10-2008

Using spreadsheets to discover meaning for parameters in nonlinear models

Kris H. Green
St. John Fisher University, kgreen@sjf.edu

Using spreadsheets to discover meaning for parameters in nonlinear models

Abstract

Using spreadsheets one can develop an exploratory environment where mathematics students can develop their own understanding of the relationship between the parameters of commonly encountered families of functions (linear, logarithmic, exponential and power) and a natural interpretation of "rate of change" for those functions. The key to this understanding involves expanding the concept of rate of change to include percent changes. Through the use of the spreadsheet model, students can explore and easily determine which type of change is most natural for a given family of functions. This, in turn, provides a mechanism for interpreting the parameters of the function numerically, rather than graphically, as is common.

Disciplines

Mathematics | Science and Mathematics Education

Comments

This version is pre-publication.

Final version published as Green, K. (2008), Using spreadsheets to discover meaning for parameters in nonlinear models*. Journal of Computers in Mathematics and Science Teaching, 27*(4):423-411.

Using Spreadsheets to Discover Meaning for Parameters in Nonlinear Models

Kris H. Green, St. John Fisher College

ABSTRACT

Using spreadsheets one can develop an exploratory environment where mathematics students can develop their own understanding of the relationship between the parameters of commonly encountered families of functions (linear, logarithmic, exponential and power) and a natural interpretation of "rate of change" for those functions. The key to this understanding involves expanding the concept of rate of change to include percent changes. Through the use of the spreadsheet model, students can explore and easily determine which type of change is most natural for a given family of functions. This, in turn, provides a mechanism for interpreting the parameters of the function numerically, rather than graphically, as is common.

Topics: Spreadsheets, functions, rate of change, percent change, parameters.

Using Spreadsheets to Discover Meaning for Parameters in Nonlinear Models

The following presents a model of an activity for using spreadsheets to help students develop a more intuitive understanding of what parameters mean in four of the basic function families as well as an understanding of rates of change from a deeper perspective. This deeper perspective reinforces the notions of total change and percent change, concepts often involving considerable difficulty for students. In fact, by focusing on just these two concepts we can develop very natural interpretations of the parameters of these function families as different rates of change.

The activities were developed as part of a spreadsheet-based mathematics course for college students majoring in business. Students in the course use realistic data to develop mathematical models that can be used to analyze business-related scenarios in order to explain the situation and inform decision-making activities. The first half of the course is devoted to developing and understanding linear models of data, including multivariate data and data with categorical, rather than numerical, variables, such as gender or education level. The second half of the course focuses on developing and applying nonlinear models of data, including a brief introduction to differential calculus and optimization. The students have all completed the equivalent of a college algebra course, either through high school mathematics classes or as a college course.

This course was created at the request of our college's business school and was designed to use appropriate business software (spreadsheets) to develop mathematical understanding. Although we could have focused on many different functions, only four were identified as being of primary importance: linear, exponential, logarithmic and power functions. Initially, this course made extensive use of material from the textbook by Albright, Winston & Zappe (2003), and this activity developed from an attempt to help students verify and understand a statement in that text (p. 605) claiming that the constants in a particular logarithmic model could be interpreted as a fixed change in *y* for a one percent change in *x*. After due consideration, not only did the statement make sense, but it was also a more intuitive way for understanding rates of change than differential calculus.

MOTIVATION FOR DESIGNING THESE ACTIVITIES

At the heart of these activities lie three critiques of more standard methods for teaching students about nonlinear families of functions. First, students' understanding of the concept "rate of change" is usually built on a single class of examples: linear functions. This has the immense value of providing students with a strong foundation in one technique that applies to many problems in the real world. At the same time, this is limiting, since the idea of a constant proportional change, regardless of input level, applies only to this class of functions. And while this forms the basis for the study of differential calculus which students can use to successfully study the rate of change of any function at any point, this fails to produce a general description of the parameters of a nonlinear function, especially one of the ones listed above, in terms of some intrinsic "rate of change" that is on par with the slope of a linear function. In other words, for linear functions, the rate of change is a constant, and students can easily use the rule of three to attach meaning to this constant from graphical, numerical and algebraic perspectives. However, this fails for exponential, logarithmic, and power functions, which all contain constant parameters that are related to the slope in some way, but for which the slope is level-dependent, and the connections between the slope and these parameters is obscured by concentrating on proportional rates of change.

The second critique is that students may easily develop an understanding of how the parameters in a nonlinear function influence the function through the rule of three (or four), but this understanding is built, for the most part, case by case. They do not develop a general way of describing the parameters that can be used to study anything other than the single family currently under investigation. Lacking a more general framework for understanding these parameters almost forces students into a "stamp collecting" mode for learning about the function families, which reinforces the erroneous belief that "only geniuses are capable of discovering, creating, or really understanding mathematics" (Schoenfeld, 1988). This belief can be mitigated by experiences in the mathematics classroom, but in order for students to have opportunities for exploring and creating mathematics on their own, they must have tools appropriate for this. In particular, they need tools that allow them to explore formulaic, numerical and graphical representations simultaneously. While graphing calculators are useful for this, they lack a property I refer to as "immediacy": due to the size of the screen in a graphing calculator, it is difficult to view more than one representation at a time. Thus, when a student changes a parameter value on one screen, she has to go to a different screen in order to see the results of this change on the graph. There is much opportunity for loss of and confusion of ideas during this transition. With spreadsheets, however, these different representations can be viewed together and induced changes in one representation can be immediately linked to the changes in another. For example, students can immediately connect a change in the numerical parameters of a function with a change in the graphical representation of the function. For these reasons as well as motivational reasons related to the use of spreadsheets in the workplace (Shore, Shore & Boggs, 2004, p. 226) we will make use of spreadsheets to explore these functions and develop deeper understanding rather than using graphing calculators or non-technology approaches to studying the functions above.

