Longitudinal Regression Approaches

Causality

Although regression models describe a predictive relationship in which we must choose an "independent" and "dependent" variable, concluding that there is a relationship between them is not a confirmation of a causal relationship. Mill (1843) popularized the idea that an association between X and Y is causal when X produces Y regularly, known as *ceteris paribus* (when all things are equal). The general concept is isolation of the independent variable from all other potential causes of Y. This is the basic idea of the modern experiment, either through random assignment to groups or other methods of isolation (e.g., fully controlled manipulation of the independent variable, in say, a chemistry experiment). Kenny (1979) provides a valuable, brief summary of conditions for deciding a causal relationship exists that include: a) there is a relationship between X and Y; b) temporal precedence (X precedes Y); and b) nonspuriousness (third variables can be ruled out).[1] Your text (Cohen, Cohen, Aiken, & West, 2003, p. 455) adds that a "theoretically plausible mechanism" must exist. We can establish a relationship with correlation or regression analysis and we have been discussing how multiple regression can be used to try to statistically control for confounding relationships, which is one method of attempting to establish nonspuriousness, but we have not discussed establishment of temporal precedence. Without an experiment, the best way to attempt to establish temporal precedence is by collecting data over time—a longitudinal study.

Two Ways of "Predicting Change" over Two Waves

To investigate causal precedence, researchers are often interested in exploring whether some predictor, which can involve an intervention variable or any measured variable, is related to change in some outcome that is measured at more than one time point. The focus here will be on two waves of data (e.g., pretest-posttest, Year 1 to Year 2), but there are many other possible designs with more time points and methods for analyzing them (see Newsom, Jones, & Hofer, 2012). With two waves, there are primarily two types of regression models used for exploring predictors of change over time. Unfortunately, the two primary methods do not always lead to the same conclusion and there is much confusion over which method is "best" and how to interpret the results. There is good reason for the confusion, because there are many subtleties and there has been enormous debate about the superiority of one method over the other (e.g., see Campbell & Kenny, 1999; MacKinnon, 2008; Finkel, 1995 for reviews). I will not try to resolve the debate, but I want to provide some clarifying information.

Difference Scores

One basic regression model is to compute a difference score for the dependent variable, by subtracting the earlier time point from the later time point, $Y_{2\text{-}1} = Y_2 - Y_1$, and then use another variable, X, usually measured at the earlier time point, to predict change in the dependent variable.

$$Y_{2-1} = B_0 + B_1 X + e$$

Difference scores are sometimes also referred to as "gain scores" or "change scores." The difference score captures whether a particular case has increased or decreased over the interval (e.g., one year), and the slope, B_1, represents the amount of change in Y over each unit increment in X. If the predictor is binary, this simple regression is the same as the mixed between and within subjects ANOVA. It is also reduced form of a growth curve model with just two time points. Growth curve models estimate the average difference across any two points in a longitudinal design with several time points.

Lagged Regression

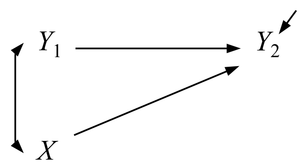
An alternative is to predict the dependent variable at the second time point, Y_2, using both the predictor X, measured at the prior time point, and the dependent variable, Y_1, also measured at the prior time point.

$$Y_2 = B_0 + B_1 X + B_2 Y_1 + e$$

[1] Kenny's is a less formal, and less complete, summary of criteria for causality than recent, more detailed discussions (see, for example, Pearl, 2009; Russo, 2010).

The concept is to take into account (control for) any pre-existing differences on Y_1, so it is sometimes referred to as the ANCOVA approach (and Cohen et al., 2003 refer to as “regressed” change). If X was a binary predictor, representing treatment and control conditions, the model would be exactly the same as an ANCOVA.

We can represent this model with a path diagram.



Y_1 is the dependent variable measured at Time 1, Y_2 is the (exact same) dependent variable measured at Time 2, and X_1 is some predictor (explanatory) variable measured at Time 1. The double-headed arrow between X and Y_1 simply indicates that the correlation between the two variables is taken into account as with any multiple regression model. The path between the two measurements of the dependent variable represents the extent to which the two measurements are correlated. If there is a perfect relationship, the dependent variable is highly stable over time. The small arrow to the right pointing down into Y_2 represents the residual variance or the variance unaccounted for in Y_2, and, ignoring X for a minute, that path could be interpreted as what is not stable or what is changing in the dependent variable. It then follows that another interpretation of this model (beyond controlling for initial differences) is that X_1 is predicting the remaining variance or the change in Y over time.

Lord's Paradox
Given the latter interpretation of the lagged regression model, one would think it is providing the same information as the difference score model—both models attempt to predict change in Y with X. It is sometimes disconcerting to discover, however, that the two approaches do not always lead to the same result (just as with ANCOVA and mixed ANOVA). Sometimes, X will be a significant predictor of "change" using one of the regression approaches but not the other (although, often they will show the same result). Such a lack of correspondence in results from the two longitudinal regression approaches is referred to as *Lord's paradox* (Lord, 1967). The explanation for this difference in the two regression models is complex and often difficult to fully grasp (Campbell & Kenny, 1999), but it comes down to two different concepts of stability and change.

