

A New Analytic Framework for Moderation Analysis — Moving Beyond Analytic Interactions

Wan Tang¹, Qin Yu¹, Paul Crits-Christoph² and Xin M. Tu¹

¹University of Rochester and ²University of Pennsylvania

Abstract: Conceptually, a moderator is a variable that modifies the effect of a predictor on a response. Analytically, a common approach as used in most moderation analyses is to add analytic interactions involving the predictor and moderator in the form of cross-variable products and test the significance of such terms. The narrow scope of such a procedure is inconsistent with the broader conceptual definition of moderation, leading to confusion in interpretation of study findings. In this paper, we develop a new approach to the analytic procedure that is consistent with the concept of moderation. The proposed framework defines moderation as a process that modifies an existing relationship between the predictor and the outcome, rather than simply a test of a predictor by moderator interaction. The approach is illustrated with data from a real study.

Key words: Analytic interactions, mixed-effects model, non-linear model, varying-coefficient regression model.

1. Introduction

Moderation and mediation analyses are widely used in biomedical and psychosocial research (Baron and Kenny, 1986; Chaplin, 1991; Cole and Maxwell, 2003; Crits-Christoph *et al.*, 2003; Holmbeck, 1997; Kraemer *et al.*, 2001; 2002; Krull and MacKinnon, 2001; Rogosch *et al.*, 1990; Rothman and Greenland, 1998). Although often implemented in correlational studies in social psychology and other fields of inquiry, moderation and mediation analyses have become increasingly popular and an integral part of data analysis in treatment research (Kraemer *et al.*, 2002). In intervention studies, moderation analysis helps determine whether an intervention has a differential effect among subgroups that are defined by baseline characteristics. Thus, moderators provide useful information for treatment decisions and maximizing treatment effect. In contrast to moderation analysis, mediation analysis helps identify mechanisms by which an intervention achieves its effect. By identifying the correct mediation process

through which treatment affects study outcomes, not only can we further our understanding of the pathology of the disease and treatment, but also provide information for developing new and alternative treatments to treat the disease with efficient use of resources. Moderation and mediation analyses are also often performed for epidemiologic studies to determine risk factors and elucidate the causes of a disease.

In a seminal paper, Baron and Kenny (1986) proposed a general framework for characterizing a moderation and mediation process. In particular, they laid a theoretical foundation for conceptualizing such processes and for approaching the underlying analytic problems. In addition, their work clarified the fundamental difference between the closely related, yet fundamentally distinct notion of moderation and mediation. However, as recently pointed out by Kraemer *et al.* (2001), the limited analytic strategies proposed in their paper have been used and extrapolated to situations to which they often do not apply, leading to confusion in interpreting analysis results and even conflict with the conceptual definition of such processes. For example, their work showed that the presence of analytic interaction between a moderator and a predictor (the product of the two variables) is model-dependent; the same data may show zero or non-zero moderator by predictor interactions depending on which analytic models (e.g., logistic or linear model) are used to fit the data. Thus, the popular approach of simply looking for non-zero interactions as used in most moderation analyses has limited applications and often leads to dubious and uninterpretable results. Defining a general analytic definition consistent with the conceptual notion of moderation is the focus of this paper.

In this paper we restrict our attention to moderation analysis and discuss a new approach to address the limitations of current methods. More specifically, our approach more broadly models the effect of a moderator so that its effect is not limited to analytic interactions. For convenience, we mainly focus on moderation analysis. We show how analytic interactions can become uninterpretable as moderation effect and how one variable can be a moderator without assuming the form of such interactions. After describing the new analytic framework in Section 2, we illustrate the proposed approach with a real data example in Section 3, followed by concluding remarks.

2. Moderation Analysis

In this section, we first review existing models for moderation analysis and in the process outline the problems with such methods. We then propose our approach to address these issues.

2.1 Modeling moderation with interaction terms

For convenience, assume a relatively simple moderation process involving only one predictor, x , a response, y , and a moderator, z . Assume that y is continuous and consider the following linear model relating x to y :

$$y_i = \alpha_0 + \alpha_1 x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n, \quad (2.1)$$

where i indexes the subjects from a sample of size n and $(0, \sigma^2)$ denotes a random variable with mean 0 and variance σ^2 . For robust inference, we only assume that ϵ_i has a mean of 0 and is uncorrelated with x . The latter is known as the pseudo-isolation condition, which enables one to establish the influence of x on y isolated from z in a causal relationship (e.g. Bollen, 1989).

A moderator is a variable that affects or modifies the relationship between x and y . In a conceptual sense, if z is a moderator, it interacts with the predictor x to alter the effect of the latter variable on the response y . Because of such an “interaction” interpretation, a popular approach as used in many moderation analyses is to include the first-order (xz) or even higher-order (e.g. x^2z) moderator by predictor interactions to examine moderation effect. For example, by including the first-order x by z interaction in (2.1), we obtain:

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 z_i x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n. \quad (2.2)$$

Under this model, the effect of x on y defined by α_1 in (2.1) has been altered and replaced by a function of z in the form of $\alpha_1 + \alpha_2 z_i$ (Aiken and West, 1991; Neter *et al.* 1990). Because of this, a moderator is also known as an effect modifier.

