

Editorial

Motivating, testing, and publishing curvilinear effects in management research

Klaus E. Meyer

Senior Editor

Forthcoming:

Asia Pacific Journal of Management, 2009

Vol. 26, no. 3

Helpful comments by Sumon Bhaumik, C.S. Wong and Editor-in-Chief Mike Peng are greatly appreciated

K. E. Meyer

Full address: School of Management, University of Bath, Claverton Down, Bath, BA2 7AY, U.K.

e-mail: k.meyer@bath.ac.uk, web: www.klausmeyer.co.uk

Editorial

Motivating, testing, and publishing curvilinear effects in management research

Abstract

Curvilinear relationships are a powerful tool to conceptually express and empirically test complex relationships, and thus to advance understanding of the underlying theories and their implications for management practice. Yet, my practical experiences as an editor show that non-linear functions need to be handled carefully in ways that many authors do not seem to be aware of. This editorial thus discusses opportunities and tripwires of theoretical and empirical studies using curvilinear effects, with the aim to enhance the quality of manuscripts submitted for consideration at *APJM*.

Many papers in *APJM* empirically test relationships derived from theory. In our editorial practice, we frequently encounter submission that either could have greatly enhanced impact if they explored curvilinear relationships, or that use curvilinear effects but with substantial shortcomings in either their theoretical treatment, or their interpretation of the results. This editorial aims to outline the opportunities and tripwires of empirical analyses using curvilinear effects, and thus to enhance the quality of manuscripts submitted for consideration at *APJM*.

The relationships are usually assumed to be uniform over the range of the explanatory variable, such that they can be tested with a linear regression coefficient. Yet, this is only one of many possibilities; effects may interact with countervailing effects, or be limited to a particular range of a variable. For example, increasing employee ownership may enhance corporate performance due to better employee motivation, yet it may also cause corporate governance conflicts that harm corporate performance. In such situations, hypothesizing about an aggregate effect with a curvilinear form may provide a powerful tool to conceptually express and empirically test complex relationships, and thus to advance understanding of the underlying theories and their implications for management practice. Yet, non-linear functions need to be handled carefully—both in their theoretical development and in the interpretation of empirical results. Unconvincing theoretical motivation or inadequate treatment of the empirical results can lead authors to spend a lot of time on avenues that prove ultimately fruitless.

Theoretical foundations

Theoretical considerations lead to hypotheses that suggest an association between two (or more) variables, yet theory rarely suggests the precise functional form. In some cases, authors may be able to derive a curvilinear effect from curvilinearity of an underlying relationship that has previously been established in the literature. For example, Takeuchi, Lepak, Marinova and Yun (2006) start out from the non-linear nature of individual learning to argue that expatriate spouses' perceived novelty is curvilinearly related to their adjustment (conceptualized as a learning outcome). In such cases, the development requires careful transposition to the variables of the study, but is comparatively straight forward.

***** Insert Table 1 here *****

If such a clear theoretical foundation is not available, a curvilinear relationship between two variables may arise from the interaction of countervailing effects. Table 1 provides an overview of possible outcomes when a positive effect A interacts with a negative effect B (Table 1). While a curvilinear effect is one possibility, it is not the only one. If the nature of the interaction is not known, then theory does not lead directly to statements regarding the functional form of the relationship. If the opposing effects are of similar strength, which may be a reasonable default assumption, an empirical study would yield an “insignificant” result (Table 1, case 1).

Cases 2 and 3 illustrate the most common practice of hypothesizing. Management scholars typically focus their theoretical development on the effect that they expect to dominate, and thus hypothesize a positive or negative effect. In the interest of clarity of

the theoretical exposition, possible countervailing effects are relegated to the discussion section (if there is no possibility of a contrary effect then the relationship is probably tautological and not worth testing). Following this established practice eases communication with reviewers, but it should not be seen as a dogma as other ways of structuring a paper are common in other social sciences.

If one effect dominates at lower levels of the explanatory variable, while the other dominates at higher levels, then this suggests a quadratic relationship as illustrated in cases 4 and 5 of Table 1. Hypotheses proposing such a quadratic relationship, also called U-shaped or inverse-U-shaped, require an argument to explain why the author(s) expect one effect to dominate over a certain range of the explanatory variable. In my experience, it is rarely possible to exclude the possibility of an alternative functional form. However, a careful discussion of the nature of the opposing effects at different stages of the explanatory variables, possibly with a graphical illustration, as in Chan, Makino, and Isobe (2006) and Meyer and Sinani (2009), can make the argument much more convincing.