Finally, there is a substantial body of literature suggesting that the emphasis placed on linear and proportional reasoning has a negative effect on student problem solving. Van Dooren, De Bock, Hessels, Jansen, & Verschaffel (2005) discuss a large body of literature related to the way students misapply proportional reasoning in common sense and other problem solving activities, such as those involving the way in which the area or volume of a figure increase if the linear dimensions are doubled. Their empirical study also suggests that misapplication of proportional reasoning is the largest category of mistakes students make in mathematics problems. Thus, students must encounter other ways of measuring change and look at situations where these other types of change occur naturally. They must confront the differences in these situations and develop natural, intuitive ways for analyzing change in these contexts. While it is relatively common to develop exponential functions in this way, by starting with a process such as bacteria growth for which each time period results in the population increasing by a fixed percent (often referred to as the growth rate, r,) once the concept of the exponential has been developed from repeated multiplication by the factor $(1 + r)$ and converted to an expression in terms of the constant e, the reverse process is not always applied in order to interpret the rate of growth. Nor is it extended to other families of functions.

The notion of rate of change underlies almost all major quantitative disciplines in the world. We constantly hear about the change in the economy as measured by income or employment, changes in our investments, changes in the environment. With only a linear concept of change, students are at a disadvantage for making informed decisions regarding these and many other situations. And while we can linearize many quantities to approximate their change locally, this often provides little information about the long-term behaviors of the quantities under investigation. Take for example the rate of change of an arbitrary function near a critical point. The slope is zero at the critical point (by definition) but that does not remotely imply that the slope is zero everywhere. Thus, rather than focusing only on measuring change locally, students should also learn how global properties like the parameters of a function relate to local changes. If done effectively, this can also help students build understanding for exploring other situations not covered by their previous experience.

SETTING UP THE ACTIVITY

Throughout the activity, students will be dealing with four basic families of functions: linear, logarithmic, exponential, and power functions. The basic symbolic representation for each will be taken as shown in figure 1, so that each has two parameters, labeled *A* and *B*. Before conducting the activity, my students have used spreadsheets to graph these four functions and explore how shifting and scaling helps these functions fit a variety of data. These students are familiar with the algebraic and graphical forms of each, but still struggle to understand what each of the constants in the formulae means, often reverting to "slope" and "y-intercept" even for functions such as the exponential for which the value of its slope depends on the value of its independent variable. This stands in marked contrast to the slope of a linear function which *is* constant for all values of the independent variable. In business applications, this property of linear functions is referred to as "level independence" since the level of the input does not affect the slope.

The students have explored these functions primarily through the use of spreadsheets, developing tables of data for each function showing both the dependent and independent variables, linking them through a formula which emphasizes which numbers in the functional representation are constant for all values of the independent variable (the parameters) by using absolute cell references to refer to parameter values. A typical spreadsheet is shown in figure 1 along with the graph of these data.

Table 1. Four basic functions and their representations for this activity.

Family of functions	Symbolic form
Linear	$y = Bx + A$
Logarithmic	$y = B\ln(X) + A$
Exponential	$y = Ae^{Bx}$
Power	$y = Ax^{B}$

