On Examples, Counterexamples, and Proof by Example

Thomas Mitchell

July 12, 1996

This piece is motivated by two seemingly unrelated items that have recently appeared in the *Journal* ([1] and [3]). In [1], the editor of the *Journal* wrote of “mathematical cranks” and how frustratingly difficult it can be sometimes to get would-be “circle-squarers” and “angle-trisectors” to see the errors of their ways. As an economics professor with a degree in mathematics, I am only too well aware of circle-squarers, angle-trisectors, and those who would make implicit use of the “Law of Universal Cancellation,” also mentioned in [1].

On the “Law of Universal Cancellation”

A few years ago I wrote a short paper ([2]) that I hoped would discourage economics students from employing a proof “technique” that I call “proof by example.” Sometimes objectives are best obtained through humor, so consider the following “theorem,” which formalizes the Law of Universal Cancellation.

Theorem 1 *If the numerator and denominator of a fraction have a digit in common, then the fraction can be reduced by simply striking the common digit from the numerator and denominator.*

“Proof.” Clearly, the theorem holds if the numerator and denominator both end in “zero”; we can strike the trailing zero from both the numerator and the denominator and correctly reduce the fraction. Less obviously, consider the following:

$$\frac{1\not{6}}{\not{6}4} = \frac{1}{4}, \quad \frac{1\not{9}}{\not{9}5} = \frac{1}{5}, \quad \frac{2\not{6}}{\not{6}5} = \frac{2}{5}, \quad \frac{4\not{9}}{\not{9}8} = \frac{4}{8};$$

. . . .

$$\frac{\not{3}4}{1\not{3}6} = \frac{4}{16}, \quad \frac{\not{3}4}{2\not{3}8} = \frac{4}{28}, \quad \frac{\not{6}7}{2\not{6}8} = \frac{7}{28}, \quad \frac{\not{6}7}{4\not{6}9} = \frac{7}{49},$$

$$\frac{\not{9}6}{1\not{9}2} = \frac{6}{12}, \quad \frac{\not{9}7}{1\not{9}4} = \frac{7}{14}, \quad \frac{\not{9}7}{2\not{9}1} = \frac{7}{21}, \quad \frac{\not{9}8}{1\not{9}6} = \frac{8}{16},$$

$$\frac{\not{9}8}{2\not{9}4} = \frac{8}{24}, \quad \frac{\not{9}8}{3\not{9}2} = \frac{8}{32};$$

. . . .

$$\frac{\not{2}3}{1\not{2}65} = \frac{3}{165}, \quad \frac{\not{4}6}{1\not{4}95} = \frac{6}{195}, \quad \frac{\not{5}4}{1\not{5}12} = \frac{4}{112}, \quad \frac{\not{5}6}{4\not{5}92} = \frac{6}{492},$$
$$\frac{\not{5}7}{1\not{5}96} = \frac{7}{196}, \quad \frac{\not{6}7}{3\not{6}85} = \frac{7}{385}, \quad \frac{2\not{6}7}{14\not{6}85} = \frac{27}{1485}, \quad \frac{5\not{0}1}{39\not{0}78} = \frac{51}{3978},$$
$$\frac{6\not{1}2}{50\not{1}84} = \frac{62}{5084}, \quad \frac{8\not{4}6}{27\not{4}95} = \frac{86}{2795}, \quad \frac{9\not{3}4}{72\not{3}85} = \frac{94}{7285}.$$

In the following, simply delete the zero in the middle:

$$\begin{array}{cccccc}
\frac{101}{202} = \frac{11}{22}, & \frac{101}{303} = \frac{11}{33}, & \cdots, & \frac{101}{808} = \frac{11}{88}, & \frac{101}{909} = \frac{11}{99}, \\
 & \frac{202}{303} = \frac{22}{33}, & \cdots, & \frac{202}{808} = \frac{22}{88}, & \frac{202}{909} = \frac{22}{99}, \\
 & & \ddots & \vdots & \vdots \\
 & & & \frac{707}{808} = \frac{77}{88}, & \frac{707}{909} = \frac{77}{99}, \\
 & & & & \frac{808}{909} = \frac{88}{99};
\end{array}$$

. . . .