In some relationships, the relative strength of opposing effects may vary several times over the full range of an explanatory variable, which would suggest higher order functional forms, such as cubic (cases 6 and 7). These types of hypotheses need even more careful treatment as two turning points over the range of the explanatory variable need to be justified. The difficulty of doing so is highlighted by a recent controversy between Contractor (2007) and Hennart (2007) for the relationship between multinationality and performance.

Quadratic and cubic functions have been popular in recent management research. Yet, in many situations it may be more appropriate to test for a positive or negative effect with a declining rate (cases 8 and 9). This can empirically be tested with a logarithmic regression specification, though it may also result in significant results in a quadratic expression.

In other cases, the explanatory variable may go through qualitatively different stages that shift at certain benchmarks. For instance, equity ownership may be qualitatively different at 49.9% and 51.1%. In such cases, a step function may be more appropriate (case 10). Thus, the variable should be replaced by dummies that represent the distinct categories, or a linear effect may be combined with a dummy for all observations above a certain threshold (as in Munich, Svejnar, & Terrell 2005).

Let me illustrate these ideas by the example of employee ownership. A study of the motivation effect may empirically test the suggested positive effect on a sample of firms with small stakes of up to 5% of equity held by employees. On the other hand, a study of corporate governance conflicts may test the suggested negative effect in a sample of firms with, say, at least 30% of employee ownership. If authors have a sample that covers a broad range of firms, they may combine the effects and hypothesize a curvilinear relationship with an inverse-U shaped form. Other studies may investigate the regulatory framework and suggest qualitative changes once the employee-ownership passes certain thresholds. In this case, a step-function may be most suitable to capture the relationship suggested by theory.

Do hypotheses have to specify the exact functional form of the relationship? In my opinion, the answer is “not necessarily.” If the curvilinear effect is central to the

contribution of the paper, then an exact theoretical treatment is clearly warranted. Naturally, a clearly predicted and empirically supported effect provides more powerful evidence than a curvilinear effect motivated only by the existence of opposing effects. Thus, if the curvilinear effect is central to the paper's overall contribution, this approach would be preferred.

In other situations, however, it may be appropriate to stipulate a curvilinear form without hypothesizing its exact form. For example, countervailing effects may emerge for one of several focal variables of a theoretical framework, without offering clear suggestions which effects would dominate at what level. Or, authors may test for a curvilinear form when reinvestigating a relationship of high theoretical relevance, yet with contradictory *prior* empirical evidence. Rather than to hypothesize the exact nature of the curvilinearity based on speculative arguments, it is possible to align the hypotheses closely to the theoretical arguments, and then to investigate alternative specification of the functional form as part of the empirical analysis. In particular, when a predicted linear effect turns out to be insignificant, supplementary analyses may investigate curvilinear specifications and thus confirm that the expected effect does hold albeit only for a certain range of the variable (see Wong & Campion, 1991). Possibly explanations of the functional form can then be discussed with the interpretation of the results.

Curvilinear models provide powerful explanations, yet, authors should also reflect over the possibility that non-linearities in the data may be caused, for instance, by different effects in different subsets of the sample or by moderating effects. Hypothesizing and testing such effects may lead to insights of greater theoretical impact.

Interpreting quadratic relationships

When I receive a paper with a quadratic relationship for review, I usually first do a simple back-of-the-envelope calculation. Surprisingly often, this test leads me to reject the empirical validity of the findings. The test goes like this: Obtain the turning point (maximum or minimum) point of the estimated curve by taking first derivatives, and then check where this point is relative to the range of the explanatory variable.¹ Surprisingly often, this point is beyond two standard deviations from the mean, or even beyond the range for which the variable is defined (or for which the authors have observations).

The interpretation of curvilinear effects obtained in a regression analysis needs great care. Usually, it helps to plot the estimated curve graphically, and to calculate the respective turning points. The explanatory variable should be depicted over the relevant range, and reporting the mean, standard deviations and variable range. In the case of abstract constructs it may often be more illustrative to use a scale from two standard deviations below the mean to two standard deviations above the mean.

If the maximum of a curve is outside the relevant range of the variable, the relationship does not show an inverse-U-shaped (as might be suggested by authors failing to do this test), but is “positive at a declining rate,” and a logarithmic specification may provide a better fit. If the turning point is within the relevant range but far from the mean, then great care is required in the interpretation. First, such results are often influenced by single outliers (or a small number of outliers). This problem can be addressed by running the regression again with these outliers excluded, and considering both results in the

¹ This is an application of algebra. If the estimation yields a coefficient β_1 on a linear effect and β_2 on the corresponding quadratic effect, this corresponds to a function of the form $\beta_1x + \beta_2x^2$. Taking the first derivative and setting it to zero yields $\beta_1 + 2\beta_2x = 0$, and thus the turning point at $x = (-\beta_1/2\beta_2)$.

interpretation. Second, visual inspection of the curve often suggest a relatively flat level at higher levels, such that the appropriate interpretation may a positive effect at low levels, and an in-substantive effect at high levels. Again, a logarithmic function may provide a better fit. Either way, authors discussing a negative effect that turns out to be (almost) outside the relevant range appear not to understand what they are actually doing when adding the x^2 element, which can lead to very negative responses from reviewers.

