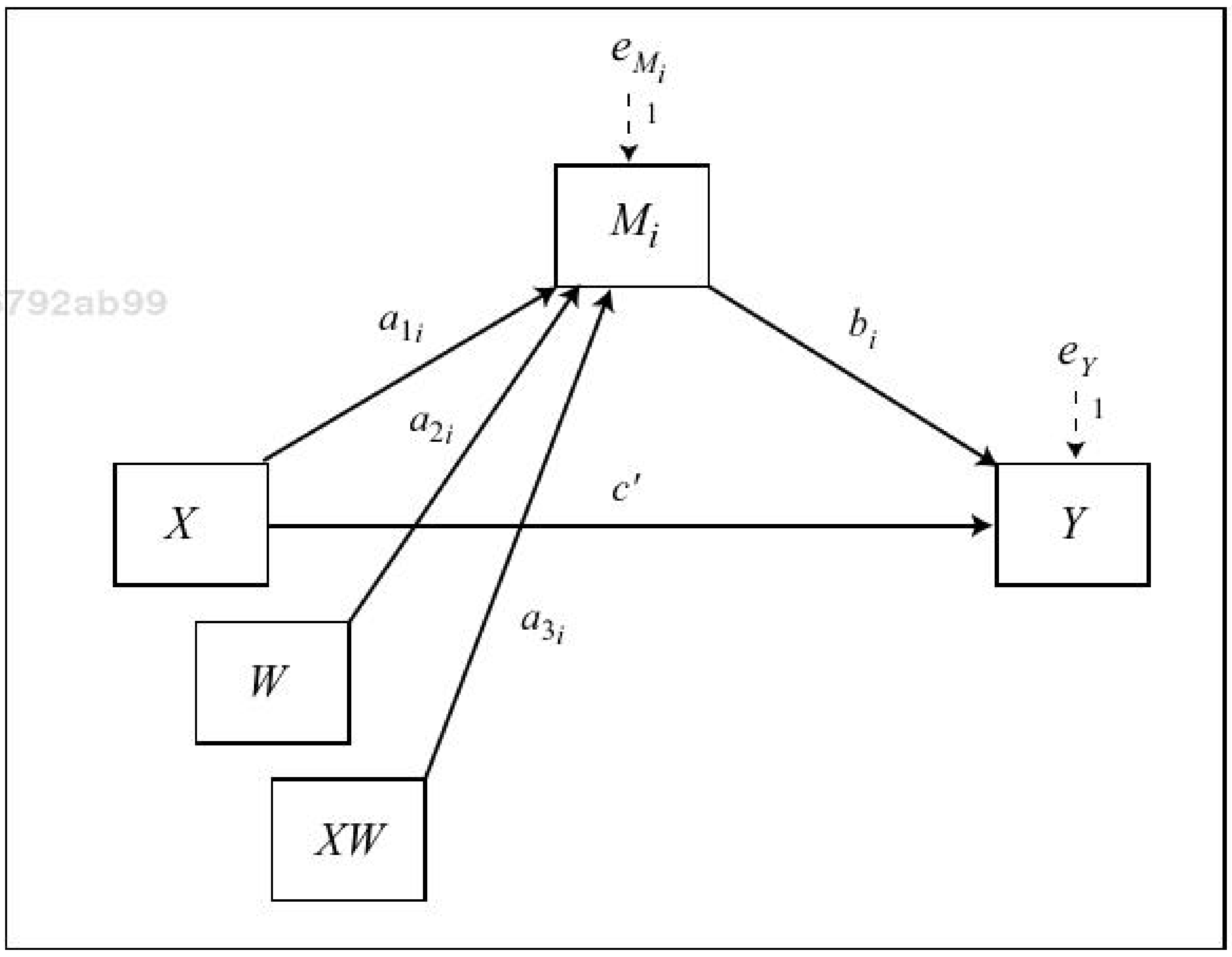
Note: Model 6 allows up to four mediators operating in serial.

Model 7

Conceptual Diagram



Statistical Diagram



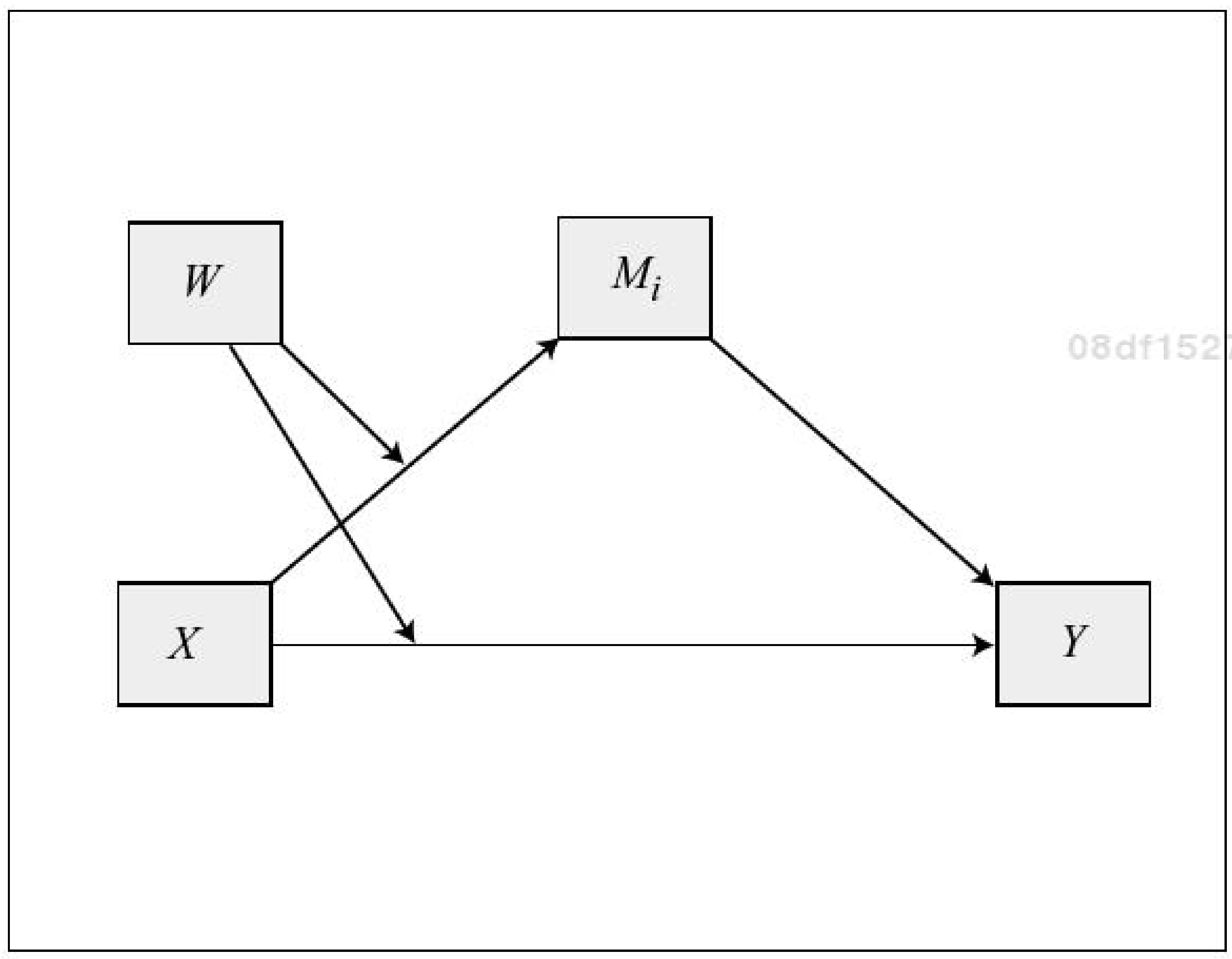
Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)b_i$

