

## Papers

5. Teaching Business Cycle Dynamics: A Comparison of Graphs and Loops

# Teaching Business Cycle Dynamics: A Comparison of Graphs and Loops

I. David Wheat, Jr.

Senior Lecturer in System Dynamics

University of Bergen

## Abstract

*This article compares the pedagogical value of an AS/AD graph and system dynamics-based feedback loops. Both methods were employed in a controlled experiment designed to assess their effectiveness when teaching the sticky price theory of business cycles to undergraduates. The performance of the students using the feedback loops was significantly higher than the performance of those using the graph.*

*JEL: A22, C91, E32*

*Key words: education, experiment, feedback, macroeconomics, model*

This paper presents the results of an experiment comparing two methods of teaching undergraduates about a particular theory of business cycles. One method utilized the standard macroeconomics textbook version of the aggregate supply and demand (AS/AD) graph. The other method relied on feedback loop diagrams that are commonly used when explaining to non-specialists the structure and behavior of a complex system. Following an elaboration of the rationale for the experiment in section 1, the second section outlines the feedback method of teaching macroeconomics. Section 3 describes the experiment, and the last two sections present and discuss the findings.

## 1. Problems with Graphs in Economics

At the undergraduate level, instructors rely almost exclusively on graphs to explain and demonstrate economic models (Kennedy 2000). The magnitude of the modern instructor's penchant for graphs is suggested by Cohn et al. (2001), who found that graphs in popular textbooks outnumber graphs in early 20th century texts about ten to one and sometimes twice that ratio. (An exception is Kennedy's (2000) text, which is highly selective and judicious in its use of graphs.) The increasing textbook reliance on graphs could be problematic, however, for two reasons. First, the instructional value added by graphical representation of economic behavior may be lower than previously thought, presumably because students find it difficult to interpret the graphs. The second problem concerns the predominant graphical model of modern undergraduate macroeconomics textbooks—the AS/AD model. Even if students can interpret that graph, they may get a misleading interpretation about the dynamic behavior of the economy.

The first problem—doubts about the efficacy of graphs as a teaching tool—was highlighted by Cohn et al. (2001). Cohn and his colleagues were the first to study the learning impact of graphs in introductory economics courses. In one experiment, they found graphs to be no more effective than verbal instruction alone. Another experiment found that students in a graph-supplemented lecture actually showed *less* improvement than those in a lecture-only session. These findings suggest that the pedagogical value of graphs in undergraduate settings may have been overrated in the past. If so, that bodes ill for current and future undergraduate instruction, given the dominance of graphical instruction in the classroom and the proliferation of graphs in textbooks.

The second problem—the misperception of the economy’s dynamic behavior due to a *correct* interpretation of a misleading graph—has been forcefully argued by Colander (1991, 1995). He (p. 106) calls the textbook AS/AD model “confusing and logically flawed, ... a crutch ... that encourages students to understand incorrectly how aggregate disequilibrium forces operate.” Even worse (p. 108), the adjustment process they see is “...one which is superficially satisfying to students but fundamentally wrong.” Colander (1991, p. 105) concludes that it should “...never [have become] the central focus of what is taught to undergraduates.”

In principle, graphical comparative statics would not be used to illustrate dynamics—the transition process from one equilibrium condition to another over some time period. Nevertheless, common classroom—and textbook—practice reflects an implicit assumption that graphical representation of two settled equilibrium conditions is a pedagogically useful way to engage in a discussion about the settlement process in between. For more than a century, economics instructors have used chalkboards to shift supply and demand curves and trace price movements, aiming to give students a visual impression of dynamics. In the computer age, it is common practice to use slideshow software to animate static graphs, making curves “move” from one equilibrium point to another (without any economic structural reason inherent in the slideshow software). Point-and-click may replace chalk-and-talk, but the increasingly animated static graphs may be “superficially satisfying to students” while conveying the wrong message about the transition process between equilibria.