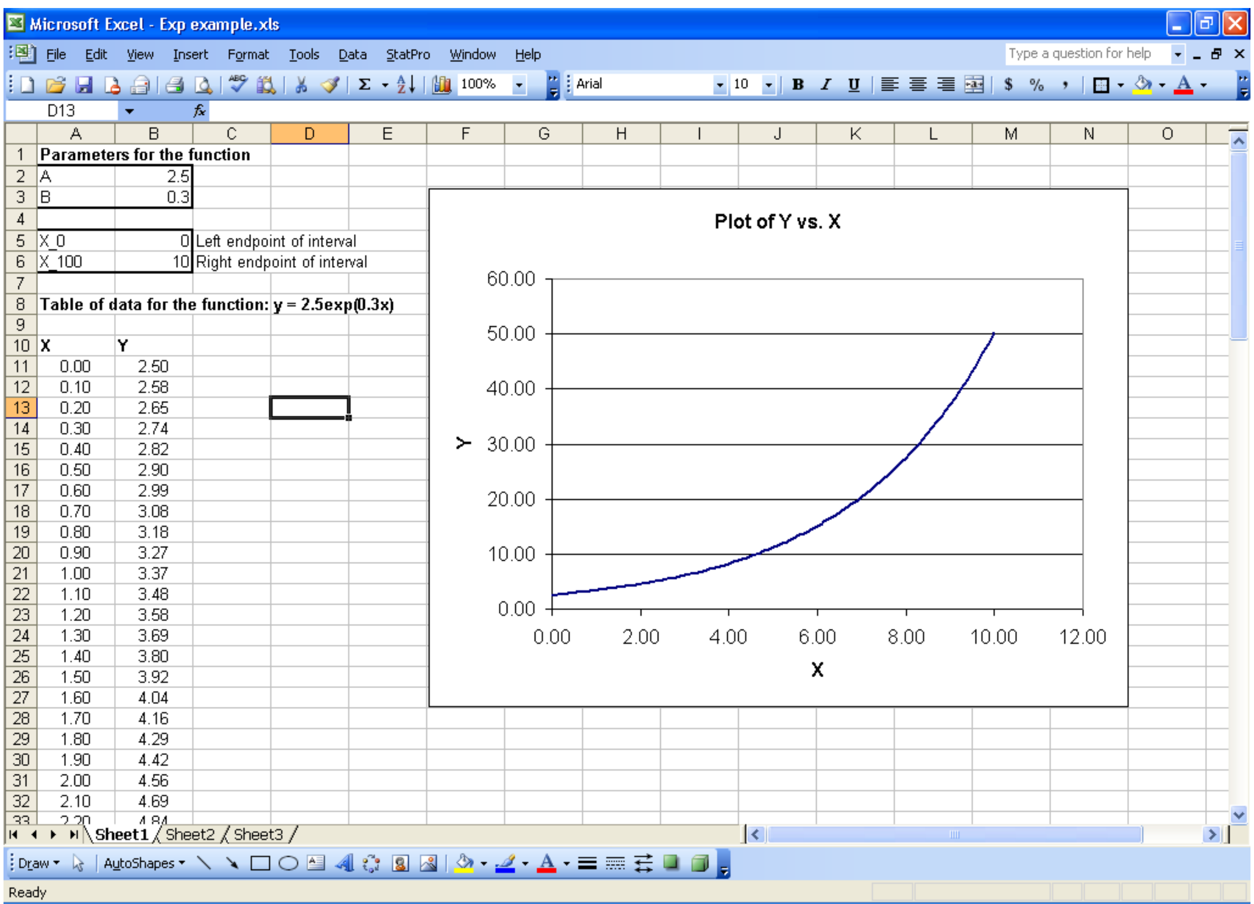

Figure 1. Spreadsheet showing parameter values, table of data, and graph of an exponential function.

After building the table of data, the students explore what happens to the graph, the formula and the table of values as each parameter is changed. Students are encouraged at this stage to develop descriptions of the parameters based on what they observe. Through class discussion, we compare these descriptions, refining them until they accurately reflect the nature of the parameters. For example, students often note that as the parameter *B* is changed the graph of the logarithmic function gets steeper or shallower. Thus, they identify this with "slope", failing to note that the word "slope" has a specific mathematical definition. Prompting them with questions about the differences between the slope they observe in the logarithmic function and the linear function, the students try to account for the fact that the slope in the logarithmic function is level dependent. But they lack language for this, and they have trouble finding something that is constant that they can associate with the parameter.

As a further example of their difficulties, students at this stage tend to identify many of the constants in the algebraic forms from table 1 with "y-intercept". And while it is true that the parameter *A* in the linear and exponential functions is associated with the point (0, *A*) on the graphs, this does not work for the parameter *A* in the logarithmic function, although it does shift the graph vertically. In the case of the power function, students encounter another serious problem: the y-intercept is always at the point (0, 0). Thus, it seems, anecdotally, that students' familiarity with linear functions and the parameters of slope and y-intercept becomes a hindrance to their exploring and describing what they observe. They seem to assume instead that all functions with two parameters must have a slope and a y-intercept, and their job is to determine which number plays which role. These observations are commensurate with Alan Schoenfeld's (1988) discoveries regarding the unintended lessons that students glean from even the best implemented mathematical curriculum: rather than believing that they can develop their own descriptions of the phenomena at hand, descriptions that account for all aspects observed, they seek to use expert-derived descriptions that are inappropriate. Once I point out what all the students have noticed as obvious, that the slopes are different at different points on the graph, the students seem to change attitudes. This is often accompanied by a statement like, “Oh, I thought we could only use ‘slope’ and ‘y-intercept.’”

The final component of preparing students for the activity is to elicit their ways of conceiving of “change” mathematically. It is important here to focus on a simple scenario involving one variable. For example, one can look at the stock market closing values, such as the Dow Jones Industrial Average (DJIA), each day for a period time. One then asks simply, “How might we measure the change in this quantity from one day to the next?” Students almost always suggest taking the difference (total change) first. A few quick calculations can show how easy this is to use. To get them to consider other ways to measure change, the question can be expanded. “Suppose we want to compare the change in the DJIA over one week to the change in the DJIA last year?” Here, the amounts of time are different, so students quickly recognize that comparing the total change over a long period of time to the total change over a shorter period of time fails to provide useful information. Instead, they need to put these onto the same footing. One way is to use slope: measure the change as the total change in DJIA divided by the total change in time for the two periods and then compare these ratios. Another way, one commonly suggested by students who remember some of their science labs, is to use the percent change in a quantity. Thus, we have two simple ways to measure the change in a single quantity. We can measure the total change in the quantity by a simple difference, or we can measure the percent change by dividing the difference by the initial amount. These two methods are all that will be needed to develop more natural understandings for the parameter *B* in each of the families of functions shown in table 1.

This can be seen easily as a sort of combinatorial problem. We are searching for an understanding of a parameter in a function relating two variables, *x* and *y*. Between any two points on the graph of the function, we can measure the total change or percent change in each variable and form a simple ratio. For example, the slope of a linear function is really a ratio of the total change in *y* to the total change in *x*. But we could look at a ratio of the total change in *y* to the percent change in *x*, or any other combination. We have two variables and two ways of measuring change, giving four possible ratios to explore for each function family. And, since we have emphasized throughout the process of building linear models that one can interpret the slope in terms of what happens to the dependent variable when the independent variable increases by one unit, we can simplify our ratios involving total change in the independent variable to always focus on a one unit change. Similarly, to make comparisons standard, we consider only changes in the independent variable of one percent for ratios involving percent changes in *x*. We are then left with four ratios that could be used to measure "rate of change" for any given function. We can measure either the total change or the percent change in *y* as a result of a one unit change in *x*, and we can measure the total or percent change in *y* as a result of a one percent change in *x*.

ACTIVITY 1: FINDING NATURAL MEASURES OF CHANGE

The goal in activity 1 is to have students explore the two ways for measuring the change of the input variable (*x*) and to find measures of change for the output variable (*y*) that are level independent. In particular, activity 1 involves setting up a spreadsheet like the one shown below (figure 2). The spreadsheet has cells where students can place the parameters of the particular function they are working with (in figure 2, the function being explored is an exponential with the parameters $A = 2$, and $B = 1.5$). By setting up calculations for each of the four possible ratios discussed above, students can then determine which, if any, of these ratios is constant across values of the independent variable, and therefore natural to use in describing the rate of change of the dependent variable. To construct these ratios, each student creates a column of *x* values for the function. The next column contains values of $x + 1$, a one unit increase in the independent variable. The third column contains values representing a 1% increase in *x*. The next three columns contain values for $y_1 = f(x)$, $y_2 = f(x+1)$, and $y_3 = f(x+0.01x) = f(1.01x)$. The final four columns compute the four different ratios possible with these two measures of change: the total change in *y* for a given one unit change in *x*, (column G) the total change in *y* for a given 1% change in *x* (column H), the percent change in *y* for a given one unit change in *x* (column I) and the percent change in *y* for a given 1% change in *x* (column J). Spreadsheet formulas for these are shown in table 2 using the notation of Microsoft Excel, which is fairly standard across spreadsheet packages.