Direct effect of X on $Y = c'$

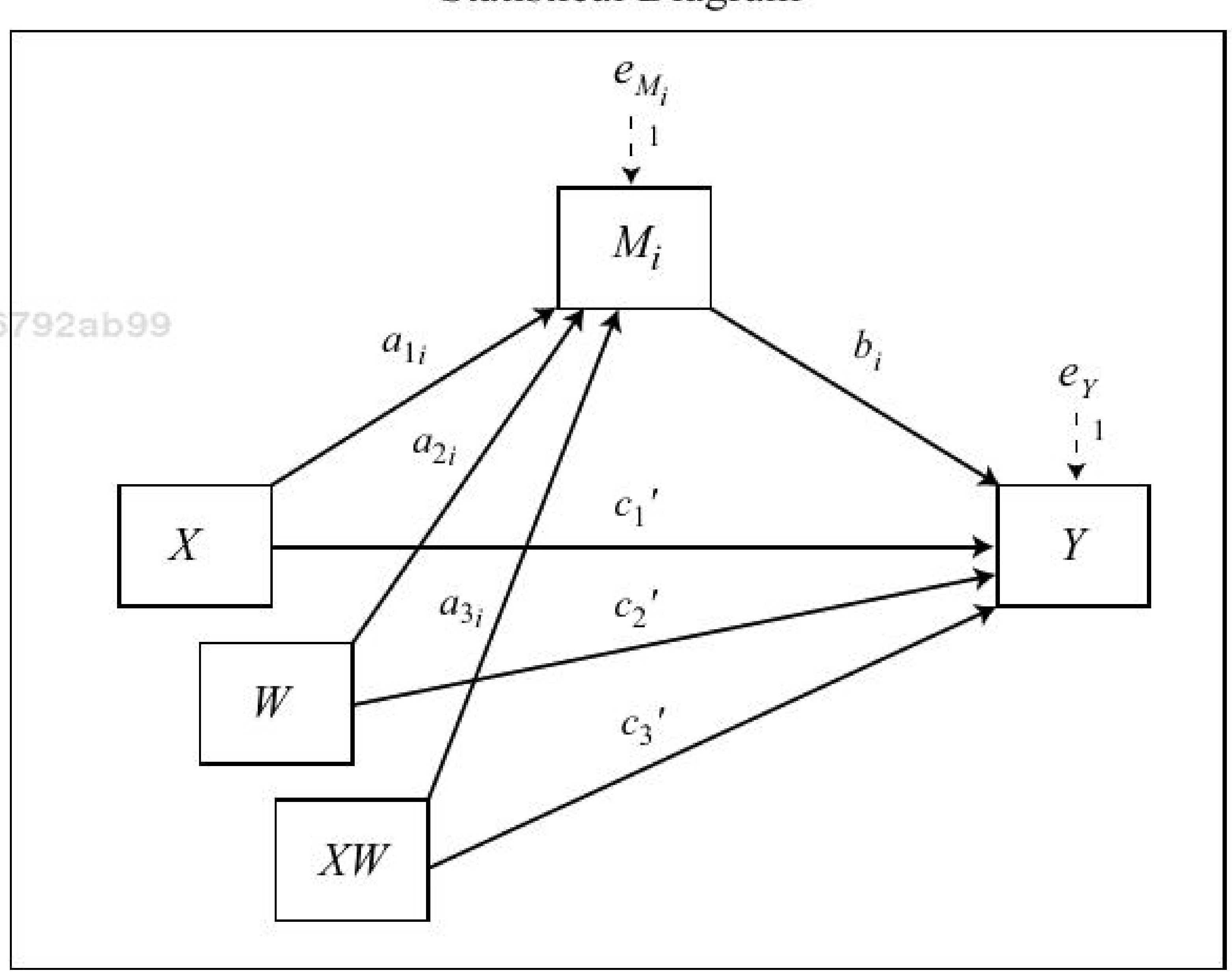
Note: Model 7 allows up to 10 mediators operating in parallel.

Model 8

Conceptual Diagram



Statistical Diagram



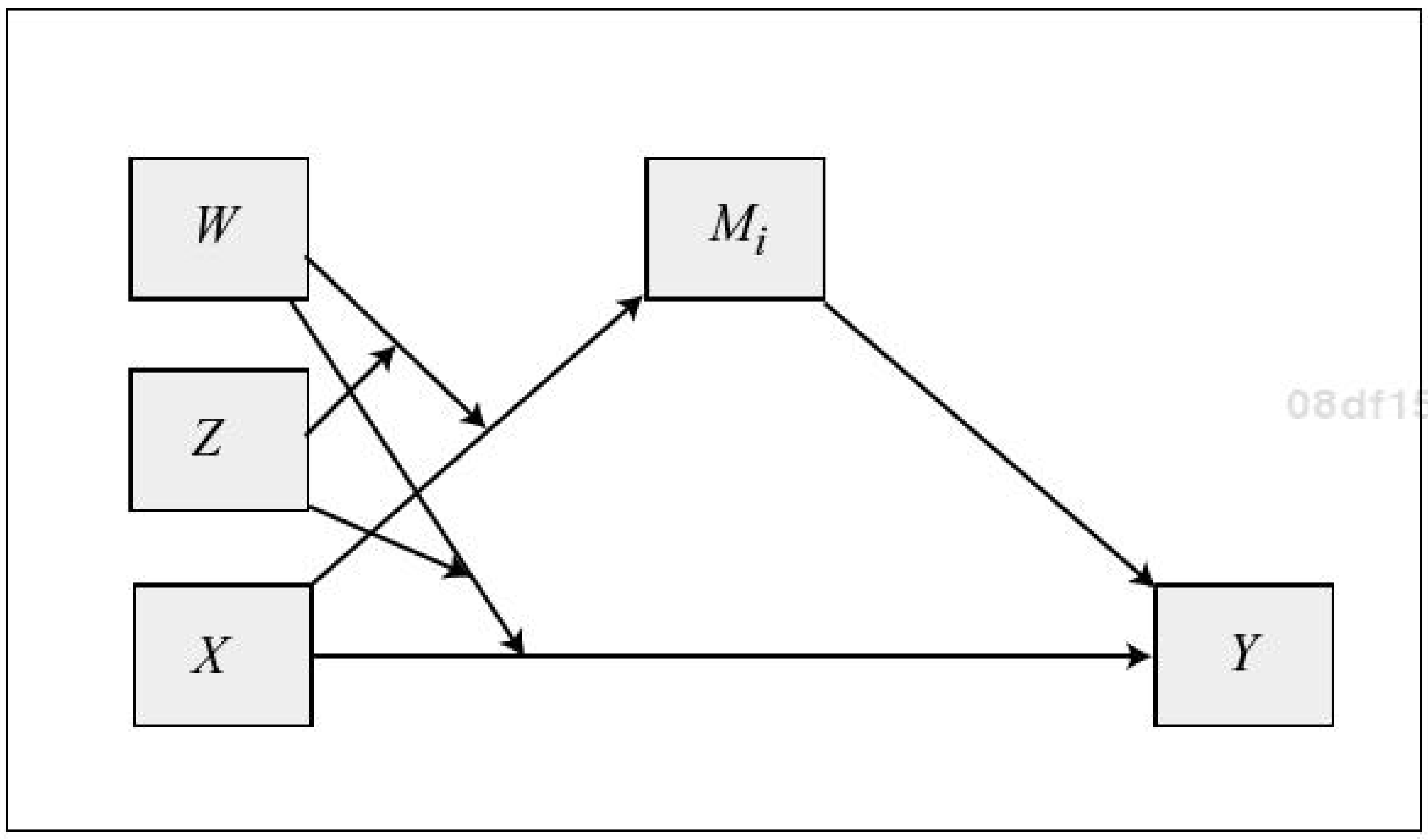
Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)b_i$

Conditional direct effect of X on $Y = c_1' + c_3'W$

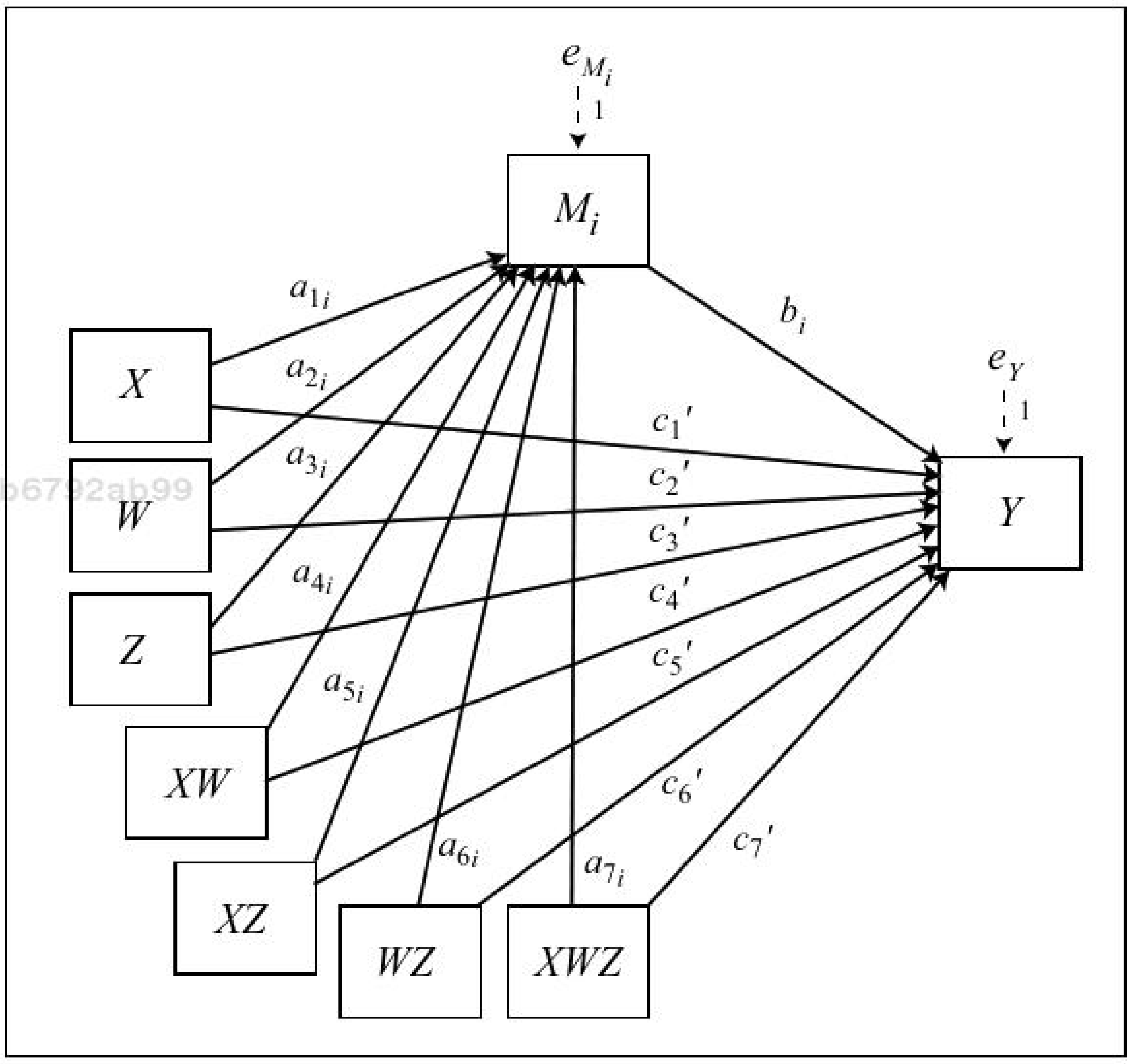
Note: Model 8 allows up to 10 mediators operating in parallel.