Although moderation does translate into analytic interactions in the case of (2.2), it is a fundamentally different concept. Indeed, simply interpreting moderation as analytic interactions can have serious ramifications. In particular, not all interactions will have the moderation interpretation. For example, consider the two scatter plots in Figure 1, in which the relationship between y and x is plotted for the two levels of a binary variable z as indicated by circles ($z_i = 0$) and squares ($z_i = 1$). The data in the left diagram is modeled by:

$$y_i = (\alpha_0 + \delta_0 z_i) + (\alpha_1 + \delta_1 z_i) x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n. \quad (2.3)$$

In this model, the interaction $z_i x_i$ changes the slopes of the linear relations between x and y at the two levels of z and thus modifies the initial linear relationship in (2.1) for the differential effect by z . Note that (2.3) has an extra term, $\delta_0 z_i$, and is thus different from (2.2). This additional main effect of z in (2.3) is used to accommodate the different intercepts corresponding to the two levels of z . The data in the second diagram is modeled by a quadratic model:

$$y_i = (\alpha_0 + \delta_0 z_i) + (\alpha_1 + \delta_1) x_i + \alpha_2 z_i x_i^2 + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n. \quad (2.4)$$

Although similar, the above is fundamentally different from (2.3) as a model for moderation. Unlike (2.3), this model does not have a moderation interpretation with respect to (2.1) since it does not merely modify the effect of x on y , but rather it postulates quite a different relationship between x and y by adding a quadratic term $z_i x_i^2$. Thus, z cannot be considered as a modifier for the linear relationship between y and x as initially modeled in (2.1).

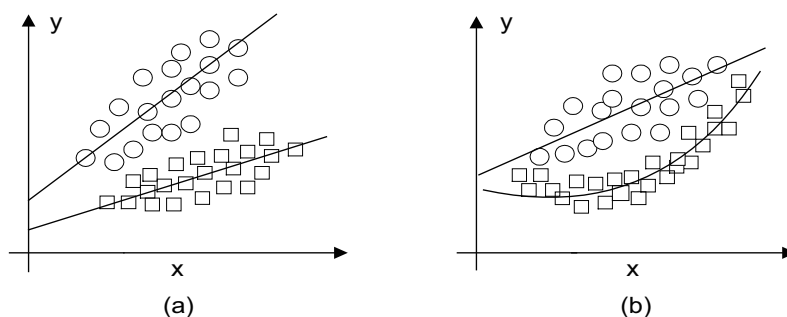


Figure 1: Two patterns of treatment response and fitted regression lines as a function of moderator z and treatment condition (circles for treatment 1 and squares for treatment 2).

An obvious difference between (2.3) and (2.4) is that the latter involves a higher-order interaction $z_i x_i^2$. However, a moderation model for (2.1) does not have to involve only the first-order interaction. For example, although the model below contains a higher-order interaction between x and z :

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 z_i x_i + \alpha_3 z_i^2 x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n. \quad (2.5)$$

It is still a moderation model for (2.1) since the inclusion of the interaction $z_i^2 x_i$ does not change the initial linear relationship between y and x . In comparison to (2.2), it modifies the effect of x on y through a slightly more complex quadratic form.

Note that although (2.4) is not a moderation model for (2.1), it may be viewed as such a model for a different quadratic relationship between y and x :

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n, \quad (2.6)$$

since z in (2.4) modifies the effect of x on y by altering the coefficients associated with the linear and quadratic terms. In general, any model we create by adding certain analytic interactions to (2.1) can be viewed as a moderation model for some model that relates y to x . However, a moderation model for (2.1) should only alter the effect of x on y without changing the original linear relationship between x on y . Thus, the essence of the definition of a moderator z is that

the model form remains the same, but the coefficients may change and become functions of z .

The examples above show that not all interactions have a moderation interpretation. The reverse is also true. Interaction in a conceptual sense has a broader interpretation, not just limited to analytic interactions in the form of cross-variable product. Consider for example a non-linear model relating x to y as given by:

$$y_i = \alpha_0(e^{-x_i})^{\alpha_1} + \alpha_2(e^{-x_i})^{\alpha_3} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq I \leq n. \quad (2.7)$$

This bi-exponential model, which is not a linear model since y is not a linear function of the coefficient α_1 , is widely used in modeling plasma concentration y as a function of time x in biomedical research (Neter *et al.* 1990; Davidian and Giltinan, 1995). Although non-linear, the effect of x on y is still defined by α_k , ($0 \leq k \leq 3$); if z alters the effect of x on y , it must do so through changes in these coefficients. For example, if α_2 is a linear function of z , it follows that:

$$y_i = \alpha_0(e^{-x_i})^{\alpha_1} + (\alpha_{20} + \alpha_{21}z)(e^{-x_i})^{\alpha_3} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq I \leq n. \quad (2.8)$$

The above model has the extra term $\alpha_{21}z(e^{-x_i})^{\alpha_3}$ in comparison to (2.7). Obviously, this term is not an analytic interaction in the form of xz . Further, it is not even possible to express it in the more general form of analytic interaction as $h(x)g(z)$ where $h(x)$ and $g(z)$ denote some functions of x and z , respectively. However, z still modifies the effect of x on y and thus conforms to the conceptual notion of moderation. As in the linear model case, if we literally add analytic interactions to (2.8), we may end up with models with no moderation interpretation. For example, if we simply add the analytic interaction xz to (2.8), we obtain:

$$y_i = \alpha_0(e^{-x_i})^{\alpha_1} + \alpha_2(e^{-x_i})^{\alpha_3} + \alpha_4 z_i x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n. \quad (2.9)$$

As with (2.4), the above actually represents a new model for the relationship between y and x , rather than a moderation model to account for the altered effect of x on y by z based on the original model in (2.5).