While one could certainly save class time by setting up such a spreadsheet and distributing it to the class, having the students develop the spreadsheet under teacher guidance serves to reinforce some of the notation of algebra and the structure of the mathematical objects we are studying. In particular, it helps emphasize the difference between constants (parameters) and variables in formulas, since most spreadsheets use a notation like that shown in table 2, where dollar signs ($) are used to refer to cell values that are intended to remain fixed, regardless how the formula is copied to other cells. Thus, the parameters in cells C2 and C3 should always be referred to using these *absolute cell references* ($C$2 and $C$3) to make sure that these same values are used in all formulas relating to the parameters. *Relative cell references* lack the dollar signs, and indicate that the formulas used should change based on how the formula is copied to another cell. These references are necessary for variables (such as the *x* values in column A) so that as the formula to calculate *y* is copied down column D, the cell reference changes one row at a time to always refer to the *x* value in that row of the data table.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Function	Exponential	$y=Ae^{Bx}$									
2	Parameters	A	2									
3		B	1.5									
4												
5	x	x+1	1.01x	$y_1 = f(x)$	$y_2 = f(x+1)$	$y_3 = f(1.01x)$	$y_2 - y_1$	$y_3 - y_1$	$(y_2 - y_1)/y_1$	$(y_3 - y_1)/y_1$		
6	1	2	1.01	8.963378141	40.17107385	9.098842254	31.20769571	0.135464113	3.48168907	0.015113065		
7	2	3	2.02	40.17107385	180.0342626	41.39446518	139.8631888	1.223391332	3.48168907	0.030454534		
8	3	4	3.03	180.0342626	806.857587	188.3208544	626.8233244	8.286591818	3.48168907	0.04602786		
9	4	5	4.04	806.857587	3616.084829	856.7508737	2809.227242	49.89328673	3.48168907	0.061836547		
10	5	6	5.05	3616.084829	16206.16786	3897.720525	12590.08303	281.6356964	3.48168907	0.077884151		
11	6	7	6.06	16206.16786	72631.00535	17732.3721	56424.83749	1526.204249	3.48168907	0.094174284		
12	7	8	7.07	72631.00535	325509.5828	80672.02828	252878.5775	8041.022933	3.48168907	0.11071061		
13	8	9	8.08	325509.5828	1458832.74	367011.0298	1133323.157	41501.44697	3.48168907	0.127496852		
14	9	10	9.09	1458832.74	6538034.745	1669687.733	5079202.005	210854.9931	3.48168907	0.144536784		
15	10	11	10.1	6538034.745	29301438.86	7596112.647	22763404.11	1058077.902	3.48168907	0.161834243		
16	11	12	11.11	29301438.86	131319938.3	34557915.36	102018499.4	5256476.499	3.48168907	0.179393119		
17	12	13	12.12	131319938.3	588535132.1	157218510.2	457215193.8	25898571.95	3.48168907	0.197217363		
18	13	14	13.13	588535132.1	2637631469	715253212	2049096337	126718079.9	3.48168907	0.215310986		
19	14	15	14.14	2637631469	11821044126	3253988074	9183412657	616356604.5	3.48168907	0.23367806		
20	15	16	15.15	11821044126	52978244260	14803762088	41157200134	2982717962	3.48168907	0.252322716		
21	16	17	16.16	52978244260	2.37432E+11	67348548001	1.84454E+11	14370303741	3.48168907	0.27124915		
22	17	18	17.17	2.37432E+11	1.0641E+12	3.06397E+11	8.26664E+11	68964888872	3.48168907	0.290461621		
23	18	19	18.18	1.0641E+12	4.76895E+12	1.39393E+12	3.70485E+12	3.29832E+11	3.48168907	0.309964451		
24	19	20	19.19	4.76895E+12	2.13729E+13	6.34157E+12	1.6604E+13	1.57262E+12	3.48168907	0.329762028		
25	20	21	20.2	2.13729E+13	9.57869E+13	2.88505E+13	7.4414E+13	7.47751E+12	3.48168907	0.349858808		
26												
27												
28												
29												
30												
31												

Sheet1 / Sheet2 / Sheet3

Figure 2. Spreadsheet showing different measures of rates of change for an exponential function with *A* = 2, *B* = 1.5.

Table 2. Formulas for the spreadsheet in figure 2. After entering the formula, it can be copied down to the other cells in the column by simply double-clicking on the cell's fill handle.

Cell	**Formula**	**Cell**	**Formula**
B6	=A6 + 1	G6	=E6 – D6
C6	=1.01*A6	H6	=F6 – D6
D6	=$C$2*exp($C$3*A6)	I6	=(E6 – D6)/D6
E6	=$C$2*exp($C$3*B6)	J6	=(F6 – D6)/D6
F6	=$C$2*exp($C$3*C6)		

In the example shown in figure 2, we clearly see that only one of the four ratios is constant across all values of the input variable: Column I, showing the percent change in *y* based on a one-unit change in *x* is always equal to about 3.48. Thus, for this case, an exponential functions, one of the most complex examples from the four families, the level independent change comes from examining the ratio formed by the percent change in *y* given a one unit total change in *x*.

After exploring each of the functions, students will discover the natural interpretations for the change in the various functions shown in table 3. By the end of activity 1, students see that each of the functions has a natural way of measuring rate of change associated with it. From this, students see that it is possible to investigate mathematical situations and develop deeper understanding by making their own observations. This leads us to activity 2.