Model 12

Conceptual Diagram



Statistical Diagram



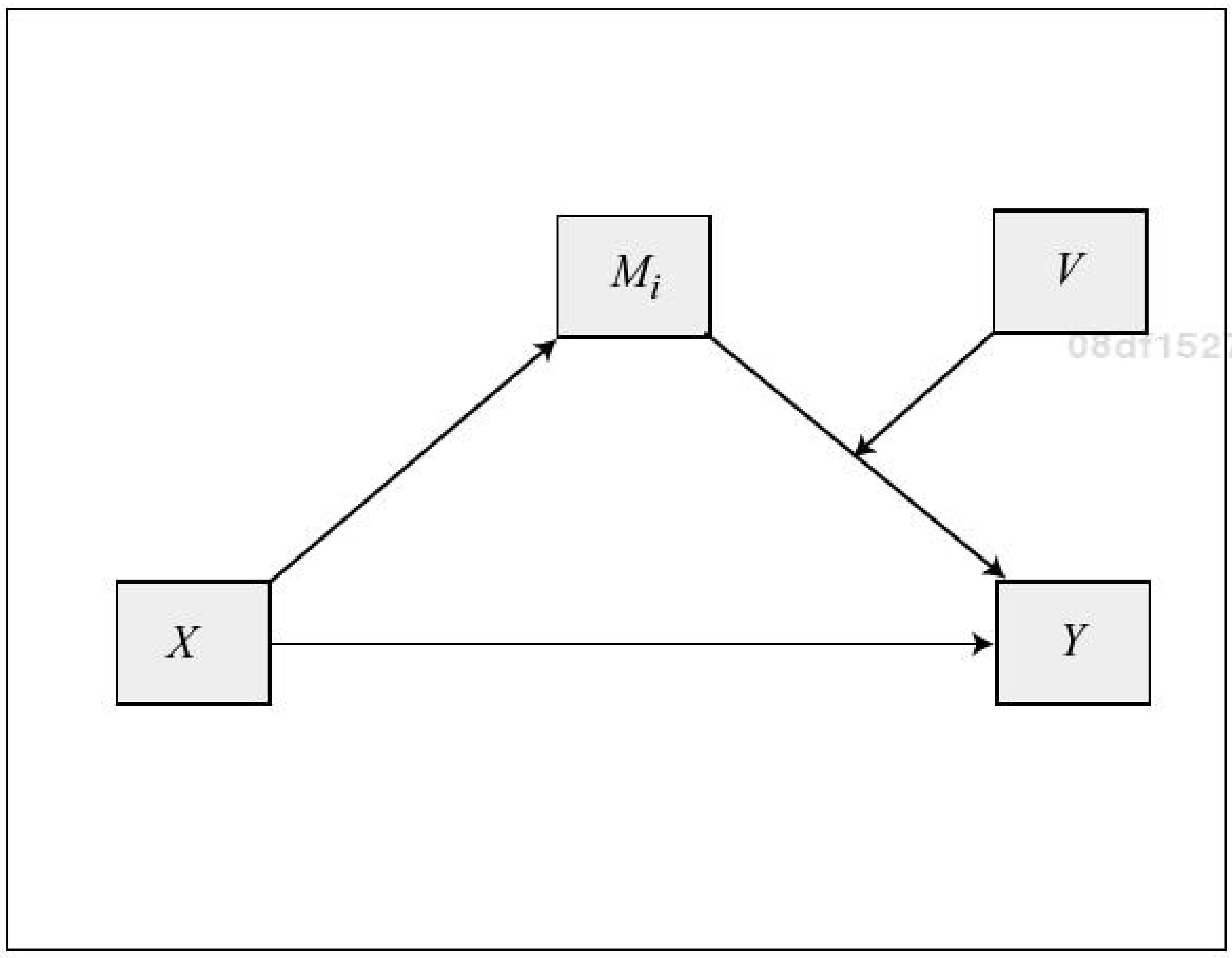
Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{4i}W + a_{5i}Z + a_{7i}WZ)\, b_i$

Conditional direct effect of X on $Y = c_1' + c_4'W + c_5'Z + c_7'WZ$

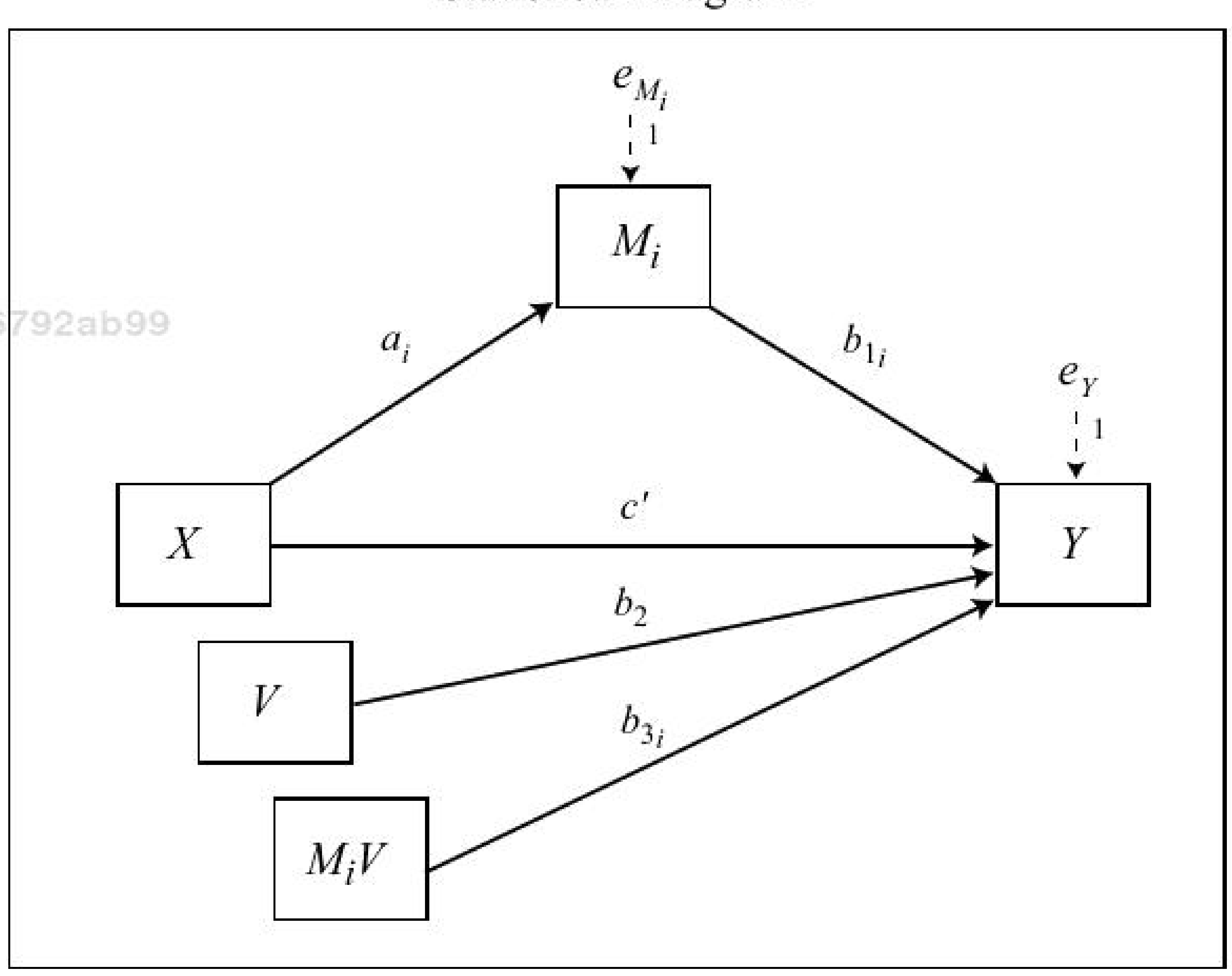
Note: Model 12 allows up to 10 mediators operating in parallel.