The traditional justification for the graphical approach to economic dynamics is that alternative methods require a level of mathematical preparation that is uncommon among undergraduates. However, now that Cohn has raised questions about the efficacy of graphical instruction in general, one wonders whether static graphs—even when animated and even if accurately reflecting disequilibrium forces—can foster the temporal reasoning skills necessary for grasping dynamics. Arguably, economics students need the same kind of spatial-temporal reasoning abilities that Shaw (2000) cites as important for math and science students. The Colander critique (1991, p. 109) focuses on the misuse of a static graph to explain dynamic phenomena and emphasizes that any “... final equilibrium depends on the process of getting there” (e.g., the supply response to expectations of changing demand, which sets in motion a feedback process that affects actual and expected demand in the future).<sup>1</sup> Colander’s criticism addresses a weakness in the AS/AD model that is the very strength of the feedback method.

## 2. The Feedback Method

Over the past five years, a new method has been developed to teach macroeconomic dynamics without requiring students to manipulate equations or rely exclusively on static graphs.<sup>2</sup> Drawing on an underutilized historical thread of feedback thinking in economics (Richardson 1991) and a central tenet in system dynamics, it is called simply the *feedback method*. It utilizes the diagramming and simulation features of system dynamics stock-and-flow modeling to make dynamics accessible to economics students who lack the

---

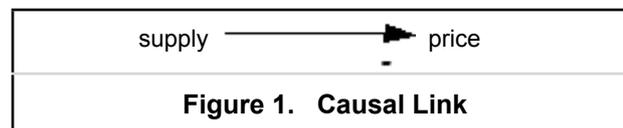
<sup>1</sup> On the issue of misperception of feedback, see Moxnes (1998) and Sterman (2000).

<sup>2</sup> The macroeconomics principles course is delivered via the Internet to students enrolled at Virginia Western Community College in Roanoke, Virginia.

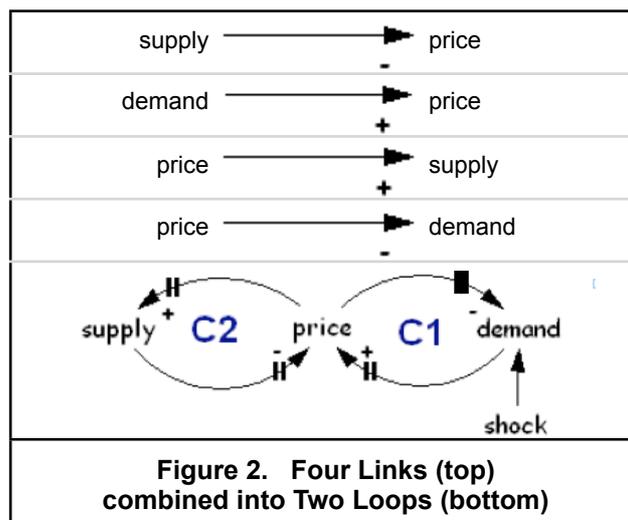
mathematical training normally considered a prerequisite for such access. Using system dynamics to study economics is not new, nor is using system dynamics methods to teach economics. A list of system dynamics pioneers in economics includes J.W. Forrester (1968, 1976, 1979), Mass (1975, 1980), N. Forrester (1982), Low (1980), J. Forrester, Mass, and Ryan (1980), Sterman (1985), and Radzicki (1993), among others. To my knowledge, however, the feedback method described here is the first use of system dynamics concepts as the central organizing method for an entire undergraduate macroeconomics course. Radzicki (Worcester Polytechnic Institute) and Yamaguchi (Doshisha Business School) have developed complete SD-based curricula for graduate economics courses they teach in the US and Japan, respectively. System dynamics textbook discussions of feedback are accessible in Richardson and Pugh (1989), Ford (1999), and Sterman (2000).