Note that the problem with interpreting conceptual interaction as simply analytic interaction involving cross-variable products has also been noted by Kraemer *et al.* (2001). By considering analytic interactions across different types of models (e.g. linear, logistic etc.), they demonstrated that the presence of such interaction effects depends on the type of models being fitted. Our considerations above complement their findings by further elucidating the mechanism that causes such model dependency when defining moderation through interactions.

2.2 A varying-coefficient model based general framework for moderation — moving beyond interactions

As illustrated in the preceding section, current methods for moderation analysis developed on the premise of analytic interactions are problematic. If used without caution, they may give rise to models that do not have moderation interpretation in the conceptual sense. In addition, such interaction-based strategies generally do not work for non-linear models, as interactions between a predictor and a moderator do not have to be in the form of cross-variable product. In this section, we systematically address these issues simultaneously by proposing an approach that does not rely on analytic interactions.

Let us start with the linear model (2.1) again. As we discussed earlier, this model is determined by the coefficients or parameters, α_0 and α_1 . Thus, if z interacts with x to alter the relationship between x and y , these parameters become a function of z , i.e.,

$$y_i = \alpha_0(z_i) + \alpha_1(z_i)x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq n. \quad (2.10)$$

Unlike the models in (2.2) and (2.3), no specific form is assumed for $\alpha_k(z)$ ($k = 0, 1$). Thus, it represents a general class of models for moderation effect derived based on the original model (2.1). For example, if we know a priori that

$$\alpha_0(z_i) = \alpha_0, \quad \alpha_1(z_i) = \gamma_1 + \gamma_2 z_i$$

then we immediately obtain the model in (2.2). The model in (2.10) automatically excludes models that contain analytic interactions but do not have a moderation interpretation such as the model in (2.4). The linear model in (2.10) with the coefficients being a function of a variable is known as the varying-coefficient linear model (Fan and Zhang, 1999; Hastie and Tibishirani, 1993).

Thus, our approach to defining the effect of a moderator z on the linear model (2.1) is to change the definition of the coefficients so that they become a function of z . This principle is readily applied to general models such as the generalized linear and non-linear models (McCullagh and Nelder, 1989; Davidian and Giltinan, 1995). For example, the generalized linear model for a binary response is expressed as:

$$y_i = \text{Bin}(h(\alpha_0 + \alpha_1 x_i), 1), \quad 1 \leq i \leq n, \quad (2.11)$$

where $\text{Bin}(p, 1)$ denotes a Binomial (or Bernoulli) distribution with sample size $n = 1$ and the probability of success p . Thus, in (2.11), the mean of y_i is modeled as a function of x_i : $h(\alpha_0 + \alpha_1 x)$. The most popular choice is the logit function,

$$h(\alpha_0 + \alpha_1 x_i) = \exp(\alpha_0 + \alpha_1 x_i) / [1 + \exp(\alpha_0 + \alpha_1 x_i)], \quad (2.12)$$

though other link functions such as the probit link are also often used (McCullagh and Nelder, 1989; Tu *et al.*, 1999). Once a link function is chosen, the relationship between y and x is determined by the parameters α_0 and α_1 . Thus, as in the linear model case, we define the effect of a moderator z by letting α_k be a function of z :

$$y_i \sim \text{Bin}(h)\alpha_0(z_i) + \alpha_1(z_i)x_i, 1 \leq i \leq n. \quad (2.13)$$

In the logistic model case, α_k in (2.12) becomes a function of z .

The definition also carries through in a straightforward fashion for non-linear models. For example, by modeling the coefficients as a function of z in (2.6), we obtain:

$$y_i = \alpha_0(z)(e^{-x})^{\alpha_1(z)} + \alpha_2(z)(e^{-x})^{\alpha_3(z)} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n.$$

As in the linear model case of (2.10), the above includes (2.8) as a special case, but excludes (2.9) as a model for moderation analysis.

Note that in semiparametric regression analysis, models are specified by the conditional mean of the response given the predictor (Robins *et al.* 1995):

$$E(y_i|x_i) = h(x_i, \alpha), \quad 1 \leq i \leq n, \quad (2.14)$$

where α is the vector of parameters or coefficients and $h(x, \alpha)$ is a function of x and α . When defined under the semi-parametric regression setup, a moderator z can affect only α , without altering the functional form $h(x, \alpha)$, i.e.,

$$E_z(y_i|x_i) = h(x_i, \alpha(z_i)), \quad 1 \leq i \leq n.$$

For example, by expressing (2.1) in the form (2.14), we obtain:

$$E(y_i|x_i) = \alpha_0 + \alpha_1 x_i, \quad 1 \leq i \leq n.$$

Thus, (2.2) and (2.3) are both moderation models for (2.1) since z alters only the parameter vector. However, (2.4) is not a moderation model for (2.1) since it also changes the functional form $h(x, \alpha)$.