In the following, the middle digits can be deleted to reduce the fraction:

$$\frac{134}{737}=\frac{14}{77},\quad \frac{143}{242}=\frac{13}{22},\quad \frac{156}{858}=\frac{16}{88},\quad \frac{165}{363}=\frac{15}{33},\quad \frac{178}{979}=\frac{18}{99},\quad \frac{187}{484}=\frac{17}{44},$$
$$\frac{242}{341}=\frac{22}{31},\quad \frac{264}{363}=\frac{24}{33},\quad \frac{286}{484}=\frac{26}{44},\quad \frac{363}{462}=\frac{33}{42},\quad \frac{363}{561}=\frac{33}{51},\quad \frac{385}{484}=\frac{35}{44},$$
$$\frac{484}{583}=\frac{44}{53},\quad \frac{484}{682}=\frac{44}{62},\quad \frac{484}{781}=\frac{44}{71},\quad \frac{536}{737}=\frac{56}{77},\quad \frac{737}{938}=\frac{77}{98}.$$

$$\frac{102}{306}=\frac{12}{36},\quad \frac{102}{408}=\frac{12}{48},\quad \frac{103}{206}=\frac{13}{26},\quad \frac{104}{208}=\frac{14}{28},\quad \frac{134}{536}=\frac{14}{56},\quad \frac{134}{938}=\frac{14}{98},$$
$$\frac{136}{238}=\frac{16}{28},\quad \frac{154}{253}=\frac{14}{23},\quad \frac{154}{352}=\frac{14}{32},\quad \frac{165}{264}=\frac{15}{24},\quad \frac{165}{462}=\frac{15}{42},\quad \frac{176}{275}=\frac{16}{25},$$
$$\frac{176}{374}=\frac{16}{34},\quad \frac{176}{473}=\frac{16}{43},\quad \frac{176}{572}=\frac{16}{52},\quad \frac{187}{286}=\frac{17}{26},\quad \frac{187}{385}=\frac{17}{35},\quad \frac{187}{583}=\frac{17}{53},$$
$$\frac{187}{682}=\frac{17}{62},\quad \frac{195}{390}=\frac{15}{30},\quad \frac{196}{294}=\frac{16}{24},\quad \frac{196}{392}=\frac{16}{32},\quad \frac{196}{490}=\frac{16}{40},\quad \frac{197}{394}=\frac{17}{34},$$
$$\frac{198}{297}=\frac{18}{27},\quad \frac{198}{396}=\frac{18}{36},\quad \frac{198}{495}=\frac{18}{45},\quad \frac{198}{594}=\frac{18}{54},\quad \frac{198}{693}=\frac{18}{63},\quad \frac{198}{792}=\frac{18}{72},$$
$$\frac{201}{603}=\frac{21}{63},\quad \frac{201}{804}=\frac{21}{84},\quad \frac{203}{406}=\frac{23}{46},\quad \frac{203}{609}=\frac{23}{69},\quad \frac{204}{306}=\frac{24}{36},\quad \frac{206}{309}=\frac{26}{39},$$
$$\frac{234}{936}=\frac{24}{96},\quad \frac{253}{451}=\frac{23}{41},\quad \frac{264}{561}=\frac{24}{51},\quad \frac{268}{469}=\frac{28}{49},\quad \frac{275}{374}=\frac{25}{34},\quad \frac{275}{473}=\frac{25}{43},$$
$$\frac{275}{671}=\frac{25}{61},\quad \frac{286}{385}=\frac{26}{35},\quad \frac{286}{583}=\frac{26}{53},\quad \frac{286}{781}=\frac{26}{71},\quad \frac{297}{396}=\frac{27}{36},\quad \frac{297}{495}=\frac{27}{45},$$
$$\frac{297}{594}=\frac{27}{54},\quad \frac{297}{693}=\frac{27}{63},\quad \frac{297}{891}=\frac{27}{81},\quad \frac{298}{596}=\frac{28}{56},\quad \frac{301}{602}=\frac{31}{62},\quad \frac{302}{604}=\frac{32}{64},$$
$$\frac{302}{906}=\frac{32}{96},\quad \frac{304}{608}=\frac{34}{68},\quad \frac{306}{408}=\frac{36}{48},\quad \frac{352}{451}=\frac{32}{41},\quad \frac{374}{572}=\frac{34}{52},\quad \frac{374}{671}=\frac{34}{61},$$
$$\frac{385}{682}=\frac{35}{62},\quad \frac{385}{781}=\frac{35}{71},\quad \frac{392}{490}=\frac{32}{40},\quad \frac{394}{591}=\frac{34}{51},\quad \frac{395}{790}=\frac{35}{70},\quad \frac{396}{495}=\frac{36}{45},$$
$$\frac{396}{594}=\frac{36}{54},\quad \frac{396}{792}=\frac{36}{72},\quad \frac{396}{891}=\frac{36}{81},\quad \frac{398}{597}=\frac{38}{57},\quad \frac{398}{796}=\frac{38}{76},\quad \frac{401}{802}=\frac{41}{82},$$
$$\frac{402}{603}=\frac{42}{63},\quad \frac{403}{806}=\frac{43}{86},\quad \frac{462}{561}=\frac{42}{51},\quad \frac{473}{572}=\frac{43}{52},\quad \frac{473}{671}=\frac{43}{61},\quad \frac{495}{693}=\frac{45}{63},$$
$$\frac{495}{792}=\frac{45}{72},\quad \frac{495}{891}=\frac{45}{81},\quad \frac{532}{931}=\frac{52}{91},\quad \frac{536}{938}=\frac{56}{98},\quad \frac{572}{671}=\frac{52}{61},\quad \frac{583}{682}=\frac{53}{62},$$
$$\frac{583}{781}=\frac{53}{71},\quad \frac{594}{693}=\frac{54}{63},\quad \frac{594}{792}=\frac{54}{72},\quad \frac{594}{891}=\frac{54}{81},\quad \frac{596}{894}=\frac{56}{84},\quad \frac{602}{903}=\frac{62}{93},$$
$$\frac{603}{804}=\frac{63}{84},\quad \frac{682}{781}=\frac{62}{71},\quad \frac{693}{792}=\frac{63}{72},\quad \frac{693}{891}=\frac{63}{81},\quad \frac{792}{891}=\frac{72}{81}.$$