Figure 1 illustrates as an example the simplest curvilinear (inverse-U) relationship, namely $\beta_1x + \beta_2x^2$ with $\beta_1=1$ and $\beta_2 = -1$, which can be depicted as an inverse-U shaped relationship with an optimum at 50% (Panel 1, Figure 1). However, closer inspection of the data may lead to a different interpretation. In a recent case, an author had few observations over 40%, and the largest value near 55%. Thus, statements this author made on the negative branch of the curve implied an out-of-sample extrapolation. This is normally not justified. Incorporation of the relevant descriptive statistics (Panel 2, Figure 1) provides a clearer understanding of the actual relationship. This is one reason why descriptive statistics ought to always be reported in econometrics papers. The curve does not support a hypothesis of an inverse-U-shape, yet it shows a clear positive effect at low percentages, and a levelling off at an intermediate stage. This is an interesting result too as it supports theoretical arguments based on, for instance, diminishing marginal benefits. Yet it is no an inverse-U-shape. Thus, the interpretation of curvilinear regression results requires careful attention to the descriptive statistics of the same variable.

***** Insert Figure 1 here *****

The position of the turning point is particularly interesting for certain types of variables that are often associated with a non-normal distribution. Yet these types of variables also require special attention in the interpretation:

- Percentages, such as equity share, are defined only for the range zero to hundred, and saturation effects are likely close to these limits.
- Count variables, such as age measured in the number of years, are likely to relate non-linearly to underlying theoretical concepts such as experience.
- Ratios, such as debt equity ratios, also critically vary in their meaning along the scale, notably when they include values above and below unity.

The turning points of such variables may provide particularly valuable insights for both theory and management practice. For instance, if a study could clearly identify at what level of employee ownership firms perform best, this would suggest firms to aim to arrange their ownership structure close to this level.

Summary

Curvilinear functional specifications provide opportunities for more theoretically relevant and rigorous empirical analysis. However, they need to be conducted carefully in both their theoretical reasoning and their interpretation. In particular, the theoretical discussion needs to provide a solid argument for the suggested functional form, and the empirical analysis needs to connect the regression results with other statistical features of the dataset.

References

- Chan, C. M., Makino, S., & Isobe, T. 2006. Interdependent behaviour in foreign direct investment: The multi-level effects of prior entry and prior exit on foreign market entry. *Journal of International Business Studies*, 37(5): 642-465.
- Contractor, F. J. 2007. Is international business good for companies? The evolutionary or multi-stage theory of internationalization vs. the transaction cost perspective. *Management International Review*, 47(3): 453-475.
- Hennart, J.-F. 2007. The theoretical rationale for a multinationality-performance relationship. *Management International Review*, 47(3): 423-452.
- Meyer, K. E., & Sinani, E. 2009. When and where does foreign direct investment generate positive spillovers: A meta-analysis. *Journal of International Business Studies*, 40: forthcoming.
- Münich, D., Svejnar, J. & Terrell, K. (2005). Returns to human capital under the communist wage grid and during the transition to a market economy. *Review of Economics and Statistics*, 83(1):100-123.
- Takeuchi, R., Lepak, D. P., Marinova, S. V., & Yun, S. 2007. Nonlinear influences of stressors on general adjustment: The case of Japanese expatriates and their spouses. *Journal of International Business Studies*, 38(6): 928-943.
- Wong, C. S., & Campion, M. A. 1991. Development and test of a task level model of motivational design. *Journal of Applied Psychology*, 76(6): 825-837.

Table 1 Alternative functional specifications.