Model 14

Conceptual Diagram



Statistical Diagram

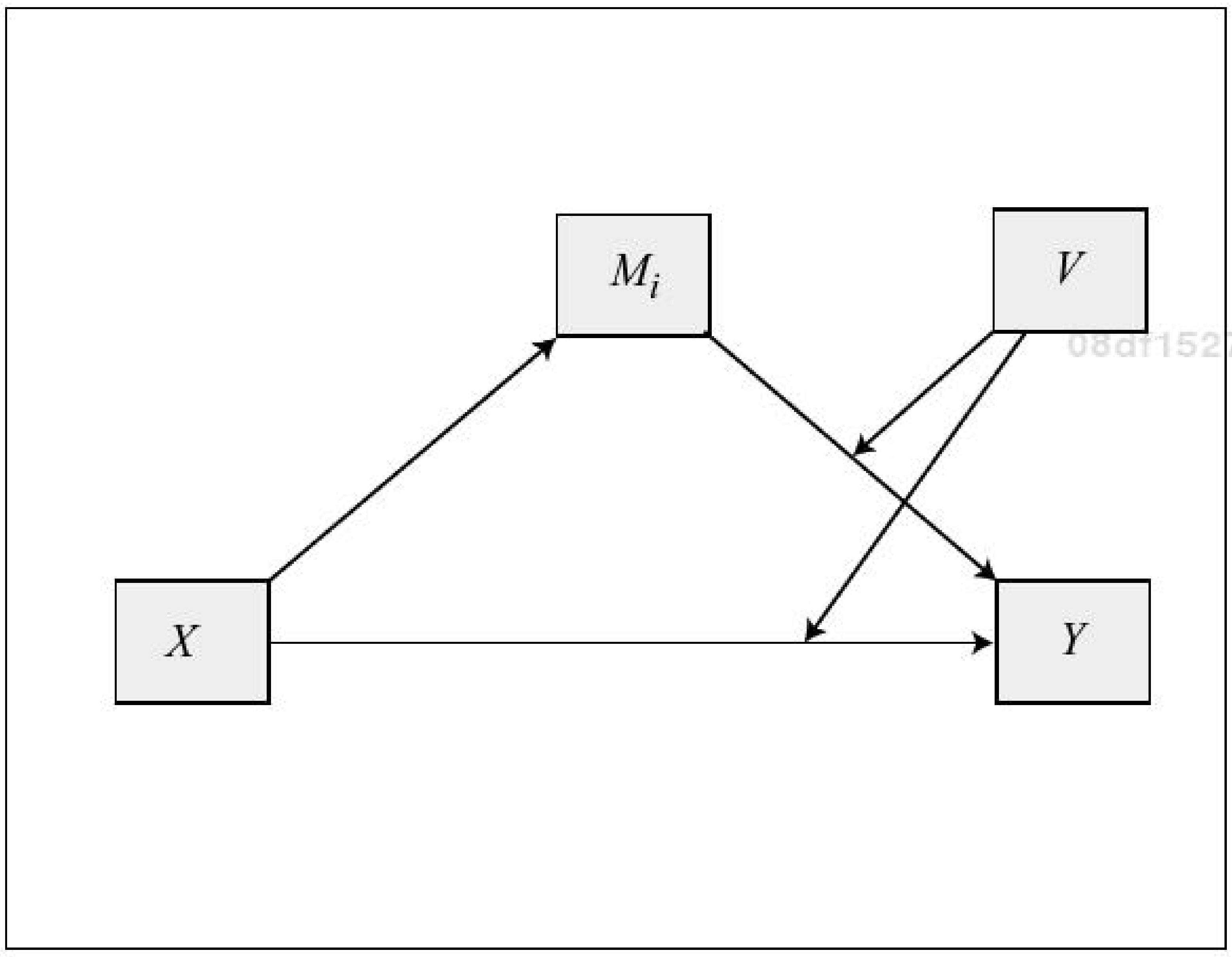


Conditional indirect effect of X on Y through $M_i = a_i\,(b_{1i} + b_{3i}V)$
Direct effect of X on $Y = c'$

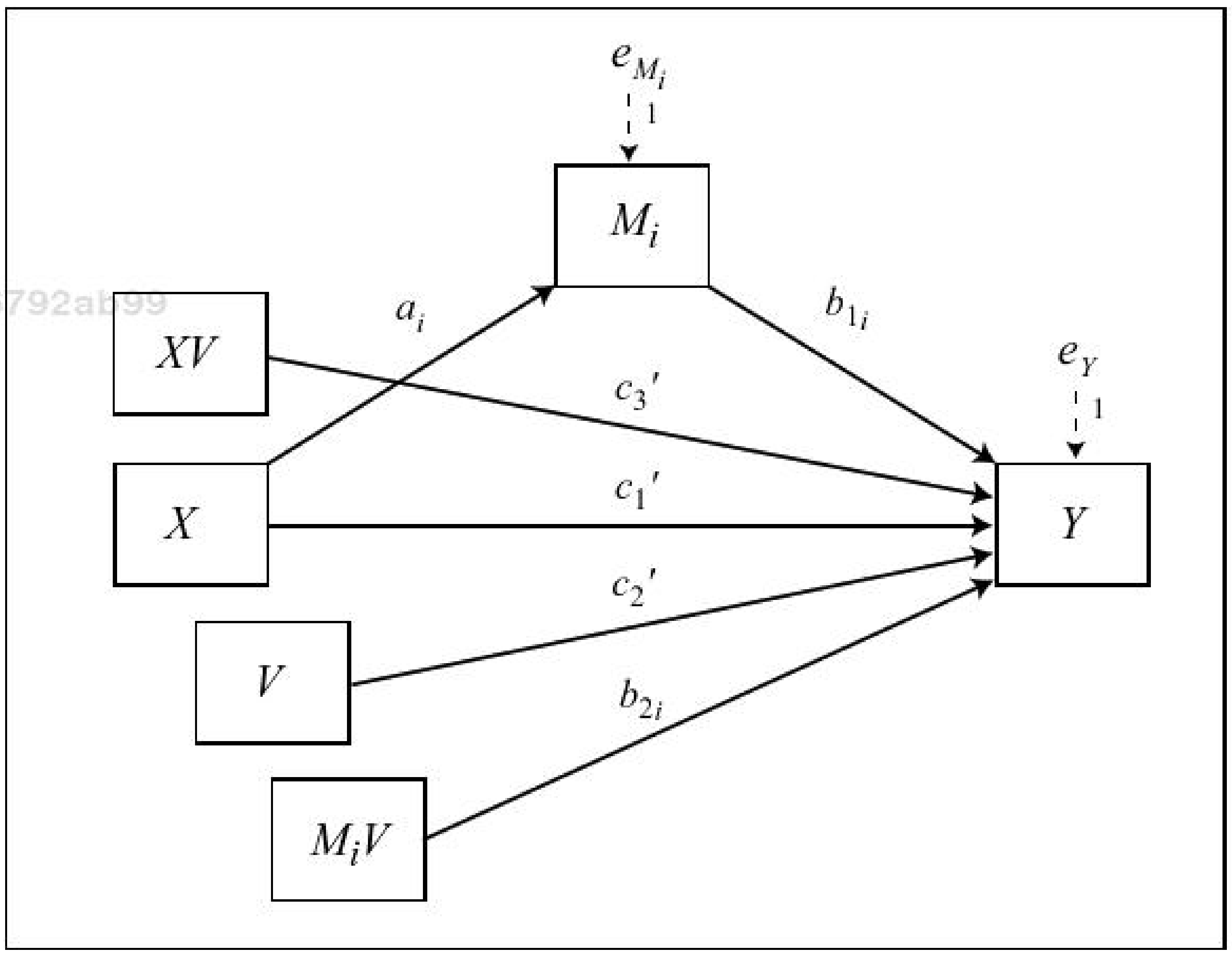
Note: Model 14 allows up to 10 mediators operating in parallel.