$$\frac{2\not{6}5}{10\not{6}}=\frac{25}{10},\quad \frac{2\not{9}8}{14\not{9}}=\frac{28}{14},\quad \frac{3\not{6}5}{14\not{6}}=\frac{35}{14},\quad \frac{4\not{6}5}{18\not{6}}=\frac{45}{18},\quad \frac{5\not{9}6}{14\not{9}}=\frac{56}{14},$$
$$\frac{6\not{9}5}{13\not{9}}=\frac{65}{13},\quad \frac{6\not{9}8}{34\not{9}}=\frac{68}{34},\quad \frac{7\not{3}2}{18\not{3}}=\frac{72}{18},\quad \frac{7\not{6}5}{30\not{6}}=\frac{75}{30},\quad \frac{8\not{5}4}{30\not{5}}=\frac{84}{30},$$
$$\frac{8\not{6}4}{21\not{6}}=\frac{84}{21},\quad \frac{8\not{6}5}{34\not{6}}=\frac{85}{34},\quad \frac{8\not{9}5}{17\not{9}}=\frac{85}{17},\quad \frac{9\not{6}5}{38\not{6}}=\frac{95}{38},\quad \frac{9\not{7}6}{42\not{7}}=\frac{96}{42}\cdots$$

The list could go on and on, so surely the theorem is correct. **Q.E.D.**

Of course the theorem is false, as a single counterexample demonstrates: $13/35 \neq 1/5$. Is it not a beautiful demonstration of the elegance of mathematical reasoning that all of the examples above are insufficient to establish the truth of the "theorem," yet a single, simple counterexample conclusively establishes its falsity?

A True Theorem

In elementary set theory, the binary operation of "set difference" removes those elements from the first set that are also elements in the second set: $A \setminus B = \{x : x \in A \text{ and } x \notin B\}$. Readers of the *Journal* should be familiar with the binary operation of set "intersection": the intersection of two sets is the set of all things that are elements in *both* sets: $A \bigcap B = \{x : x \in A \text{ and } x \in B\}$. Utilizing these two binary operations, the following theorem is true.