2.3 Inference for varying-coefficient models

Procedures for fitting varying-coefficient models are based on the idea of “local averaging.” For example, for the linear varying-coefficient model in (2.10), first we fix z and use data close to z (window) to fit the model by treating z as a constant. By moving z over the range of z in the data, we obtain estimates of $\alpha_k(z)$ as a function of z . This approach will enable us to determine the appropriate form for as well as make inference about $\alpha_k(z)$ (e.g. Carroll *et al.*, 1998;

Fan and Zhang, 1999; Hart, 1997; Hastie and Tibishirani, 1993). To overcome the difficulty with varying degrees of sparseness and gaps in the distribution of z , methods have been developed to utilize all data in the estimation of $\alpha_k(z)$ by employing varying window size and weighted averaging (with more weights attached to observations closer to z). This so-called “kernel smoothing” approach produces smooth functions of $\alpha_k(z)$ that can be used to test whether $\alpha_k(z)$ is a function of z and to suggest appropriate functional form for their modeling. Since inference for general varying-coefficient models is an area of on-going methodological research, we will not pursue estimation for general varying-coefficient models in this paper. Rather, we discuss several special cases of the varying-coefficient linear model and illustrate how such specific models can be fitted using standard procedures.

Binary and Categorical Moderator

Since the case with a binary z can be subsumed into the discussion of a categorical moderator, we consider only a categorical z and assume that z has a total of K categories.

For such a moderator, (2.10) becomes:

$$y_{ki} = \alpha_{0k} + \alpha_{1k}x_{ki} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n_k, \quad 1 \leq k \leq K. \quad (2.15)$$

In this case, the original sample is partitioned into K sub-samples, each of size n_k , and a different linear relationship is postulated for each sub-sample as characterized by the different coefficients or parameters α_{0k} and α_{1k} ($1 \leq k \leq K$). Least squares or estimating equations can be used to estimate the parameters (McCullagh and Nelder, 1989). If z is truly a moderator, then at least two of the α_{1k} 's will be different. Thus, we can ascertain the mediation role of z by testing the following hypothesis:

$$H_0 : \alpha_{1k} = \alpha_1 \text{ for all } 1 \leq k \leq K \text{ vs. } H_a : \alpha_{1j} \neq \alpha_{1k}, \text{ for some } 1 \leq j, k \leq K. \quad (2.16)$$

Note that sometimes it may happen that z changes only the intercepts without affecting the slope, i.e., $\alpha_{1k} = \alpha_1$ for all $1 \leq k \leq K$. In this case, z becomes a covariate, since a moderator must change the effect of x on y .

Binary and categorical predictor

As in the preceding section, we consider only a categorical predictor x with K levels. Let n_k denote the sub-sample size for the k th level of x ($1 \leq k \leq K$). The varying-coefficient model in (2.10) reduces to:

$$y_{ki} = \alpha_k(z_{ki}) + \epsilon_{ki}, \quad \epsilon_{ki} \sim (0, \sigma^2), \quad 1 \leq i \leq n_k; \quad 1 \leq k \leq K. \quad (2.17)$$

As a special case, if $\alpha_k(z_{ki}) = \alpha_k$, we immediately obtain from (2.17) an analysis of variance model (ANOVA) with α_k interpreted as the cell mean of the k th subsample or group. Thus, the variable z in (2.17) can be viewed as modifying the cell or group means of an ANOVA.

Now, consider a linear $\alpha_k(z_{ki}) = \gamma_{0k} + \gamma_{1k}z_{ki}$, in which case (2.17) becomes:

$$y_{ki} = \gamma_{0k} + \gamma_{1k}z_{ki}, \quad \epsilon_{ki} \sim (0, \sigma^2), \quad 1 \leq i \leq n_k; \quad 1 \leq k \leq K. \quad (2.18)$$

In this case, the difference between the means of two groups, k and r , is given by:

$$\Delta_{kr} = (\gamma_{0k} - \gamma_{0r}) + (\gamma_{1k} - \gamma_{1r})z. \quad (2.19)$$

The second term in (2.19) represents the differential effect of z on the means of the two groups and constitutes the moderation effect. If the slopes, γ_{1k} (2.18) equal to a constant across all groups, i.e., $\gamma_{1k} \equiv \gamma_1$ ($1 \leq k \leq K$), then this differential effect will be zero and the model in (2.18) reduces to:

$$y_{ki} = \gamma_{0k} + \gamma_1 z_{ki} + \epsilon_{ki}, \quad \epsilon_{ki} \sim (0, \sigma^2), \quad 1 \leq i \leq n_k; \quad 1 \leq k \leq K. \quad (2.20)$$

The above is an analysis of covariance (ANCOVA) model and $\gamma_1 z_{ki}$ is the adjustment factor for the effect of z_{ki} on the mean response (e.g. Neter *et al.*, 1990). Unlike (2.18), z in (2.20) exerts the same effect on the mean response across all the groups. Thus, in (2.20), z still modifies the effect of x on y (in the form of group means defined by the levels of x), but it does so uniformly across the groups (or there is no x by z interaction).