Table 3. Natural interpretations of change for the functions in figure 1 after exploration.

Function Family	Symbolic form	Natural sense of change
Linear	$y = Bx + A$	Change in *y* vs. change in *x*
Logarithmic	$y = B\ln(X) + A$	Change in *y* vs. % change in *x*
Exponential	$y = Ae^{Bx}$	% Change in *y* vs. change in *x*
Power	$y = Ax^{B}$	% change in *y* vs. % change in *x*

ACTIVITY 2: CONNECTING PARAMETERS AND CHANGE

The goal of activity 2 is to challenge students to determine the exact nature of the relationship between the parameter *B* in each function and the level-independent rate of change for that function family. The activity further emphasize the nature of mathematics as a logical process and leads the students into making and verifying their own conjectures about the situation by finding expressions for the way the natural ratios of change relate to the parameters of the functions. Given the table of the particular exponential function shown in figure 2, students can explore the spreadsheet by changing the values of the parameters to develop further insight. And while we could simply turn them loose to explore, expecting them to track their attempts and results, it helps to provide a slightly more structured exploration.

One of the easiest and most useful things to start with is to eliminate one of the parameters. By changing the value of *A* (in cell C2 of figure 2) students can notice that the value in column I does not change. In fact, this should be true for all of the functions in table 1; the parameter *A* could be left out in all cases. But leaving the constant in the formula emphasizes this connection to the graphical form of the function. For example, in the logarithmic function, the rate of change is the same at a particular value of *x*, regardless of whether the function is shifted up or down by any amount.

Students can also keep a systematic record of values for *B* and the resulting ratio in column I (or whichever column represents the appropriate constant ratio for the function being investigated). If they organize the table well, they can see that as *B* increases, so does the value of the ratio in column I. As *B* decreases, so does the value of column I, and it seems that column I has a lower bound of –1. Determining the range of possible values for the rate of change is important for developing a connection between the input, the parameter *B*, and the output, the rate of change. For this, students should be encouraged to try all sorts of values for *B*: negative values, small values, large values, non-integer values. But one should emphasize that a systematic study is helpful (i.e. always increasing or always decreasing input values).

Once students have a table of the level-independent change versus *B*, they can construct a scatterplot of the data. In the example shown, this should suggest that the value in column I is related to *B* by an exponential factor. Noting that the exponential has a horizontal asymptote at $y = 0$ while this graph has a horizontal asymptote at $y = -1$, suggests that one should shift the values up by 1. The data can then be fit perfectly by an exponential: (% change in y) + 1 = e^B.

Once this expression has been proposed, based on the observations and evidence, we can verify it algebraically, emphasizing the power of algebra to prove general statements rather than just specific examples. All students really need are the rules of exponents to derive this result from the initial function. Since $y = Ae^{Bx}$, we have the relation

$$\frac{y_2 - y_1}{y_1} = \frac{Ae^{B(x+1)} - Ae^{Bx}}{Ae^{Bx}} = \frac{Ae^{Bx}(e^B - 1)}{Ae^{Bx}} = e^B - 1.$$

This matches exactly with the relation found from the graph, but that relation was derived ultimately from a few specific examples. Verifying this relationship algebraically for any input values should provide more confidence in its accuracy.

After discovering these results, students can then follow them up by using algebra and the properties of each family of functions to derive the results in the rightmost column of table 4. None of these requires much algebraic effort, except perhaps the approximate meaning of the power function results. For example, $B\ln(1.01x) - B\ln(x) = B\ln(1.01) + B\ln(x) - B\ln(x)$, which simplifies to the expression given above. One can also argue for these results from another perspective, using Taylor series representations of each function, if that is appropriate.

Table 4. Relationship between the parameters and rates of change.

Function Family	Symbolic form	Interpretation of the parameter
Linear	$y = Bx + A$	*Y* changes by *B* when *x* increases by 1 unit.
Logarithmic	$y = B\ln(X) + A$	*Y* changes by 0.01*B* when *x* increases by 1%.
Exponential	$y = Ae^{Bx}$	*Y* changes by ($e^B - 1$)% when *x* increases by 1 unit.
Power	$y = Ax^B$	*Y* changes by a factor of $1.01^B - 1 \approx 0.01B$ (or *B*%) when *x* increases by 1%.

GENERALIZING THESE IDEAS AND EXAMPLES

It may seem that now students have four different statements, one per family, to learn in order to interpret the parameters. However, there are deeper connections that can be made. These deeper connections simplify the rules from table 4 considerably, making the patterns clearer.

One should note that for functional models in which *y* is proportional to *B* (taking *x* to be a number in the linear model and ln(*x*) as a constant in the logarithmic model) the correct interpretation of the rate of change involves the total change in *y* rather than the percent change. In models where the factor "*Bx*" appears (linear and exponential) we find that the rate of change is naturally interpreted using total change in *x* rather than percent change. Also note that in each case, we have chosen to label the parameters of the models so that *B* is always the parameter related to the rate of change.

As a way of understanding why these rules work and providing a general framework, one can think about the use of regression to derive models of the types explored above. In order to apply linear regression techniques to build nonlinear models, one must transform the data in some way. The typical methods for transforming explanatory (independent variable) data and response (dependent variable) data and the resulting linear equations are shown below in table 5. In all cases, this process provides the crucial link for understanding the "natural interpretation of change" for each function family.

Table 5. Table showing the relationship between the function families and the needed linearization to construct such a regression model.