Theorem 2 *If A and B are nonempty sets, then $A \setminus B$, $B \setminus A$, and $A \bigcap B$ are pairwise disjoint.*

That is, $(A \setminus B) \bigcap (B \setminus A) = (A \setminus B) \bigcap (A \bigcap B) = (B \setminus A) \bigcap (A \bigcap B) = \varnothing$.

I have given the proof of theorem 2 as homework many times and many times I have read "proofs" that begin, "Let $A = \{1,2,3,4\}$ and $B = \{3,4,5,6\}$ ... " This is what I call "proof by example." "Proof by example" is just as invalid for establishing a true theorem like theorem 2 as it is invalid for establishing a false theorem like theorem 1. The problem with "proof by example," of course, is that one may unintentionally, or unwittingly, choose examples that are "special cases" in one or more ways. In geometry, for example, it would not be appropriate to begin a proof of a theorem about triangles with, "Let $\triangle ABC$ be an isosceles triangle with $AB = BC$ ... " Even worse—but probably even easier to prove!—would be a "proof" beginning with "Let $\triangle ABC$ be an equilateral triangle ... " We should not study the equilateral triangle to establish general propositions about all triangles because equilateral triangles possess properties and characteristics not shared by all triangles. It turns out that the equilateral triangle is to the study of geometry what the "Cobb-Douglas utility function" is to the study of consumer behavior in economics. The Cobb-Douglas functional form possesses properties and characteristics not shared by all types of functions that may be used as a utility function. This brings us to the results derived in [3].

On the Cobb-Douglas Utility Function

Suppose that a consumer has already made the labor-leisure decision discussed in [3], and that as a consequence of that decision the consumer-worker chooses to work $h = 50$ hours per week (use 60 hours for a professor and 70 hours for an untenured professor ...). If the consumer-worker earns an hourly wage of w, then the consumer can expect a weekly income of $50w$ dollars, out of which s/he chooses those amounts of food (F), clothing (C), and shelter (S) that will make her/him best off. (Let shelter include any applicable utilities; e.g., gas, water, phone, electricity, cable TV.) If the consumer faces food, clothing, and shelter prices of p_F, p_C, and p_S, respectively, and if we let I denote the consumer's weekly income ($I = 50w$), then the consumer's problem is to maximize utility subject to a budget constraint

$$p_F \cdot F + p_C \cdot C + p_S \cdot S = I.$$

Making use of the Cobb-Douglas utility function, let the consumer's utility function be $U(F,C,S) = F^{a_F}C^{a_C}S^{a_S}$. The consumer's decision can now be formulated mathematically,

$$\max_{F,C,S} F^{a_F}C^{a_C}S^{a_S} \text{ subject to } p_F \cdot F + p_C \cdot C + p_S \cdot S = I.$$

Elementary calculus generates the following expressions for the optimal quantities of food, clothing, and shelter, given in terms of the problem's parameters, namely the prices of food, clothing, and shelter, the consumer's weekly income, and the exponents in the Cobb-Douglas utility function:

$$F^* = \frac{a_F}{a_F + a_C + a_S} \cdot \frac{I}{p_F}, \quad C^* = \frac{a_C}{a_F + a_C + a_S} \cdot \frac{I}{p_C}, \quad S^* = \frac{a_S}{a_F + a_C + a_S} \cdot \frac{I}{p_S}. \tag{1}$$

These functions are what economists call "ordinary demand functions"; each of them would be plotted as a rectangular hyperbola in their own coordinate system just as in figure 1 of [3]. Expressed in a different way,

$$p_F \cdot F^* = \frac{a_F}{a_F + a_C + a_S} I, \quad p_C \cdot C^* = \frac{a_C}{a_F + a_C + a_S} I, \quad p_S \cdot S^* = \frac{a_S}{a_F + a_C + a_S} I.$$

So long as the consumer's income and the exponents in the utility function do not change, these solutions indicate that on a weekly basis, the consumer will spend identical *amounts of money* on food, clothing, and shelter, regardless of prices. Quantities are adjusted whenever prices change, but in a very peculiar way. If food prices rise sharply, the quantity of food purchased is adjusted. This we expect. However, the quantities of clothing and shelter are not altered by a consumer with a Cobb-Douglas utility function! This is *not* how we tend to behave. If food becomes more expensive and our budget becomes more strained than before, many of us will tend to buy clothes less frequently—decreasing the quantity of clothing purchased and the amount of money spent on clothing—even though the price of clothing has not changed. Similarly, many of us will tend to respond to sharply higher food prices by reducing the quantities of, and the amount of money spent on, those shelter-related items—we may push the thermostat lower in the winter and higher in the summer, reduce water usage, phone-calling, even disconnect the cable under really drastic circumstances—even though the prices of these shelter-related items have not changed.