Model 15

Conceptual Diagram



Statistical Diagram

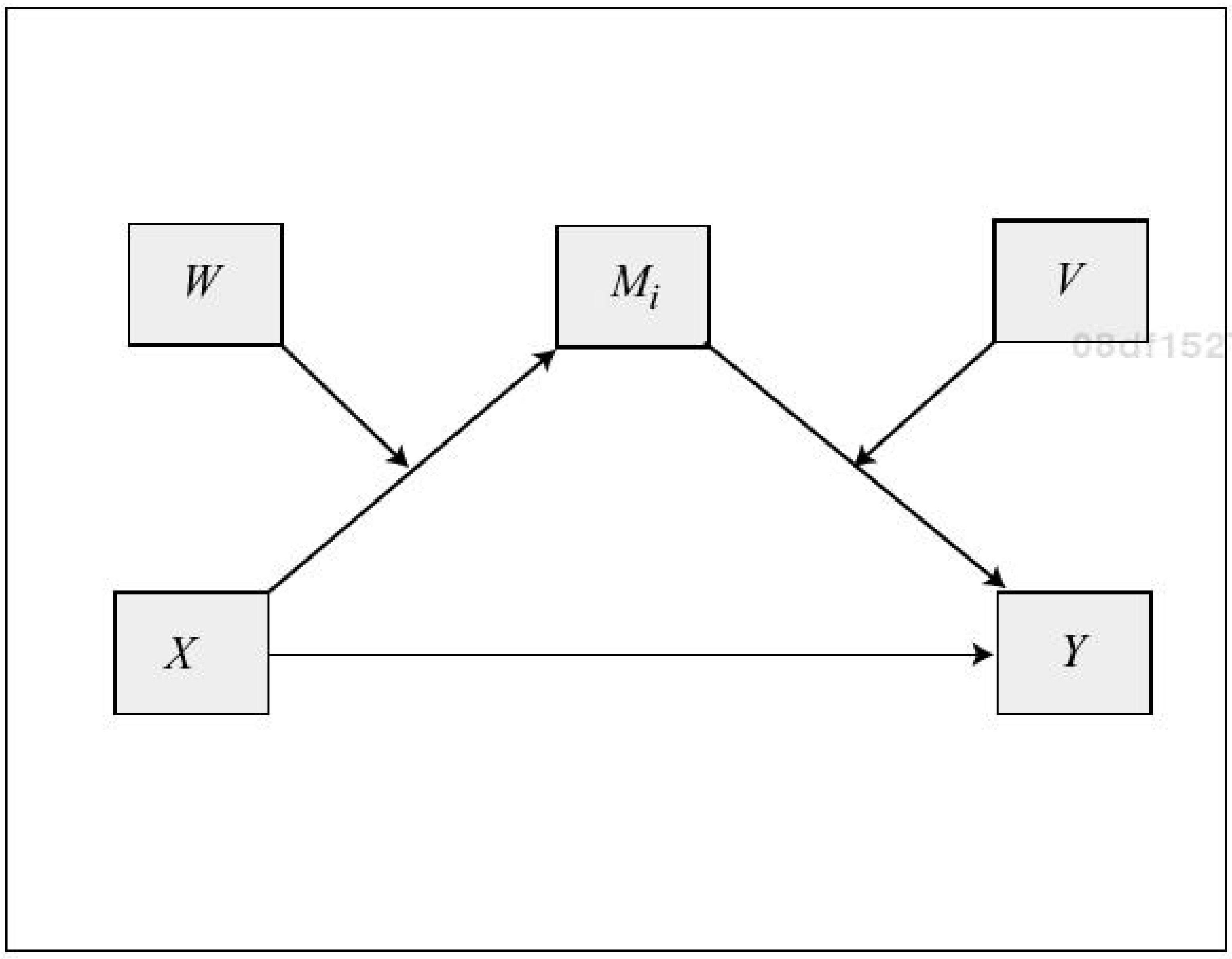


Conditional indirect effect of X on Y through $M_i = a_i\,(b_{1i} + b_{2i}V)$
Conditional direct effect of X on $Y = c_1' + c_3'V$

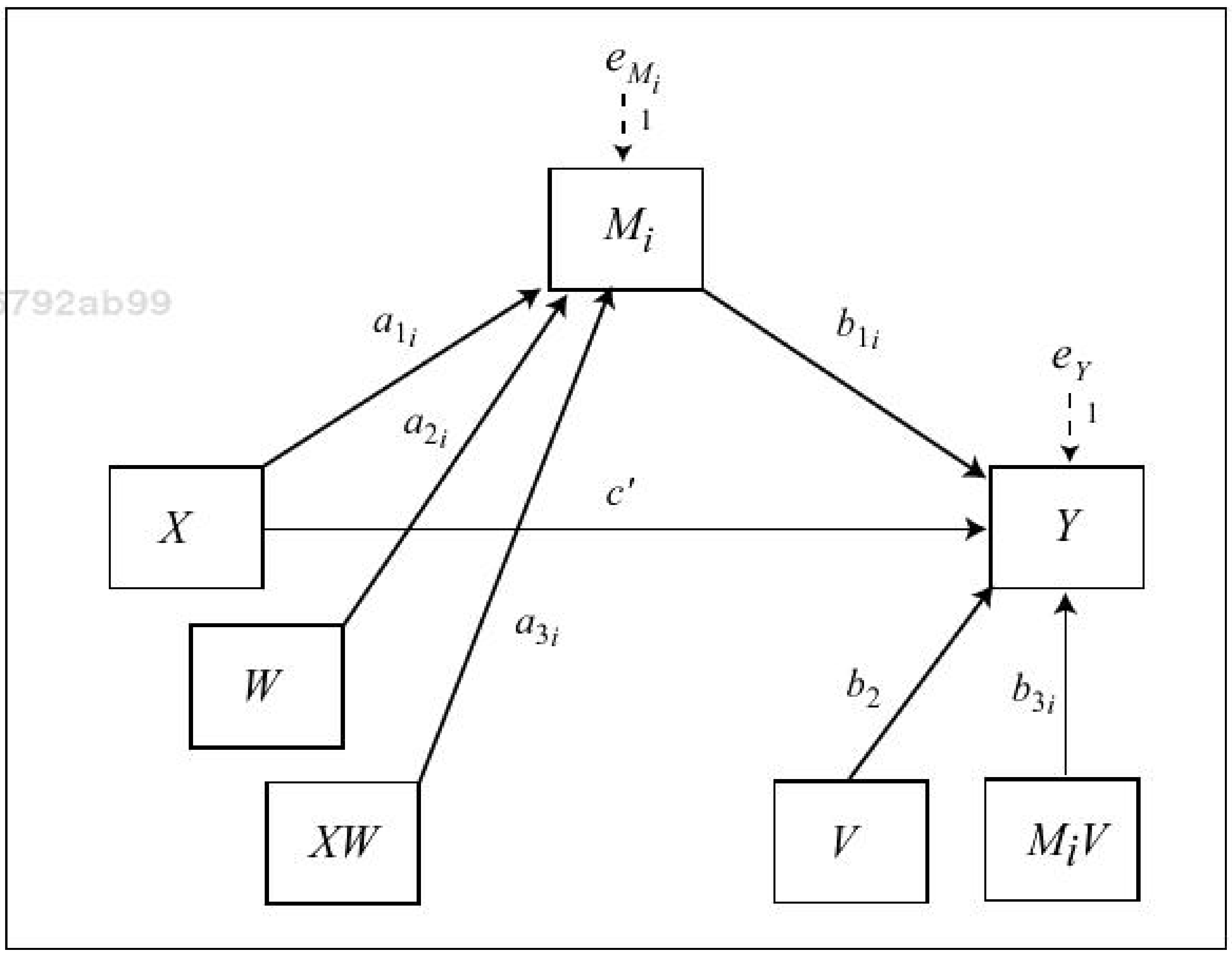
Note: Model 15 allows up to 10 mediators operating in parallel.

Model 21

Conceptual Diagram



Statistical Diagram

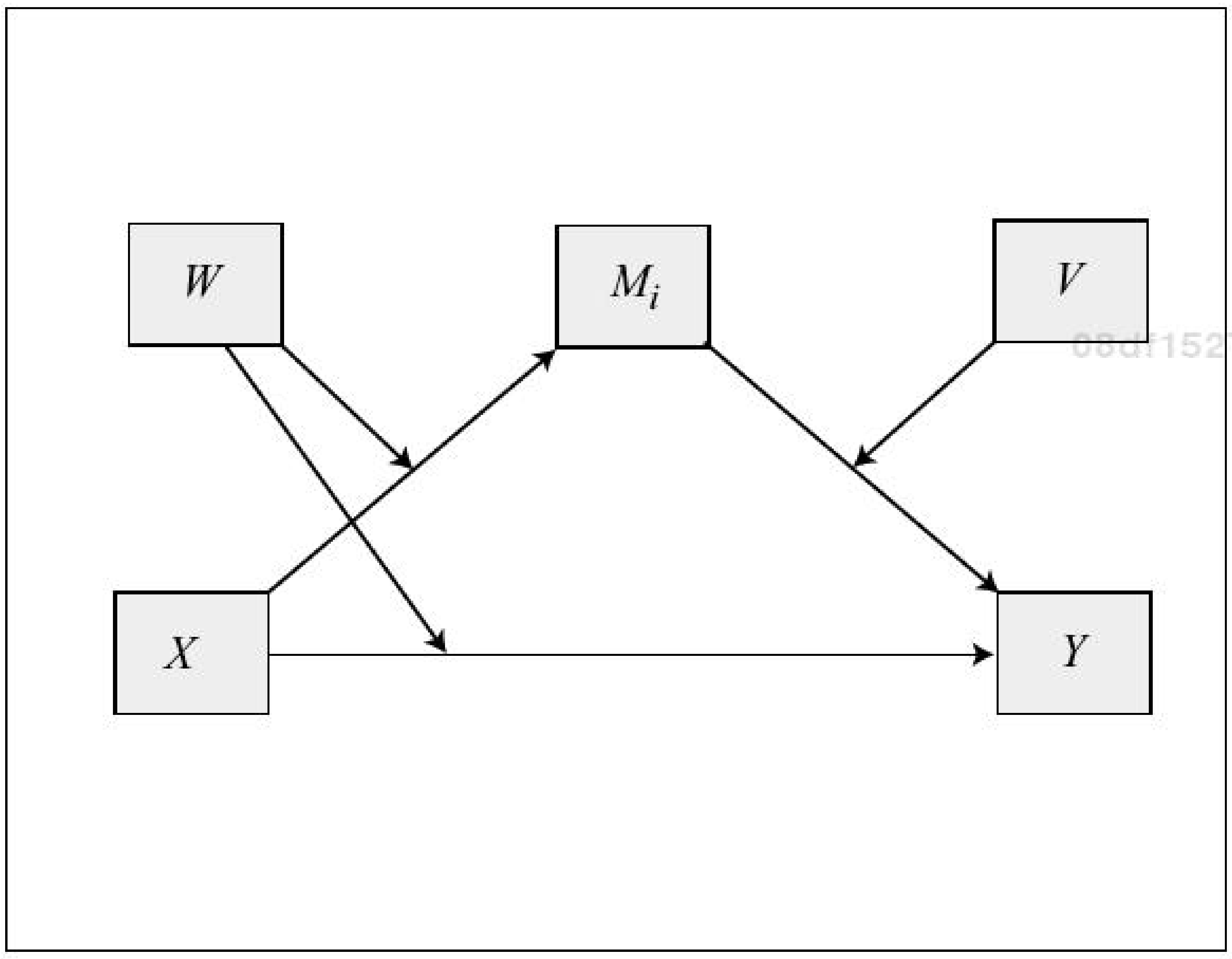


Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)(b_{1i} + b_{3i}V)$
Direct effect of X on $Y = c'$