Two Definitions of Stability and Change
The difference score concept of perfect stability is that the scores for each individual at Time 1 are exactly equal to their values at Time 2. In other words, $Y_{2i} = Y_{1i}$ and $Y_{2i} - Y_{1i} = 0$. Change is any difference in values of Y_2 and Y_1 for an individual, and, if the score goes up or down, it is not perfectly stable over time. In the lagged regression model, stability is defined as a perfect correlation (or *autocorrelation*) and perfect regression between Y_2 and Y_1. The autocorrelation will not be (or will be minimally) affected if individual values change from Time 1 to Time 2, as long as they maintain the same rank in the data set at each time point. Change in the lagged regression approach, then, reflects the extent to which the autocorrelation is less than 1.0. In fact, we can see the connection between the two modeling approaches algebraically. If the difference score regression is rewritten slightly as,

$$Y_2 - Y_1 = B_0 + B_1 X_1 + e$$

and, if we rearrange the terms a bit, we get:

$$Y_2 = B_0 + B_1 X_1 + (1)Y_1 + e$$

And this second equation suggests that the difference score regression implies (or assumes) that the autoregression for Y_2 predicted by Y_1 is perfect (i.e., B_2 from the lagged regression equals 1).[2]

The scores may change over time without necessarily affecting the autocorrelation. For example, say you add 5 points to each score in the data set at Time 2. The mean of Y_2 will be increased by 5 points, but the correlation between Y_1 and Y_2 will not change. The average difference and every individual difference score will increase by 5 points, however. This example suggests one other conceptual difference between the two definitions of stability and change. The difference score regression concerns predicting level change in the dependent variable, whereas lagged regression concerns predicting relative change in the dependent variable (for more elaborate discussion of the contrast between the two approaches to change and stability, see Newsom, 2024, Chapter 4).[3]

Summary

I am often asked which regression approach to the analysis of change should be used or which is "correct." My answer is that neither is *incorrect*. The two analysis approaches ask slightly different questions. It is certainly wrong, however, to choose one approach over the other simply because one is significant and one is not or one shows results consistent with what you would like to find. Because the difference score model concerns absolute level change in the dependent variable, it asks the question: "Whose score is most likely to increase or decrease over time?" A positive association between X and the difference score, for example, tells us that individuals with a higher score on X tend to increase more over time (or that those with lower values on X increase less or decrease more). The difference score regression does not address pre-existing differences in the dependent variable related to X, however, because the initial values on Y_1 (and their relation to X) are not statistically controlled in the model. Lagged regression, on the other hand, because it does control for initial values of the dependent variable, is better suited for addressing the question "Is X a likely cause of Y?" (Campbell & Kenny, 1999), but it is less well-suited for describing who increases or decreases over time. The lagged regression gets at the causal precedence question, and combined with statistical control of confounding relationships, we are addressing all of the three criteria for causal inference, at least to some extent. I should note that, even with longitudinal data, lagged regression, and control of covariates, it is still standard practice to avoid conclusive causal language in describing and interpreting the results. We would need to have perfectly measured the covariates and have included all possible confounders as covariates in the model.[4]

References

Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. New York: Guilford.
Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). Applied multiple regression/correlation analysis in the behavioral sciences (Third Edition). Mahwah, NJ: Erlbaum.
Finkel, S. (1995). *Causal Analysis with Panel Data.* Thousand Oaks, CA: Sage.
Kenny, D. A. (1979). *Correlation and causality.* New York: Wiley.
Lord, E. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin, 68*, 304–305.
MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis.* New York: Erlbaum.
Newsom, J.T. (2024). *Longitudinal structural equation modeling: A comprehensive introduction, second edition.* Routledge.
Newsom, J.T., Jones, R.N., & Hofer, S.M. (Eds.) (2012). *Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Science*. New York: Routledge.
Pearl, J. (2009). Causality: Models, reasoning, and inference, second edition. Cambridge, UK: Cambridge University Press.
Russo, F. (2010). *Causality and causal modelling in the social sciences*. New York: Springer.

---

[2] The terminology of "perfect" assumes that the variances of Y_1 and Y_2 are equal. They will likely be at least very similar or close to equal but that does not always hold, adding some possible subtle exceptions or complications the strict interpretation as "perfect" stability.

[3] There might be an impulse to both predict difference scores and control for the dependent variable measured at the first time point (Y_1) in the same model, but the result is statistically equivalent to the lagged regression model and provides no new information. By substituting the (1)Y1 equivalence shown in the last equation into the first (lagged regression) equation, one can show algebraically (e.g., Finkel, 1995) by subtracting this value from both sides that the end result is the same except that the autoregression coefficient (B_2) is just equal to the original coefficient minus 1, B_1 (prediction of change) provides no new information and B_2 (autoregression) may be of opposite sign:

$$Y_2 = B_0 + B_1X + B_2Y_1 + e$$
$$Y_2 - (1)Y_1 = B_0 + B_1X + B_2Y_1 - (1)Y_1 + e$$
$$Y_2 - Y_1 = B_0 + B_1X + (B_2 - 1)Y_1 + e$$

[4] There are several other things that are important, of course. Perhaps especially, the length of time between measurements should be appropriate for the causal hypothesis. As with other circumstances, absence of an effect is not very good proof of the lack of causality, because statistical power may not be sufficient or there may be other study problems. Moreover, finding an effect in one population or in one context does not necessarily generalize to all populations or all contexts.