Function Family	Independent Variable	Dependent Variable	Regression Model
Linear	x	y	$y = Bx + A$
Logarithmic	$\ln(x)$	y	$y = B\ln(x) + A$
Exponential	x	$\ln(y)$	$\ln(y) = Bx + \ln(A)$
Power	$\ln(x)$	$\ln(y)$	$\ln(y) = B\ln(x) + \ln(A)$

In each case where a variable must be "logged" in order to prepare it for linear regression, the natural interpretation of change requires that one think in terms of a percent change of the logged variable, rather than a one unit change. For example, the data shown in figure 3 are taken from an example in Albright, Winston & Zappe (2002, pp. 612-613) relating the time to complete production of a batch of some item to the number of batches that have been completed. One expects that as experience with the production process increases, indicated by later batch numbers, the time to complete a batch will decrease, at least up to a point, since familiarity with the process should improve performance. The graph of these data suggests that a logarithmic function might be appropriate as a fit for the data. Thus, we should express the data as the ordered pairs (ln(*x*), *y*) and attempt a linear fit to this transformed data. The graph in figure 4 shows the transformed data and a linear fit, which leads to the regression equation

$$\text{Time} = -14.09\ln(\text{Batch}) + 122.95.$$

From figure 4, we can interpret this model to mean that each 1% increase in the number of batches completed results in a 0.01(-14.09) minute increase in completion time. This amounts to a 0.1409 minute decrease in time after completing 1% more batches. Thus, after 100 batches, we see that the marginal increase in performance is negligible when compared to the scale of the completion times. However, when few batches have been completed, say 10, a 1% increase in experience (number of batches) requires less than 1 batch to achieve, thus the gains are more significant at the beginning.

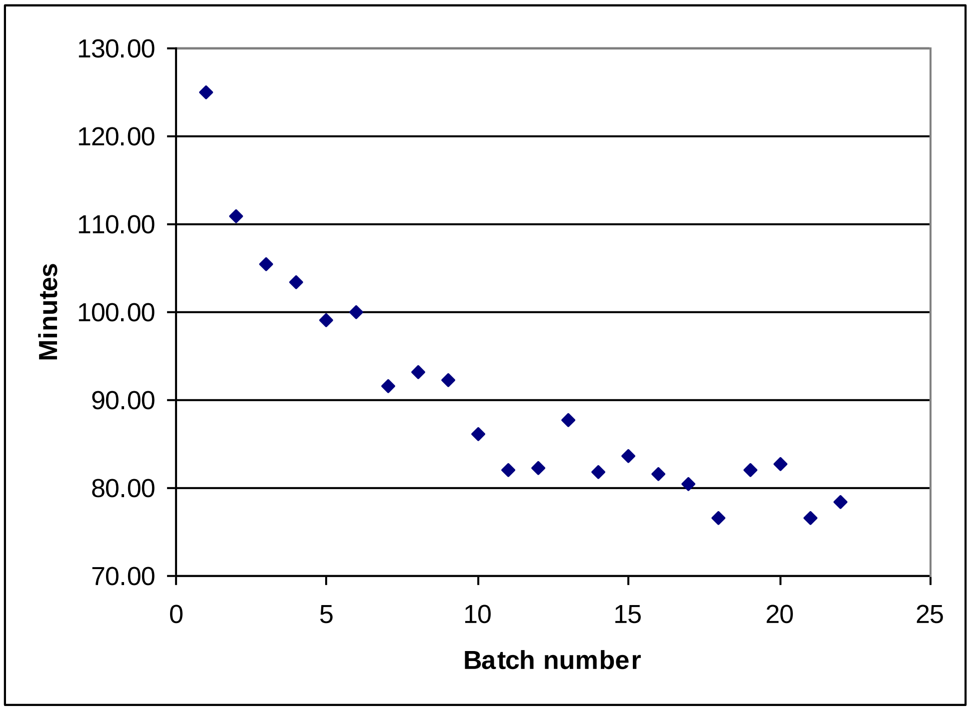

Figure 3. Data suggesting a nonlinear fit might be appropriate for predicting the number of minutes to complete a production batch vs. the number of batches already completed.

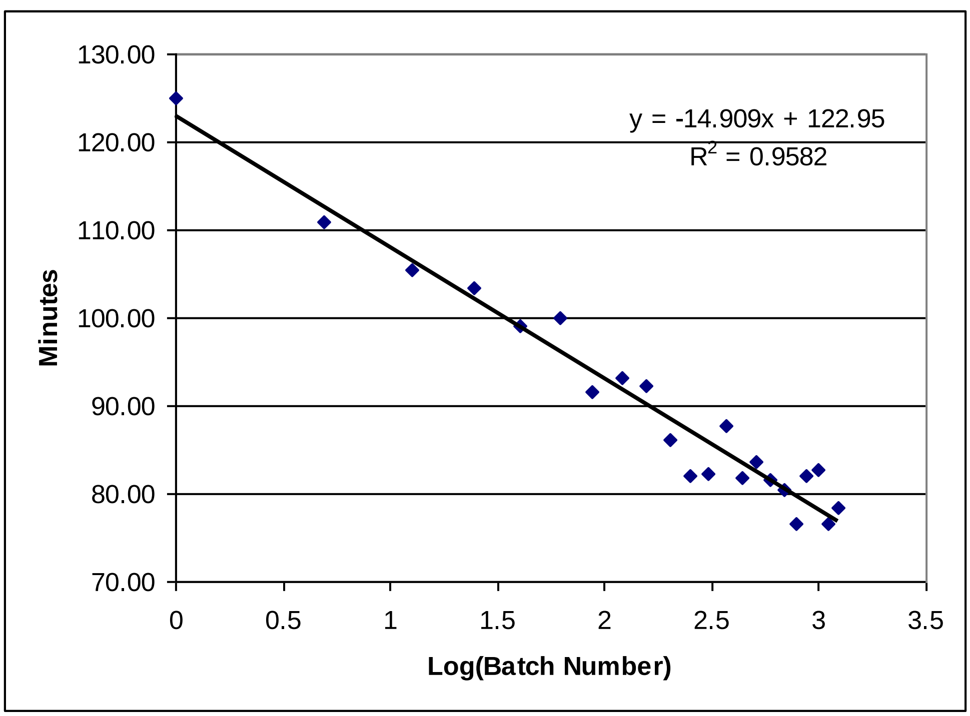

Figure 4. The data from figure 3 after transformation by logarithm with linear fit overlaid.