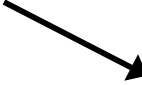

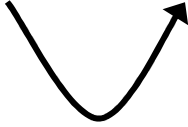





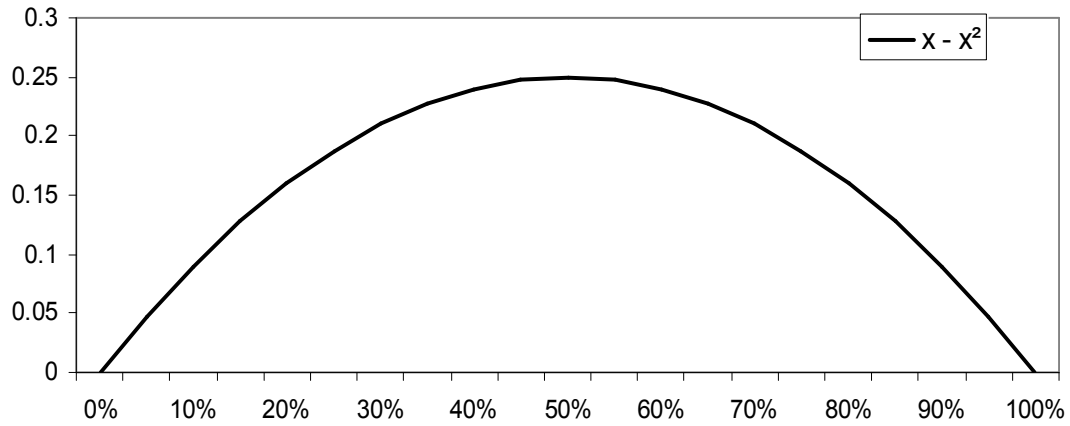
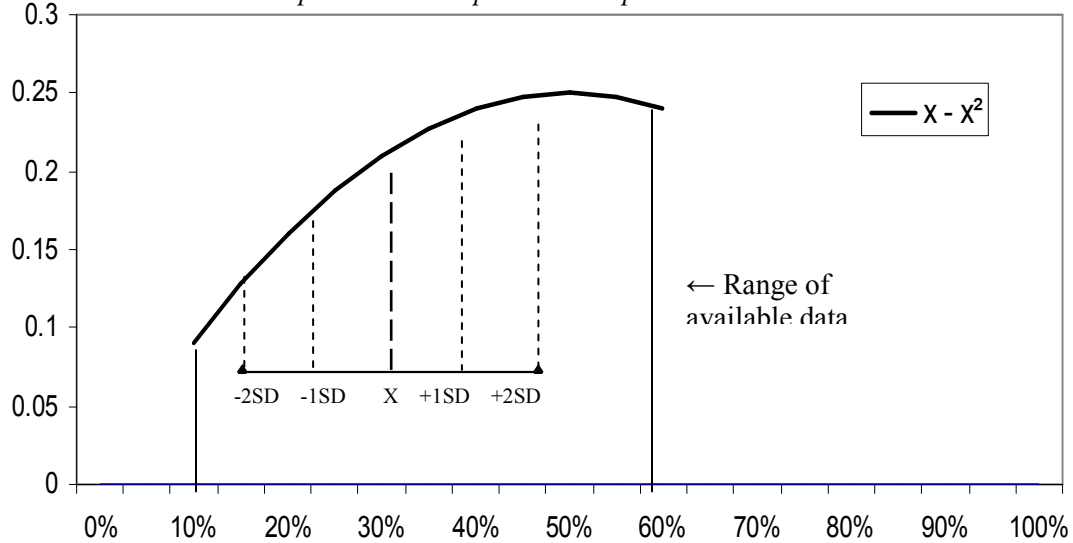
Case	Primary effects (A = positive, B = negative)	Aggregate effect	Functional form
1	A and B cancel each other out	No aggregate effect	No significant effect
2	A dominates over B		Linear: $\beta_1 x$, with $\beta_1 > 0$
3	B dominates over A		Linear: $\beta_1 x$, with $\beta_1 < 0$
4	A dominates at low level, B dominates at high levels		Quadratic (concave or inverse-U): $\beta_1 x + \beta_2 x^2$, with $\beta_1 > 0$, and $\beta_2 < 0$
5	B dominates at low level, A dominates at high levels		Quadratic (convex or U-shaped): $\beta_1 x + \beta_2 x^2$, with $\beta_1 < 0$, and $\beta_2 > 0$
6	A dominates at high and low levels, B dominates at intermediate levels		Cubic (S-shape): $\beta_1 x + \beta_2 x^2 + \beta_3 x^3$, with $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$
7	B dominates at high and low levels, A dominates at intermediate levels		Cubic (inverse-S): $\beta_1 x + \beta_2 x^2 + \beta_3 x^3$, with $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$
8	A dominates, but weakens for higher values		Logarithmic: $\beta_1 \log(x)$, with $\beta_1 > 0$
9	B dominates, but weakens for higher values		Logarithmic: $\beta_1 \log(x)$, with $\beta_1 < 0$
10	Qualitative shifts at certain benchmark values		Step-function

Figure 1: Interpreting a quadratic relationship

Panel 1: An inverse-U shaped relationship



Panel 2: An inverse-U shaped relationship with descriptive statistics



Note: \bar{X} = mean, SD = standard deviation