Is there an obvious correspondence between the ordinary demand functions above and the results presented in [3]? Absolutely! We could summarize the three functions in equation (1) by using X as an index:

$$X^* = \frac{a_X}{a_F + a_C + a_S} \cdot \frac{I}{p_X}, \ (X = F,C,S). \tag{2}$$

Since the labor-leisure choice has already been made, the total amount of money available to spend on X is I. If we want one more unit of X, then we must give up an amount of money equal to its price. The unit price of X is p_X, so if we take the total amount of money, I, and divide by the unit price of X, then I/p_X represents the largest feasible value that X can take. This quantity is multiplied by the share of X's exponent relative to the sum of all of the exponents.

Now consider the optimal quantity of labor hours worked given in [3] *when the amount of non-labor income is zero*:

$$h^* = \frac{\alpha}{\alpha+\beta}T, \tag{3}$$

and use $L+h=T$ to write the consumer's optimal quantity of leisure:

$$L^* = T - h^* = \frac{\beta}{\alpha+\beta}T. \tag{4}$$

(We convert hours-worked to hours of leisure because we always want to deal with "goods," those things that make us happier when we have more of them. Although we all can enjoy our jobs, between labor and leisure, leisure is the "good.") Does equation (4) look like equation (2)? It certainly does! In equation (4), T is the largest feasible value that L can take. This quantity is multiplied by the share of L's exponent relative to the sum of all of the exponents, i.e., $\beta/(\alpha+\beta)$.

If you are not convinced, then let us look at this labor-leisure problem from a more economic perspective.

If the consumer has no non-labor income ($M = 0$ in [3]), what is the largest feasible value that C can take? If the worker takes no leisure, $L = 0$, choosing to work during all of the T hours available, then the worker's time worked is $h = T$ and the available income is wT; consequently, the consumer can have current consumption of $C = wT$.

Now suppose that the worker/consumer chooses to have a single hour of leisure. Since the sum of leisure hours and labor hours must be the total time available, $L+h=T$, $L=1$ implies that $h=T-1$. How much current consumption can the consumer have while taking one hour of leisure? Clearly, $C = wh = w(T-1) = wT - w$. Comparing the current consumption possible with $L=1$ and $L=0$, we find that current consumption is reduced by w, the wage rate. As a result, economists say that the "cost" of an hour's leisure is the labor income forgone, namely the wage rate. This can be interpreted as a economic formalization of the "time is money" phrase we hear so often.

If time is money, then we can convert to money terms the T hours available for allocation between labor and leisure. Using the wage rate as the "price" or value of an hour's time, the worker/consumer has an amount of money wT that s/he can spend on current consumption or leisure. Thus, if no leisure is taken, all wT dollars can be spent on current consumption, affording an amount of current consumption equal to wT. If no current consumption is taken, all wT dollars can be spent on leisure. But an hour of leisure costs w, so the largest feasible value that L can take is the total amount of money available, wT, divided by the price of one unit of L, namely w. Therefore, the largest feasible value that L can take is $wT/w = T$, which we already knew, but now we've worked this out using a "price" for an hour of leisure.