In real study applications, one or more $\alpha_k(z)$ in (2.17) may be non-linear or other more complex functions. For example, if $K = 2$, (2.17) with $\alpha_k(z)$ given by:

$$\alpha_1(z) = \gamma_0, \quad \alpha_2(z) = \gamma_0 + \gamma_1 I_{[z \leq c]}, \quad (2.21)$$

can be used to model the scenario where the effect of the dichotomous variable x on y is through a step function defined by some cut-off c of the moderator z as depicted in Figure 2 of Baron and Kenny (1986), where $I_{[z \leq c]}$ denotes the set indicator with $I_{[z \leq c]} = 1$ if $z \leq c$ and 0 if otherwise. Note that the advantage of formulating the model using the vary-coefficient model (2.17) is that we can use smoothing techniques to estimate the functional form of $\alpha_k(z)$ so that it is not necessary to specify a priori the value of the cut-off c as in (2.20). We illustrate how such an approach works for a real study example in Section 3.

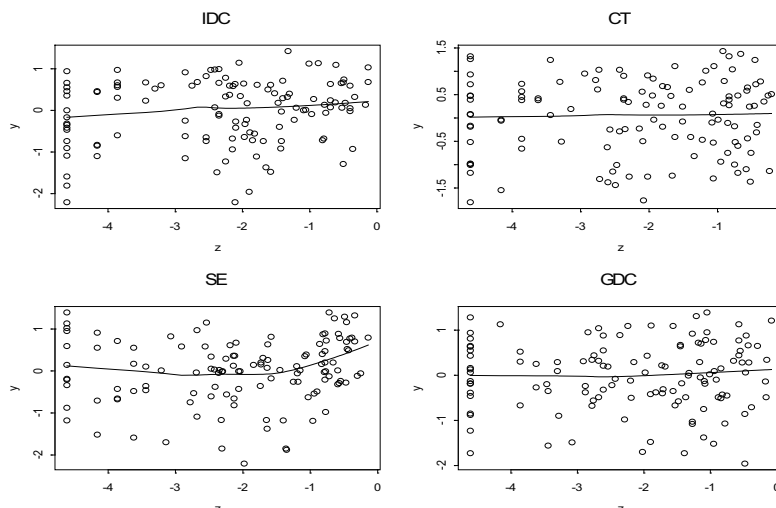


Figure 2: Scatter plots with fitted Lowess curves for each of the four treatment conditions.

Relationship to linear mixed-effects model

The linear varying-coefficient model (2.10) is also closely related to a linear mixed-effects (LMM) or hierarchical linear model (HLM) (Laird and Ware, 1982; Gibbons *et al.*, 1994; Raudenbush, 1994). In particular, the varying-coefficients, $\alpha_k(z)$, can be viewed as the mean of random individual coefficients given the value of the moderator at z . In this sense, we can derive the model in (2.10) from the perspective of this popular modeling framework.

As in the usual derivation of the linear mixed-effects model, at the first level, we assume a linear model with random individual effects as follows:

$$y_i = \alpha_{0i} + \alpha_{1i}x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2), \quad 1 \leq i \leq n, \quad (2.22)$$

In the above model, α_{0i} and α_{1i} are random variables and are assumed to be uncorrelated with ϵ_i ($k = 0, 1$). In the usual linear mixed-effects model, α_{0i} and α_{1i} are assumed to have a joint normal distribution and ϵ_i is also assumed to be normal. In (2.22), we do not assume such parametric distributions. At the second level, we model the conditional distribution of each α_{ki} given z_i as:

$$\alpha_{ki} = \alpha_k(z_i) + e_{ki}, \quad e_{ki} \sim (0, \sigma_k^2), \quad k = 0, 1, \quad (2.23)$$

where e_{ki} is assumed to have a mean of 0 and to be uncorrelated with both z_i and x_i . It follows from the assumption of α_{ki} that e_{ki} is also uncorrelated with ϵ_i for $k = 0, 1$. By combining the two models in (2.22) and (2.23), we immediately obtain the following mixed-effects model:

$$y_i = \alpha_0(z_i) + \alpha_1(z_i)x_i + \tilde{\epsilon}_i, \quad \tilde{\epsilon}_i \sim (0, \sigma_{mix}^2(x_i^2)), \quad (2.24)$$

where $\tilde{\epsilon}_i = e_{0i} + e_{1i}x_i + \epsilon_i$. As the random effects e_{ki} are not of interest, they have been combined with ϵ_i to form the model error in (2.24). It is readily shown that $\tilde{\epsilon}_i$ is uncorrelated with z_i and x_i , and thus satisfies the pseudo-isolation condition. Thus, (2.24) defines the same linear varying-coefficient model as (2.10), except that the error variance is a function of x^2 rather than a constant as in (2.10). Note that if $e_{1i} = 0$ in (2.23), we obtain the same model as in (2.10).