These types of data transformation are commonly seen in science as well; indeed, Kepler's third law of planetary motion, which relates the orbital period of a planet around the sun (the length of its solar year) to its average distance from the sun, can easily be derived from such a process, using the orbital distance and orbital period of each planet as the data. In this case, though, the data showing the average distance of a planetary body, labeled *a*, measured in Astronomical Units (AU), from the sun and the length of its solar year, labeled *T*, measured in earth days, suggest a power function as the most appropriate fit for the data. Thus, we should express the data as (ln(*a*), ln(*T*)) and attempt a linear fit to this transformed data. In class discussions, I often tell my students that these models are quite "clear cut" since we have "logged" everything in sight. It is easy to obtain the necessary data for the solar system from many sources. An example in Giordano, Weir & Fox (2003, pp. 67-68) provides both the data and context for treating this as an example of a proportional relationship, although the two quantities that are proportional are not *a* and *T*, but a^3 and T^2, which we can see as being approximately correct in a moment.

Linear regression on the transformed data shown in figure 5 produces the model

$$\ln(T) = 1.5001\ln(a) + 5.9006 \qquad (R^2 \approx 1).$$

We can now exponentiate both sides of the equation and simplify in order to solve for *T* rather than ln(*T*). This gives us Kepler's third Law: $T = ka^{1.5001}$ with *k* = 365.25656. Note that expressing the orbital period in earth years would have resulted in a final form of the law with *k* = 1, since one earth year is approximately 365.25 days and the earth is, by definition, 1 A.U. from the sun. Now, using the ideas from table 4, we see that a characteristic of this data is that each 1% increase in a planet's average orbital distance produces an increase in orbital period by a factor of $1.01^{1.5001} - 1 = 0.01504$, or about a 1.5% increase in the orbital period. Notice that the percent increase in the dependent variable is approximately equal to the slope constant in the "untransformed" equation, which becomes the exponent of the independent variable in the final form of the model. This intuition about the parameters involved provides students with a powerful tool for estimating relationships, imparting deeper meaning to the parameters than the graphical meaning usually provided to students for these four functions. As a final word on the matter, most science texts report Kepler's Law in the form $T^2 \propto a^3$ since $1.5001 \approx 3/2$, and indeed, Giordano, et al (2003) conclude their example with this observation.

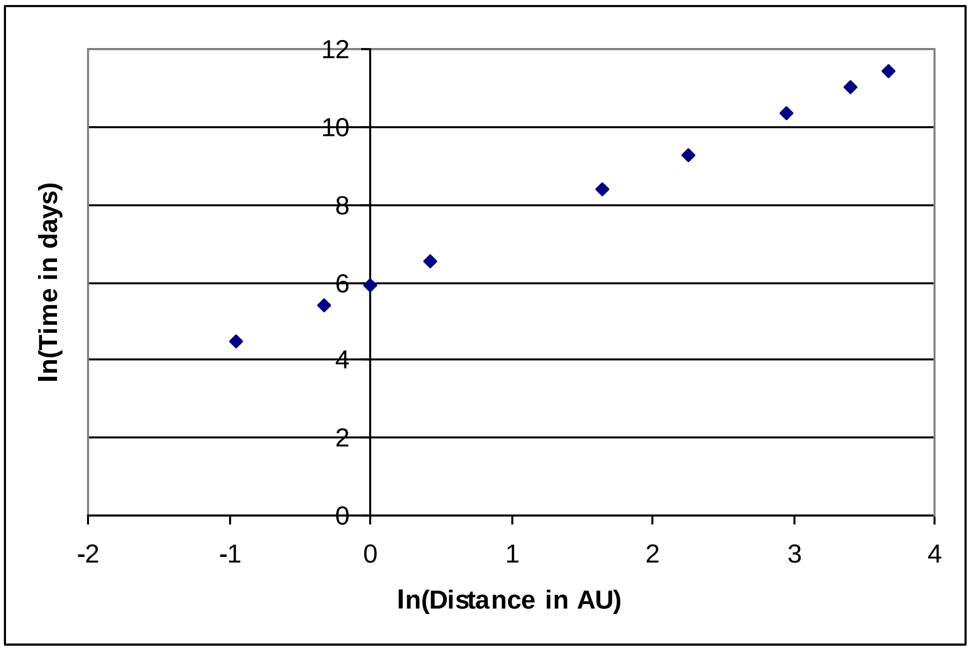

Figure 5. Plot of solar system orbital characteristics showing a nearly linear relationship between the logarithm if the distance from the sun (average) and the logarithm of the year length.

DISCUSSION

Spreadsheets then provide a powerful tool for the development of mathematical reasoning. By guiding students to build appropriate examples, teachers can encourage constructivist approaches to mathematics in which students develop their own rules and explanations. By then comparing these to other explanations, including those of experts, learning can be deepened considerably. Rather than a list of facts handed down, the mathematical ideas become a dynamic construct that students have participated in. The example provided here shows how students can be encouraged to understand the notion of change at a deeper level than simply “rise over run”. It unifies the concepts of total change, percent change and rate of change through the use of linear regression in a way that provides students with a powerful tool for understanding some of the mathematical models that are more commonly used.

The National Council of Teachers of Mathematics (NCTM, 2000, pp. 16 – 19) emphasizes the need for more constructivist approaches to teaching mathematics. And while a substantial body of literature reports that deeper learning takes place in such classroom environments, there also seems to be a sense among many educators at all levels that such approaches are “slower” and result in fewer topics covered. This criticism could certainly be leveled at the activities presented here, which take considerable time in class. Building the spreadsheets, exploring them, developing personal explanations for what is going on, and comparing these personal ideas with others in order to develop a common understanding for communication takes a full 80 minute period at the very least. It certainly would be faster to give students the results in table 4 and then practice applying these to a variety of functions. I suspect that would take about 15 to 20 minutes, leaving more time for other topics.