The upshot is still that h^* (and L^*) are independent of w. To what, then, do we attribute this seemingly paradoxical result? The Cobb-Douglas functional form! To see that the Cobb-Douglas utility function is to blame, use the following utility function,

$$U(C,L) = \left[aC^{-\rho} + (1-a)L^{-\rho}\right]^{-1/\rho}, \quad 0 < a < 1,\ \rho \geq -1, \tag{5}$$

and follow the same procedure as in [3]. If one performs the necessary calculus, one obtains the following equivalent expressions for the optimal quantity of leisure:

$$L^* = \frac{w \cdot T}{\left(\frac{a}{1-a}\right)^{1/(\rho+1)} \cdot w^{1/(\rho+1)} + w} = \frac{(1-a)^{1/(\rho+1)} \cdot w \cdot T}{a^{1/(\rho+1)} \cdot w^{1/(\rho+1)} + (1-a)^{1/(\rho+1)} \cdot w}$$

$$= \frac{(1-a)^{1/(\rho+1)} \cdot T}{a^{1/(\rho+1)} \cdot w^{-\rho/(\rho+1)} + (1-a)^{1/(\rho+1)}}.$$

Try as we might, w cannot be eliminated from L^*. This implies that L^* depends on w, and so, therefore, will h^* depend on the wage rate. Thus, it is the Cobb-Douglas utility function that is to blame for the result in [3] because the utility function in equation (5) is more general than the Cobb-Douglas function. Indeed, a Cobb-Douglas utility function is a special case of the utility function given above; it is found as the limit of the above function as ρ tends to zero:

$$\lim_{\rho \to 0} \left[aC^{-\rho} + (1-a)L^{-\rho}\right]^{-1/\rho} = C^a L^{1-a};$$

see Problem #703 in the *Journal* **9**(2) (Spring 1990, p. 134). Note also that in the limit as ρ tends to zero, L^* approaches $(1-a) \cdot T$, which is independent of w, the result derived in [3].

Expunge the Cobb-Douglas?

Economists as a group wish to explain the behavior we all observe in markets. Since most of us—as buyers of goods-and-services and sellers of labor services—do not behave in ways that can generally be explained by the Cobb-Douglas utility function, should we hit the textbooks and blot out all mentions of this particular function? Of course not!

The Cobb-Douglas utility function is useful for illustrating many economic principles in simple ways. In a real sense, it is the best function on which beginning students can "cut their analytical teeth." It is not a particularly realistic function to use, but in the early stages of study in any discipline, realism is frequently sacrificed, if temporarily, as a cost willingly paid to acquire intuition and understanding. As we move progressively deeper in the study of any discipline in pursuit of realism and the desire to explain the world around us, we generally begin dropping the assumptions made for convenience and simplicity in the earliest stages of study. To cite one obvious example, consider the study of "mechanics" in physics classes. In the beginning we are allowed—even told—to ignore friction. Is this realistic? Of course not. Indeed, each one of us has scraped elbows and knees and received burns from carpeting and gym floors; from an early age we have painful real-world experience with friction, and any theory of physics that ignores friction is destined to be imprecise. But this is not wrong, as long as we are aware of the limitations of our assumptions.

Conclusion

In [3] the author asks several questions. One of them, however, is this: "is the Cobb-Douglas model at fault?" Clearly, the answer is a resounding "yes!" The Cobb-Douglas utility function is to the economic analysis of consumer behavior what the equilateral triangle is to geometry: it is a "special case" in the extreme. Any attempts to make general statements about consumer behavior based on results obtained from the Cobb-Douglas utility function would be about as useful as a flight plan for bringing a space shuttle back through the atmosphere without accounting for friction.

Finally, can the Cobb-Douglas function's complicity come as a total surprise? Certainly not, for between mathematician Charles W. Cobb and economist Paul H. Douglas, one of them was a politician (three-term U.S. Senator)!

References

[1] Dudley, Underwood, Editorial Comment, *Pi Mu Epsilon Journal* **10**(2) (Spring 1995), 130–133.

[2] Mitchell, Thomas, Discouraging "Proof by Example," mimeo, Department of Economics, Southern Illinois University at Carbondale (1991). This paper can be downloaded off the World Wide Web; point a Web "browser" to:
`http://econwpa.wustl.edu/eprints/ge/papers/9506/9506001.abs`.

[3] Morrill, John E., How Economists Use Mathematics To Show Why Some People Work So Much for So Little, *Pi Mu Epsilon Journal* **10**(4) (Spring 1996), 269–272.

The author holds a bachelor's degree in mathematics and graduate degrees in economics; he is an associate professor of economics at Southern Illinois University at Carbondale, and a father to two daughters who cannot believe he reads the *Pi Mu Epsilon Journal* "for pleasure"! Send electronic correspondence to `tmitch @ siu.edu`.