2.4 Longitudinal data analysis

The extension to longitudinal data analysis is straightforward. As before, we only consider a continuous response y with a predictor x and moderator z . For convenience, we consider modeling such a response using the linear mixed-effects model, as this approach is widely used in modeling longitudinal data (e.g. Laird and Ware, 1982; Gibbons *et al.*, 1994; Raudenbush, 1994).

Consider a longitudinal study with n subjects and m assessment points. For illustration purposes, we only consider linear growth-curve analysis in which the trajectory of each subject is modeled as a linear function of time as follows:

$$y_{it} = \alpha_0 + \alpha_1 x_i + \alpha_2 t + \alpha_3 t x_i + b_0 + b_1 x_i + b_2 t + b_3 t x_i + \epsilon_i, \quad 1 \leq i \leq n, \quad 1 \leq t \leq m, \quad (2.25)$$

where t denotes time, α_k the fixed-effects for the population mean, and b_k the random-effects to account for individual differences ($0 \leq k \leq 3$). As in the literature, we assume that ϵ_i follows a normal distribution with mean 0 and variance σ^2 , and b_k ($0 \leq k \leq 3$) follow a joint normal with mean 0 and variance Σ_b . We define moderation to be the effect of z on the fixed-effects α_k , i.e., $\alpha_k(z)$ is a function of z .

In most applications, interest lies in whether z moderates treatment differences. For example, suppose that x is a binary indicator for two treatment conditions. The vary-coefficient linear model in this case is given by:

$$\begin{aligned} \text{If } x_i = 0 : y_{it} &= \alpha_0(z_i) + \alpha_2(z_i)t + b_0 + b_2 t + \epsilon_i, \\ \text{If } x_i = 1 : y_{it} &= [\alpha_0(z_i) + \alpha_1(z_i)] + [\alpha_2(z_i) + \alpha_3(z_i)]t \\ &\quad + b_0 + (b_2 + b_3)t + \epsilon_i. \end{aligned} \quad (2.26)$$

In this model, $\alpha_0(z_i)$ moderates the within-treatment effect, $\alpha_2(z_i)$ the change over time due to the within-treatment effect, $\alpha_1(z_i)$ the between-treatment effect at baseline, and $\alpha_3(z_i)$ the change over time due to between-treatment difference.

If all the varying-coefficients are a linear function of z , $\alpha_k(z) = \gamma_{k0} + \gamma_{k1}z$, then (2.26) gives rise to the usual approach for modeling moderation effect by including analytic interactions involving x and z . In this case, (2.26) simplifies

to:

$$\begin{aligned}
 \text{If } x_i = 0 : y_{it} &= \gamma_{00} + \gamma_{01}z_i + \gamma_{20}t + \gamma_{21}z_it + b_0 + b_2t + \epsilon_i, \\
 \text{If } x_i = 1 : y_{it} &= (\gamma_{00} + \gamma_{10}) + (\gamma_{01} + \gamma_{11})z_i + (\gamma_{20} + \gamma_{30})t \\
 &\quad + (\gamma_{21} + \gamma_{31})z_it + b_0 + (b_2 + b_3)t + \epsilon_i.
 \end{aligned} \tag{2.27}$$

In most randomized trials, the mean response does not differ between treatment conditions at baseline so that $\gamma_{10} = 0$. In addition, if randomization works effectively, z should not have a differential effect at baseline, which implies that $\gamma_{11} = 0$. So, (2.27) further simplifies to:

$$\begin{aligned}
 \text{If } x_i = 0 : y_{it} &= \gamma_{00} + \gamma_{01}z_i + \gamma_{20}t + \gamma_{21}z_it + b_0 + b_2t + \epsilon_i, \\
 \text{If } x_i = 1 : y_{it} &= \gamma_{00} + \gamma_{01}z_i + (\gamma_{20} + \gamma_{30})t + (\gamma_{21} + \gamma_{31})z_it \\
 &\quad + b_0 + (b_2 + b_3)t + \epsilon_i.
 \end{aligned} \tag{2.28}$$

In the above model, the treatment by time interaction, γ_{30} , represents treatment difference over time in the absence of the moderator z (when $z = 0$), while the treatment by time by moderator interaction, γ_{31} , represents the moderation effect of z on the treatment difference. The model in (2.28) and its generalizations for multiple treatment conditions are widely used in testing moderation effect in longitudinal studies.

As in the case of cross-sectional study designs, inference for varying-coefficient models can still be made using the usual estimation procedures when z or x or both are categorical variables. For example, we can use standard estimation procedures to fit the model in (2.28). When both z and x are continuous, inference becomes much more complex and smoothing methods may be used. Again, this issue will not be pursued here. Fortunately, in many randomized studies, treatment differences are modeled by binary indicators, in which case standard procedures can be used to fit linear mixed-effects models with varying-coefficients.