But what would students truly gain from that experience? They would see math as something handed down from authority and something to be memorized rather than developed. In terms of the revised Bloom’s taxonomy developed by a team led by Anderson and Krathwohl (2001) the explorations here provide not only factual knowledge for students, but also procedural, conceptual, and metacognitive knowledge. Students learn the different ways to interpret each family of functions (factual), but also learn how to explore problems and how to ask questions (metacognitive). They also see how the parameters of a function are related to the rate of change (conceptual) and experience setting up a spreadsheet for systematic exploration of an idea (procedural). Rather than limiting students to learning in one knowledge domain, they experience the activity through all knowledge domains. This deeper exploration of the "rate of change" concept can then be referred to when future ideas are encountered and need exploration and can serve as a basis for other explorations. For example, when Taylor series are encountered, these ideas can be revisited and comparisons made in terms of the accuracy of the approximations. Students could also search for other functions with similar descriptions for rates of change, or find rates of change descriptions for the parameters in trigonometric functions. The final interpretations in table 4 are, in some cases, approximated; students could compare the actual amounts of change to the predicted amounts, developing an understanding about approximations that could ultimately lead to the subject of numerical analysis.

Moreover, the general framework for these activities can be applied to develop constructivist lessons for any number of other topics. This framework is similar to the process of science: we look at what we know, we ask questions, we set up some experiment or activity to collect data, and then we review the data to see what patterns exist, possibly redesigning the activity to collect more data. This is, basically, what a constructivist approach to learning mathematics emphasizes. But the final step of the scientific method is also important. In fact, communicating the results becomes almost critical to the development of mathematics. Since we ultimately seek to help our students not only understand the mathematics through constructing their knowledge, but also to be able to communicate this knowledge to others, we must help them see the value of common definitions and interpretations. It is not that these are necessarily "right" or "wrong". Rather, these commonalities are conventions that help us facilitate communication. All of the students could easily come out of the activity above with the same basic idea that the parameters are connected to the ratios somehow but with possibly different ways of expressing these connections. In fact, it is likely that they will develop connections that have no common framework – one way of expressing the relationship for exponentials and another for each of the other functions. This is where closure is important; the instructor must bring these ideas together, look for common ideas, and try to express these common ideas as accurately and succinctly as possible. Students will always have their individual notes, ideas, and memories of the experience; these form the foundation of their understanding, creating a scaffold upon which the "accepted" definitions and concepts can be structured. So a larger goal for the activity should be to help students learn how compare their ideas developed through the exploration to the "expert" ideas and look for what properties make the expert descriptions (table 4) more complete or useful. This must be done carefully, though, in order to avoid sending the wrong message to the students. It has been my experience that student descriptions are usually not too far off, especially if I carefully monitor the process and provide appropriate prompts to push students into considering other cases or trying to come up with a description that includes all possible values.

From a teaching standpoint, we are left with several directions to pursue. First, the activity needs some refinement. With a large body of student work, including their notes and comments during the activity, not just their polished ideas at the end, one could build a list of student responses and difficulties and develop strategies for helping students overcome these. Given the open-ended nature of the exploration phase of these activities, this seems especially valuable and necessary. Second, even with a polished activity, students need to be able to connect these ideas to other settings and apply them in other situations. Work on this would require students to encounter situations involving many different types of change and from these, identify what family of functions best represents the situation, possibly using this to extrapolate to future values. For instance, students could be given a scenario in which a particular production line required two hours to complete their first batch of a product, but only 90 minutes to complete the tenth batch. From this, they could be asked to consider what kinds of change are reasonable and what future expectations they would have for this production line. This would help assess whether students are able to internalize and apply these concepts, rather than simply memorize the final outcome (table 4). A third direction for future work would translate the process skills – setting up a spreadsheet and exploring to find connections and meaning – to another situation, one that may not involve rates of change at all. These follow-up projects would help to further establish the value of both the constructivist learning approach in such situations and would further validate the idea that students need to learn how to learn so that fewer topics, covered at a deeper level of understanding, can be seen as valid even in mathematics classes, where learning is often tightly constrained by what future courses and applications students will need to experience.

In conclusion, the framework for these activities seems to provide students with an opportunity to become mathematical investigators while deepening their knowledge of mathematics. And while future work is certainly needed for understanding exactly how this deeper learning and experience is internalized by students, we can see clearly that such activities are a necessary and important step in re-conceiving learning experiences to help students construct meaning. To do this, though, we must be open to different approaches and ways of thinking about content that we have traditionally treated from a single approach, for example the notion of “rate of change” as discussed here.

ACKNOWLEDGEMENTS

The author would like to thank the referees and editor for comments that greatly improved the tone and strengthened the research background for this work.

REFERENCES

Albright, S., Winston, W. & Zappe, C. (2002). *Data Analysis and Decision Making with Microsoft Excel*. Albany, NY: Duxbury.

Anderson, L., & Krathwohl, D. (Eds). (2001) *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Giordano, F., Weir, M., & Fox, M. (2003). *A First Course in Mathematical Modeling*. Pacific Grove, CA: Thompson-Brooks/Cole.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and Standards for School Mathematics*. Reston, VA.: NCTM.

Schoenfeld, A. (1988). When good teaching leads to bad results: The disasters of "well taught" mathematics courses. *Educational Psychologist*, 23(2). 145-166.

Shore, M., Shore, J. & Boggs, S. (2004). Using spreadsheets and streaming video for developmental, teacher education, and general education mathematics courses. *Mathematics and Computer Education*, 38(2). 221-229.

Van Dooren, W., De Bock, D., Hessels, A., Janssens, D., & Verschaffel, L. (2005). Not everything is proportional: Effects of age and problem type on propensities for overgeneralization. *Cognition and Instruction*, 23(1). 57-86.