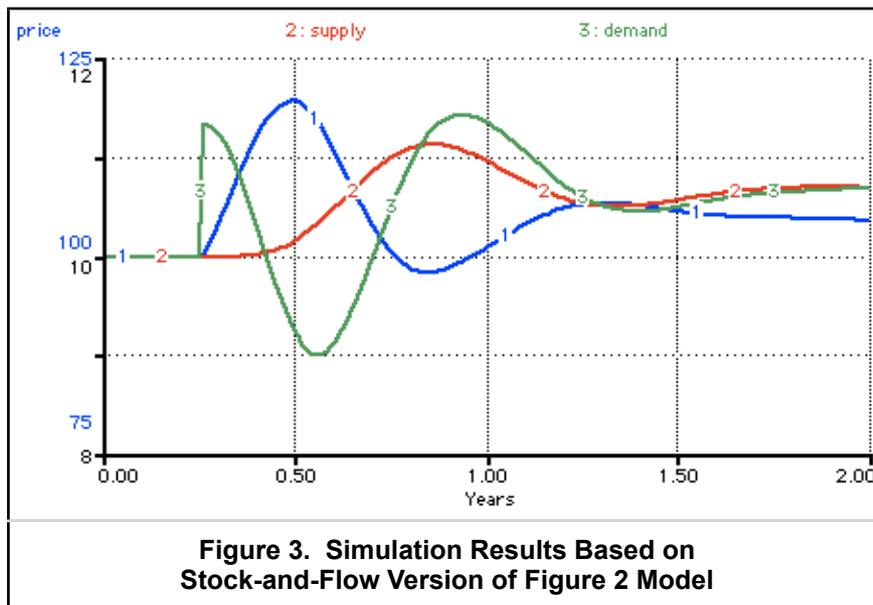
When working with stock-and-flow models of complex systems such as an economy, the relative simplicity of the feedback loop often makes it a more useful tool for communicating with non-specialists (e.g., undergraduates). In the experiment reported here, one of the instructional methods relied on a pair of feedback loops to explain the dynamics of the sticky price theory. The slides used for that method are in appendix B, and they may provide the reader with a sufficient overview of the feedback loop approach. Nonetheless, we present a brief introduction to feedback loop terminology and diagramming methods.

A *feedback loop* is a closed loop of mutual causation that occurs over time. A loop contains at least two causal links, which are *ceteris paribus* hypotheses about cause-and-effect. A causal link is a word-and-arrow diagram similar to those that appear in some economics textbooks such as Kennedy (2000) and Sexton (2002). The causal link in Figure 1, for example, expresses the hypothesis that supply influences price.



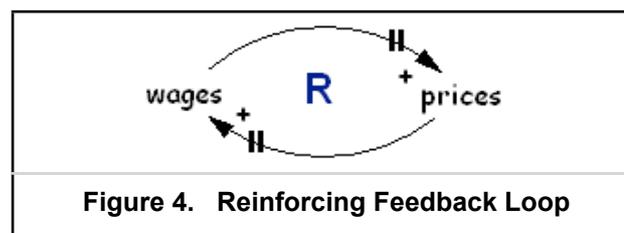
The minus sign (-) in Figure 1 means that price would *decrease* if supply *increased*. If the supply *decreased*, then price would *increase*. In other words, the negative polarity is suggestive of two variables moving in opposite directions. If the link had been labeled with a plus sign (+), that would suggest two variables moving in the same direction. Connecting links to form loops is a straightforward process. The four links in Figure 2, for example, can be combined into two feedback loops C1 and C2, where the cross marks ( || ) indicate time delays.





After an exogenous shock to demand, the hypothesized feedback structure gives rise to the simulated behavior shown in Figure 3. To interpret that behavior, use the feedback structure in Figure 2. Begin by assuming that a permanent exogenous demand shock disturbs the equilibrium. After suppliers take time to evaluate the reliability of the signal that demand has increased, price would rise.<sup>3</sup> The rising price would, in turn, put downward pressure on demand, but the full effect would occur gradually. Meanwhile, suppliers respond to the rising price by stepping up production, but it takes time to organize the requisite factors of production. When supply eventually responds, that puts downward pressure on prices, but with a delay. The damped oscillatory behavior and the amplitude and period for each curve depend on parameter assumptions for delay times and price elasticities.

When a feedback loop contains an *odd* number of negative signs, it will counteract or negate a previous trend. Such loops (e.g., C1 and C2 in Figure 2) are called *negative* or *counteracting* or *balancing* loops. Each term has the same meaning, and they are used interchangeably. The other type of feedback loop is called *positive* or *reinforcing*; again, the terms are synonymous. Figure 4 illustrates the reinforcing loop implicit in the familiar wage-price spiral hypothesis. A “walk-around-the-loop” shows that it feeds on itself and reinforces an initial trend. Confirmation comes from counting an *even* number of negative links (namely, zero) around the loop.



<sup>3</sup> In Blinder’s (1997) survey of 186 companies that represented about 85 percent of the private, non-farm GDP, he found “the typical lag of a price change behind a shock to either demand or cost is about three months.”