3. A Real Study Data Application

We illustrate the proposed methodology with real study data from the National Institute on Drug Abuse Collaborative Cocaine Treatment Study (Crits-Christoph *et al.*, 1999). This randomized and multi-center project investigated the efficacy of psychosocial treatment for cocaine dependence, with a sample of 487 patients who were randomized to one of four treatment conditions: cognitive therapy (CT) plus group drug counseling (GDC), supportive-expressive (SE) plus group drug counseling (GDC), individual drug counseling (IDC) plus group drug counseling (GDC), and GDC alone. Primary outcome analyses focused on the intent-to-treat sample and examined several measures of drug use (Crits-Christoph *et al.*, 1999).

For illustration purposes, we applied the proposed methodology to data at six month post-treatment using the Addiction Severity Drug Use composite variable (ASI; McLellan *et al.*, 1992) as the response variable. As a significant treatment difference was found among the four treatment groups, it was of interest to examine if the treatment differences were moderated by baseline alcohol consumption as measured by the ASI alcohol use composite.

Let y denote the drug use composite variable at six month post-treatment and z the pre-treatment alcohol use composite variable. We applied the ANOVA model with varying coefficients in (2.17) to examine the effect of moderation by z . To determine the appropriate analytic form for modeling the mean response of each group $\alpha_k(z_{ki})$ ($1 \leq k \leq 4$), we applied a smoothing procedure, LOWESS (locally weighted scatter plot smoother), to data from each treatment condition (e.g. Fox 2000; Loader 1999). Shown in Figure 2 are the scatter plots together with the fitted LOWESS curves for each of the four treatment groups.

Table 1: Estimates of model parameters, standard errors and p-values for the varying coefficient ANOVA model (20) with linear and quadratic mean response. The γ_{k0} represent the main effect of group k , γ_1 and γ_2 are the coefficients for the first- and second-order main effect of moderator z , δ 's denote coefficients for the first-order moderator by group interactions and η 's denote the second-order moderator by group interactions.

Quadratic								
Coefficient	γ_{10}	γ_{20}	γ_{30}	γ_{40}	γ_1	γ_2		
Estimate	0.4594	0.4049	0.7076	0.4274	0.3743	0.0599		
Std Errors	0.1543	0.1503	0.1475	0.2137	0.2085	0.0405		
p-value	0.8357	0.8810	0.0582	0.0461	0.0733	0.1399		
Coefficient	δ_1	δ_2	δ_3	η_{11}	η_{22}	η_{33}		
Estimate	-0.0258	-0.0554	0.2978	-0.0129	-0.0116	0.0571		
Std Errors	0.1465	0.1482	0.1459	0.0280	0.0286	0.0286		
p-value	0.8600	0.7090	0.0417	0.6448	0.6852	0.0464		
Linear								
Coefficient	γ_{10}	γ_{20}	γ_{30}	γ_{40}	γ_1	δ_1	δ_2	δ_3
Estimate	0.2844	0.2249	0.2588	0.1973	0.0874	0.0395	0.0036	0.0246
Std Errors	0.0933	0.0910	0.0904	0.1300	0.0503	0.0354	0.0348	0.0357
p-value	0.3509	0.7620	0.4965	0.1298	0.0832	0.2654	0.9185	0.4901

The fitted LOWESS curves indicated a quadratic mean response for the SE group, but a linear response for each of the other three groups. Thus, to formally test for moderation by z , we fitted the following quadratic response model:

$$y_{ki} = \gamma_{k0} + \gamma_1 z_{ki} + \gamma_2 z_{ki}^2 + \sum_{l=1}^3 \delta_l z_{ki} I_{kil} + \sum_{l=1}^3 \eta_l z_{ki}^2 I_{kil} + \epsilon_{ki}, \quad 1 \leq k \leq 4, \quad (3.1)$$

where γ 's, δ 's and η 's are model parameters, and $k = 1, 2, 3, 4$ denote the IDC, CT, SE and GDC treatment groups, respectively. For robust inference, we did not assume normality for and estimated the parameters using estimating equations or quasi-likelihood (e.g. McCullagh and Nelder, 1989).

Shown in Table 1 are the estimated parameters and the associated p -values. The estimated coefficients for the first- and second-order interactions are statistically significant only for the SE group (see estimates of δ_2 and η_3 and their associated p -values), indicating that unlike the other groups, the response for SE had a quadratic relationship with the moderator. Thus, pre-treatment alcohol use was a moderator, as it affects treatment response differentially between this and the other three treatment conditions.

It is interesting to note that when we applied the linear coefficient model (2.18) without the second-order interactions, none of the coefficients were significantly different from 0 (see estimates and associated p -values in Table 1). Thus, by looking only at the first-order interactions as in the traditional way, we would not be able to detect any moderation effect in this case. In this particular application, the use of the varying-coefficient model (2.17) helped identify the correct analytic interactions to model the effect of moderation by z .

4. Discussion

In this paper, we discussed a general analytic framework for moderation analysis by defining moderation as a process that modifies an existing relationship between the predictor and outcome. As illustrated by both theoretical considerations and real data analyses, moderation can follow quite a complex process, which may not be modeled by simply including analytic interactions involving the moderator and predictor as in most moderation analyses. Since the relationship between the response and predictor is defined by the coefficients or parameters of a given model, it is logical to define the effect of a moderator through such model parameters. Thus, the proposed approach is consistent with the conceptual definition of a moderation process.