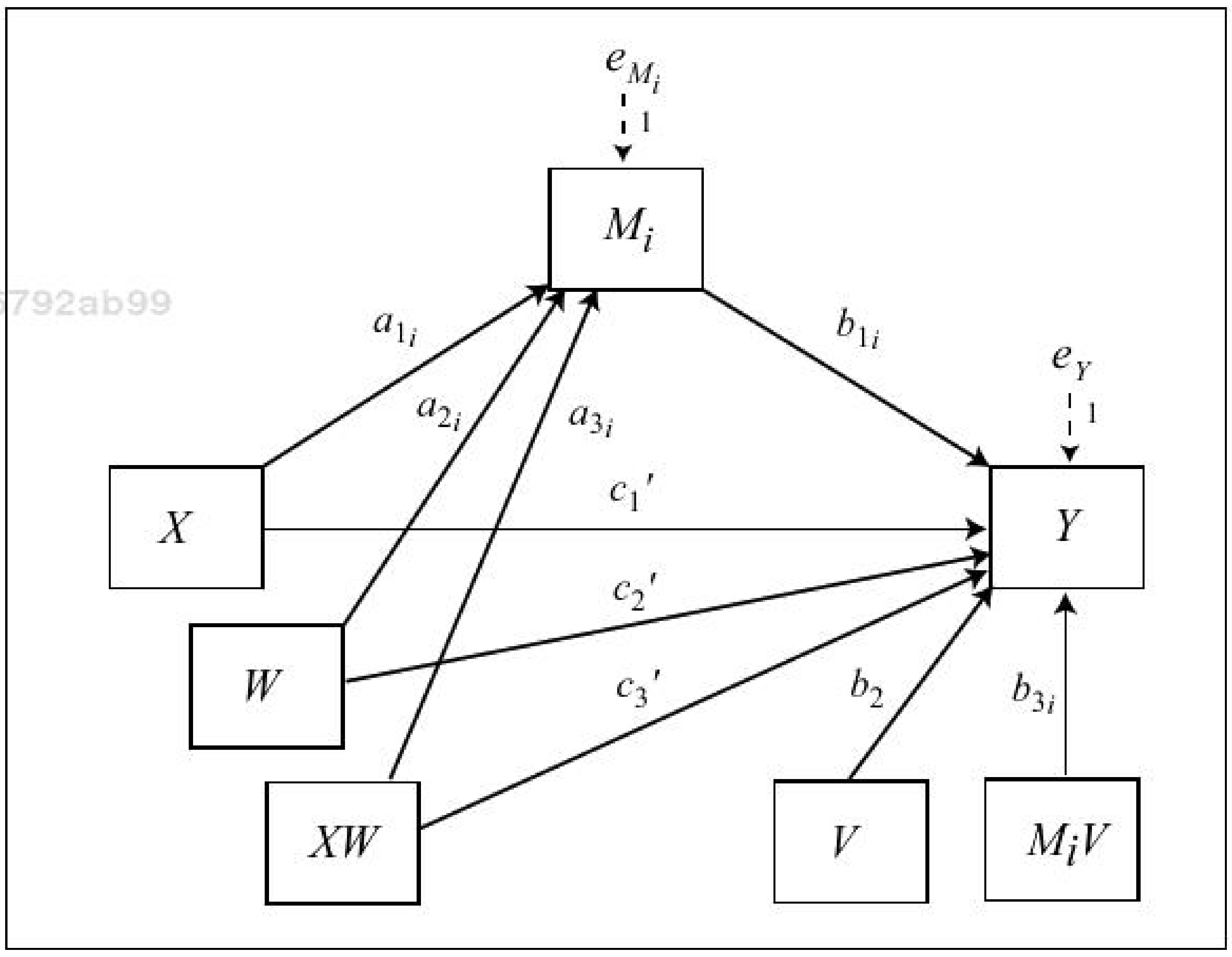
Note: Model 21 allows up to 10 mediators operating in parallel.

Model 22

Conceptual Diagram



Statistical Diagram



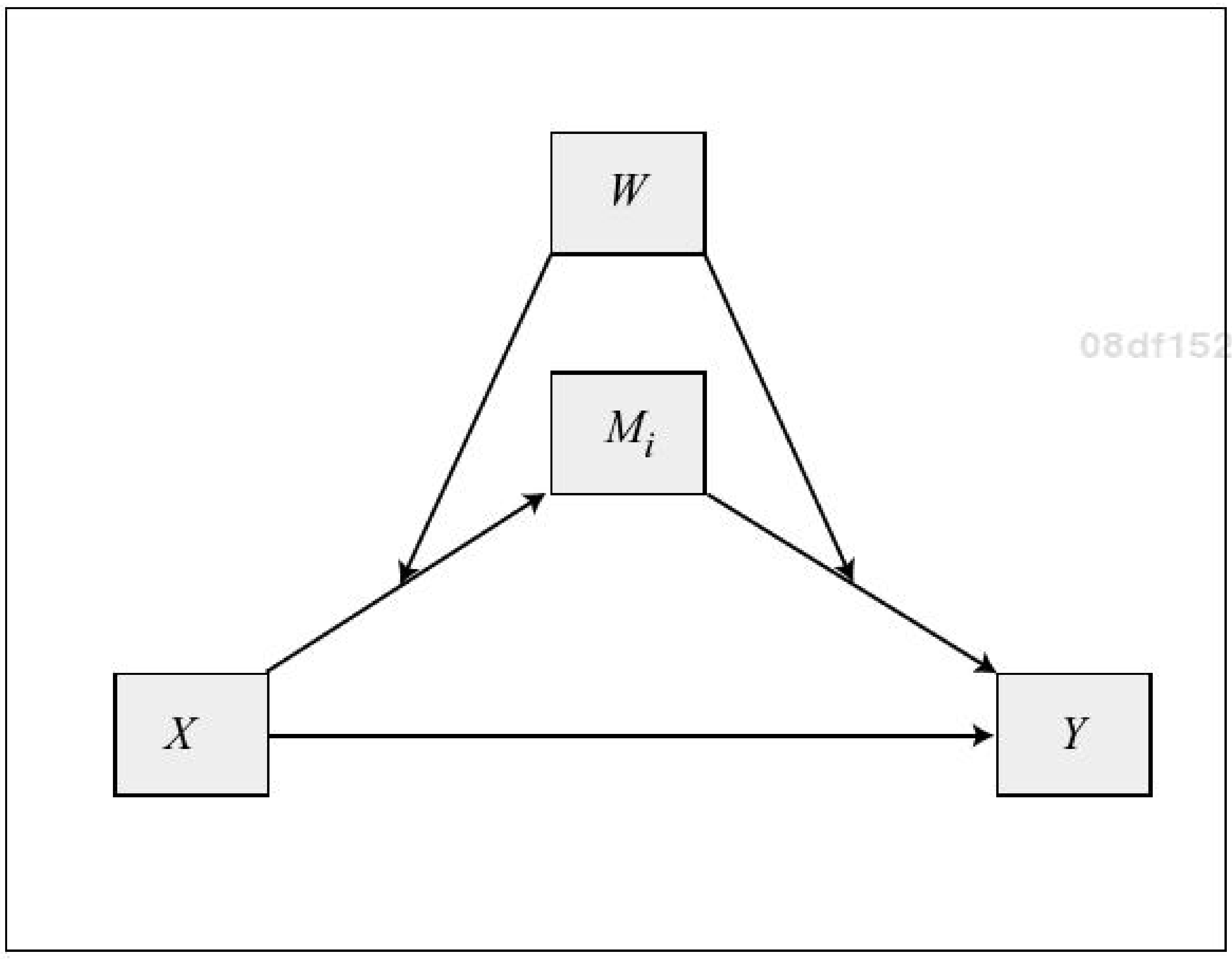
Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)(b_{1i} + b_{3i}V)$