## Previous Comparisons of Graphs and Loops

Patterned after the Cohn study, Wheat (2007d) reports on an experiment where students learn about GDP with and without a visual aid. Instead of a graph, however, the supplemental teaching tool was a stock-and-flow feedback diagram. Students aided by the *feedback diagram* outperformed those who received mere textual instruction, in contrast to the Cohn findings that *static graphs* did not add value to verbal instruction. Taken together, the two studies could be interpreted as an *indirect* comparison of the instructional value added by graphs and feedback diagrams. That motivated the experiment reported here, which makes a *direct* comparison of graphs and feedback loops as supplementary tools when teaching undergraduates.

The results of two other experiments were favorable to the feedback method and may shed light on the similar outcome reported here. The two studies examined student preferences for using graphs or loops to study dynamics. In Wheat (2007b), students had to select teaching tools to facilitate a hypothetical tutoring task. A *single slide* featuring an AS/AD graph (similar to the one in Figure A14 in the appendix) and a *single slide* featuring a feedback loop diagram (similar to the one in Figure B14) competed for selection when the student task was to explain the sticky price theory of business cycles. In other words, the students saw only the final slide extracted from a full presentation, and it was their instinctive response to that single slide that revealed their preference. A significant majority of the students (70 percent) showed a preference for the feedback loop method.<sup>4</sup> In Wheat (2007c), students received an entire slideshow presentation similar to the one featured here, but they received *both* the graphical and feedback loop versions. Regardless of the sequence of instruction, the students preferred the feedback method by a significant margin.<sup>5</sup> In Wheat (2007b), the students *explained* their preferences. The explanations of the majority—those favoring the loop diagram—stressed the importance of visualizing a real-world process. They also emphasized the mutual causation implicit in the feedback loop diagram.

All of these experiments involved different samples of students. Therefore, we do not mean to imply a longitudinal link between preference in the two previous experiments and performance here. Nevertheless, the education literature is suggestive of preference/performance links. Experimental findings by Nowaczyk et al. (1998) and Sankaran et al. (2000) show a positive relationship between preference and performance among undergraduates. Closely related to learning preferences are self-described learning styles and expressions of self-efficacy in learning situations. Terry (2001) documents the correlation between self-described learning styles and performance, and Stevens et al. (2004) show that a student's sense of self-efficacy has a significant association with performance. It is certainly conceivable—and worthy of further research—that at least part of the reason for the better performance by Group L is student preference for macroeconomic models that promote visualization of processes of change over time in the economy.

---

<sup>4</sup> The t-value was 2.86 with 45 degrees of freedom, and  $p < 0.01$  (Wheat 2007b, p. 11).

<sup>5</sup> In a paired t-test with 36 degrees of freedom,  $t = 5.19$  and  $p < 0.0001$  (Wheat 2007c, p. 10).

### 3. The Experimental Design

In this experiment, students were divided into two groups, and each student worked alone at an individual computer, using slideshow software to access the instructional material. The first ten slides were identical for both groups and included time series graphs to illustrate stylized and historical business cycle behavior in the United States. Slide 9 (reproduced in Figure 5) emphasized the multiplicity of theories about fluctuations in the economy.

*Business cycles occur for several reasons. There is no single cause. Many theories have been suggested.*

*You will now read a brief summary of the “sticky price theory” of business cycles, and then study a diagram of how the theory works.*

**Figure 5. Information Provided to Students in Both Groups (slide 9)**

Slide 10 (Figure 6) provided a very brief summary of the sticky price theory, using text only, adapted from DeLong (2002), Mankiw (2002, 2004), and Hall & Taylor (1997).