Although moderation effect often exhibits in the form of analytic interaction, especially for linear regression models, not all such interactions can be interpreted as moderation effect. By defining moderation effect using the varying-coefficient model, we are able to delineate the types of analytic interaction that have a moderation interpretation from those that do not. Also, since moderation models are defined based on the original model relating the response and predictor, they are consistent and well-interpreted. Thus, the model-dependent issue as pointed out in Kraemer *et al.* (2001) does not arise. For example, if the original relationship between the response and predictor is a linear model, the effect of a moderator is limited to modifying the coefficients of the linear model, ruling out other types

of models, such as the logistic model, as potential candidates for moderation analysis.

Since our goal in this paper was to present an appropriate analytic framework for moderation analysis, we did not get into technical details about inference for general varying-coefficient regression models. When there are multiple continuous predictors and moderators, inference for such models may become quite complex, especially with longitudinal study data. We will address these issues in future research.

Acknowledgment

We would like to thank an anonymous reviewer for helpful and constructive comments that led to substantial improvement in the presentation of the research. This research was funded in part by National Institute of Mental Health grant P50-MH-45178, U01-DA07090, P30-MH-45178, U01-DA07663, U01-DA07673, U01-DA07693, U01-DA07085, and R01-DA012249.

References

- Aiken, L. S. and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Sage.
- Baron, R. M. and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: concept, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173-1182.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley.
- Carroll, R. J., Ruppert, D. and Welsh, A. H. (1998). Local estimating equations. *Journal of the American Statistical Association* **93**, 214-227.
- Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality* **59**, 143-178.
- Cole, D. A. and Maxwell, S. E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology* **4**, 558-577.
- Crits-Christoph, P., Siqueland, L., Blaine, J., Frank, A., Luborsky, L., Onken, L.S., et al. (1999). Psychosocial treatments for cocaine dependence: National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Archives of General Psychiatry* **56**, 493-502.
- Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall.
- Fan, J. and Zhang, W. (1999). Statistical estimation in varying-coefficient models. *Annals of . Statistics* **27**, 1491-1518.

- Fox, J. (2000). *Nonparametric Simple Regression: Smoothing Scatterplots*. Sage.
- Gibbons, R. D., Hedeker, D., Charles, S. C. and Frisch, P.R. (1994). A random-effects probit model for predicting medical malpractice claims. *Journal of American Statistical Association* **89**, 760-767.
- Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-fit Tests*. Springer-Verlag.
- Hastie, T. J. and Tibishirani, R. J. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society. Ser. B* **55**, 757-796.
- Holmbeck, G. N. (1997). Toward terminological, conceptual and statistical clarity in the study of mediators and moderators: examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology* **65**, 599-610.
- Kraemer, H. C., Stice, E., Kazdin, A., Offord, D. and Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry* **158**, 848-856.
- Kraemer, H. C., Wilson, T., Fairburn, C. G. and Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* **59**, 877-883.
- Krull, J. L. and MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research* **36**, 249-277.
- Laird, N. and Ware, J. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963-974.
- Loader, C. (1999). *Local Regression and Likelihood*. Springer.
- McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, 2nd.ed. Chapman and Hall.
- McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., *et al.* (1992). The fifth edition of the addiction severity index. *Journal of Substance Abuse Treatment* **9**, 199-213.
- Neter, J., Wasserman, W. and Kutner, M. H. (1990). *Applied Statistical Models*, 3rd. ed. Irwin.
- Raudenbush, S. W. (1994). Random effects models. In (*Eds.*), *The Handbook of Research Synthesis* (Edited by H. Cooper and L. V. Hedges), Russell Sage Foundations.
- Robins, J. M. and Rotnitzky, A. and Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association* **90**, 106-121.
- Rogosch, F., Chassin, L. and Sher, K. J. (1990). Personality variables as mediators and moderators of family history risk for alcoholism: conceptual and methodological issues. *Journal of Studies on Alcohol* **51**, 310-318.
- Rothman, K. J. and Greenland, S. (1998). *Modern Epidemiology*. Lippincott Williams and Wilkins.

Tu, X. M., Kowalski, J. and Jia, G. (1999). Bayesian analysis of prevalence with covariates using simulation-based techniques: applications to HIV screening. *Statistics in Medicine* **18**, 3059-3073.

Received August 23, 2007; accepted November 29, 2007.

Wan Tang
Department of Biostatistics and Computational Biology
University of Rochester
Rochester, NY 14642, USA
wan.tang@urmc.rochester.edu

Qin Yu
Department of Biostatistics and Computational Biology
University of Rochester
Rochester, NY 14642, USA
qin_yu@urmc.rochester.edu

Paul Crits-Christoph
Department of Psychiatry
University of Pennsylvania
Philadelphia, PA 19104, USA
crits@mail.med.upenn.edu

Xin M. Tu
Department of Biostatistics and Computational Biology and Department of Psychiatry
University of Rochester
Rochester, NY 14642, USA
xin.tu@urmc.rochester.edu