Conditional direct effect of X on $Y = c_1' + c_3'W$

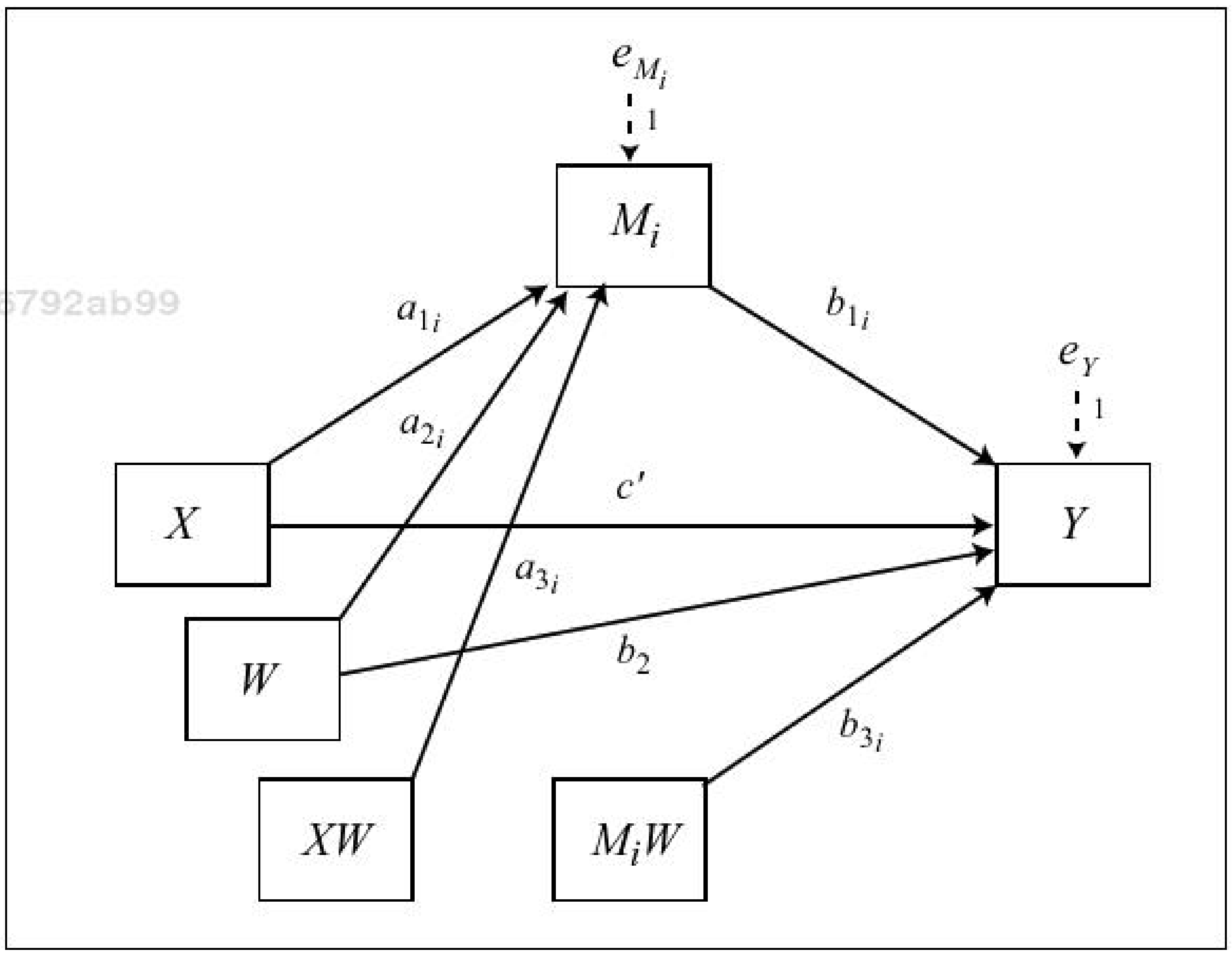
Note: Model 22 allows up to 10 mediators operating in parallel.

Model 58

Conceptual Diagram



Statistical Diagram

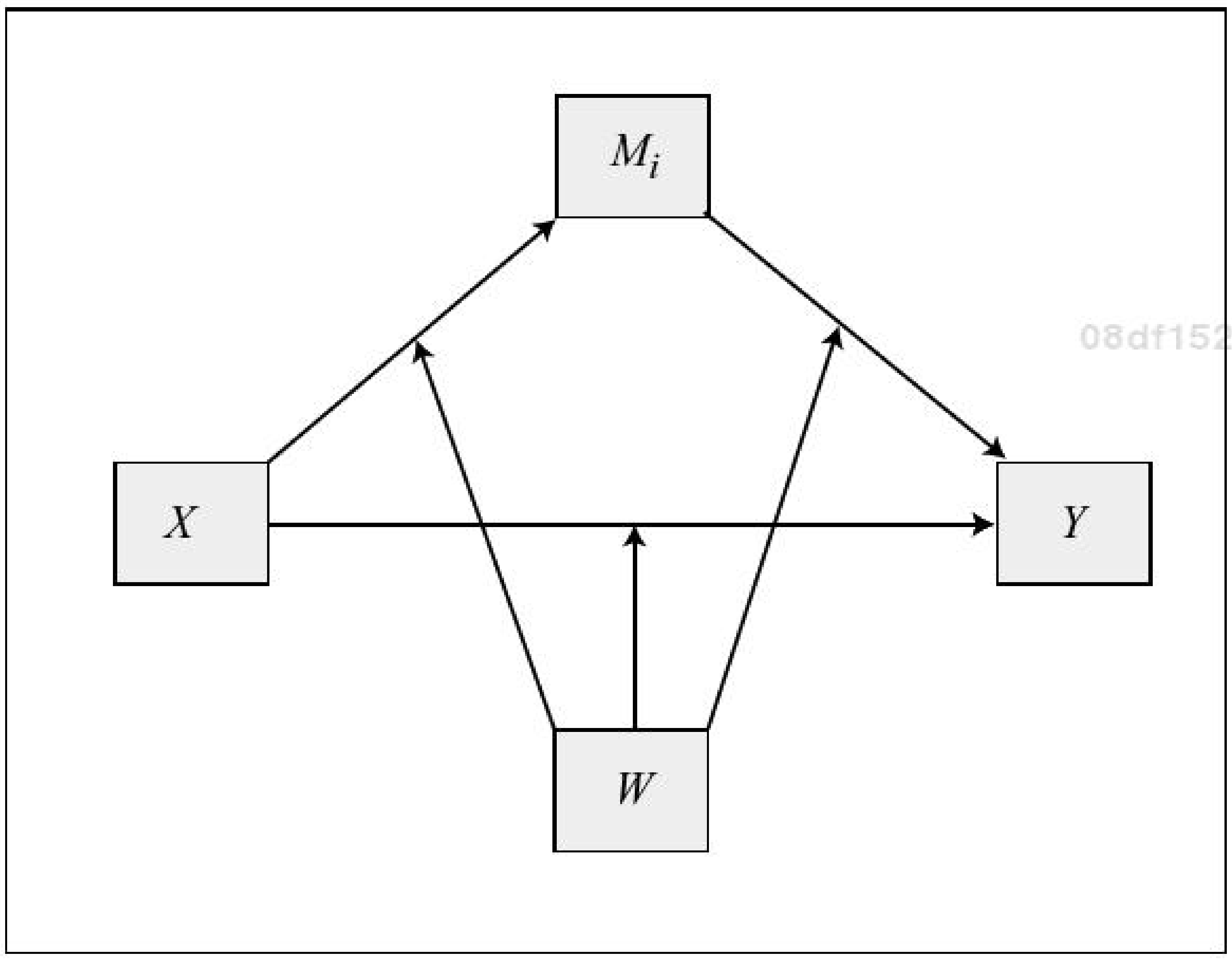


Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)\,(b_{1i} + b_{3i}W)$
Direct effect of X on $Y = c'$

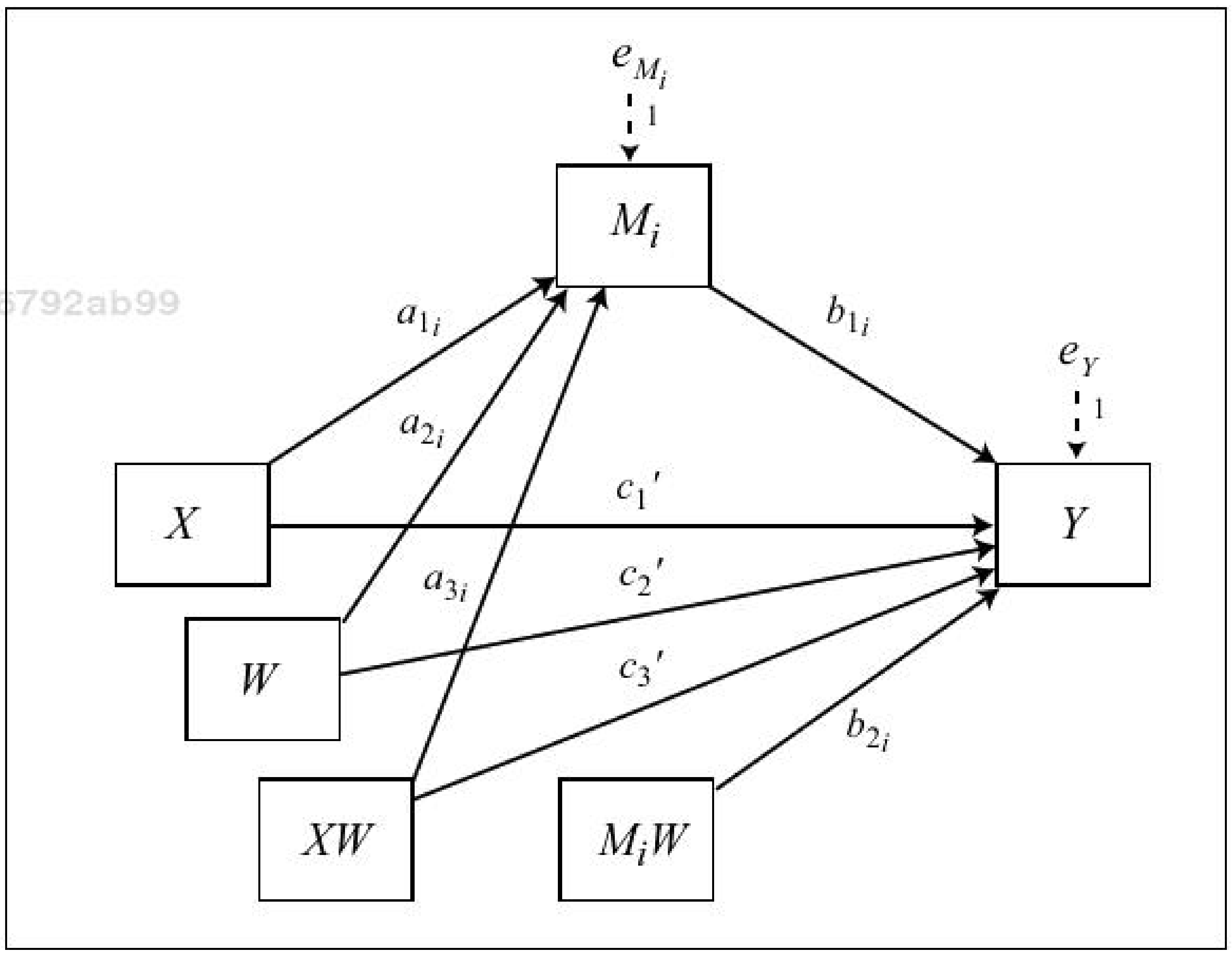
Note: Model 58 allows up to 10 mediators operating in parallel.

Model 59

Conceptual Diagram



Statistical Diagram

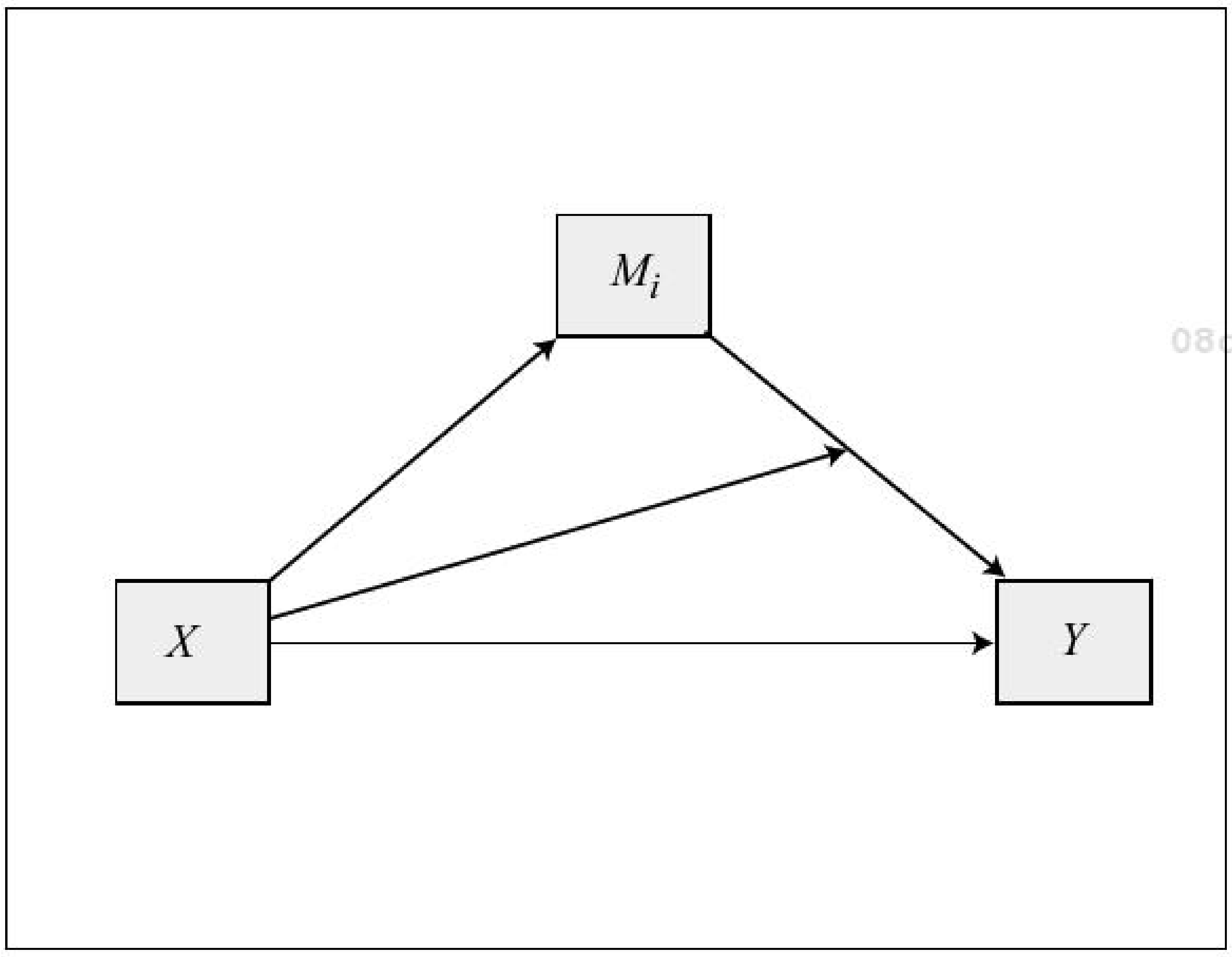


Conditional indirect effect of X on Y through $M_i = (a_{1i} + a_{3i}W)\,(b_{1i} + b_{2i}W)$
Conditional direct effect of X on $Y = c_1' + c_3'W$

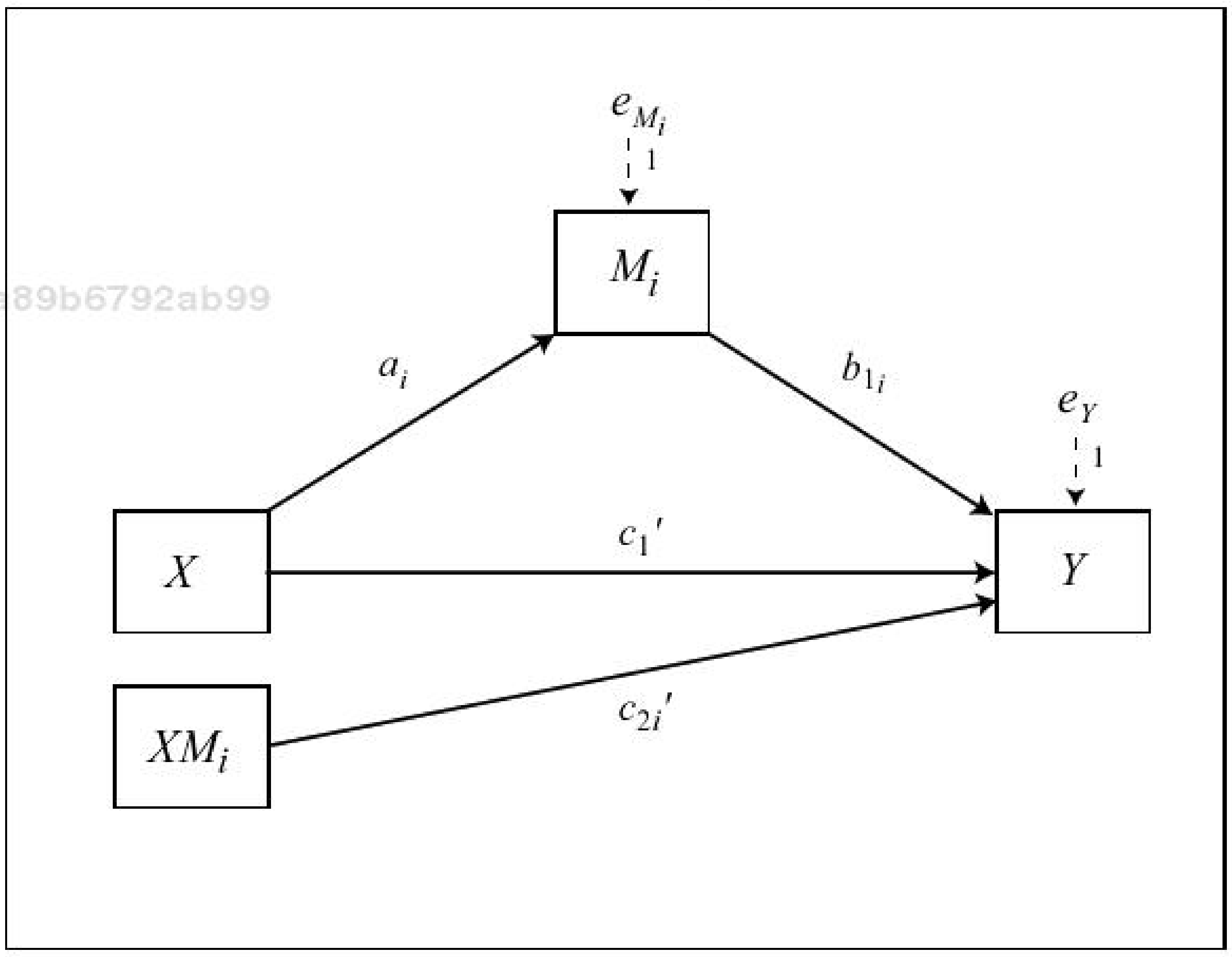
Note: Model 59 allows up to 10 mediators operating in parallel.

Model 74

Conceptual Diagram



Statistical Diagram



Conditional indirect effect of X on Y through $M_i = a_i\,(b_{1i} + c_{2i}'X)$
Conditional direct effect of $X = c_1' + c_{2i}'M$

Note: Model 74 allows up to 10 mediators operating in parallel. PROCESS does not produce a table of conditional direct effects for model 74. With only one mediator, use model 1 to generate the conditional direct effects, specifying *M* as *moderator*.