**Definitions:**

- *Sticky Prices: prices that adjust slowly to changes in the demand for goods & services.*
- *Business Cycles: up-and-down pattern of short-run economic growth that fluctuates around the long-run trend.*

**Building Blocks of the Theory:**

- *If sales decline, business firms cut back on employment & production before cutting prices.*
- *The employment & production cutbacks reduce household wages and spending, causing even more cuts in sales and production. The economy’s growth pattern slips below its long-run trend.*
- *After firms finally cut prices, sales start rising. Employment & production rebound sooner than prices. The economy’s growth pattern eventually goes above its long-run trend.*
- *After a few years, the up-and-down pattern stabilizes near the long-run trend unless additional disturbances occur:*

**Summary of the Theory:**

*Sticky prices can cause the economy’s growth pattern to fluctuate around its long-run trend. In other words, sticky prices contribute to business cycles.*

**Figure 6. Information Provided to Students in Both Groups (slide 10)**

Group G then encountered five slides explaining how to read and interpret an aggregate supply and demand (AS/AD) graph, followed by nine slides that used the graph to illustrate the sticky price theory of business cycles. The corresponding slides for Group L explained how to read and interpret reinforcing and counteracting feedback loops, followed by instruction that used those loops to illustrate the sticky price theory. The full instructional content of both methods is available in the appendices, but a summary is presented here. After the opening slides presented group G with a tutorial on AS/AD graphs, the remainder of the slideshow illustrated the medium-term response of price, sales, and GDP to a sudden

exogenous reduction in aggregate demand. A new year with a new “equilibrium” point was presented with each slide. Figure 7 displays the AS/AD graph seen after five years of lagged price adjustments.<sup>6</sup> Beginning with the decline in AD in year 1, the slides were animated to give the impression of dynamics. Each slide showed each year’s price and quantity adjustment. An overshoot occurred in year 5, followed by attainment of equilibrium soon thereafter. The small time graph tracked movement of GDP, sales, and price each year.

Group L received a tutorial on reinforcing and counteracting feedbacks in the opening slides, and the remainder of the slideshow illustrated the medium-term response of price, sales, and GDP to a sudden exogenous reduction in demand.<sup>7</sup> Figure 8 displays the feedback loop diagram and time series graph seen more than five years after the demand shock. Following the decline in sales in year 1, the narrative on each slide traced the interaction of the “turnabout” counteracting feedback loop and the “boom or bust” reinforcing loop. The small time series graph tracked simulated movement of GDP, sales, and price each year. GDP eventually approached its original trend, but prices remained lower.

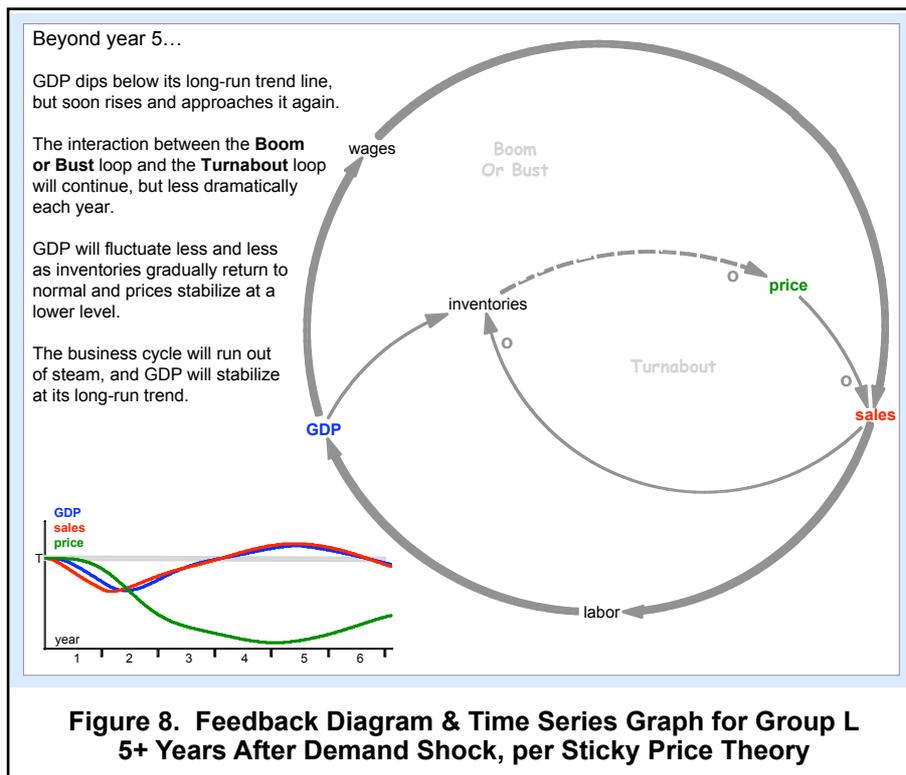
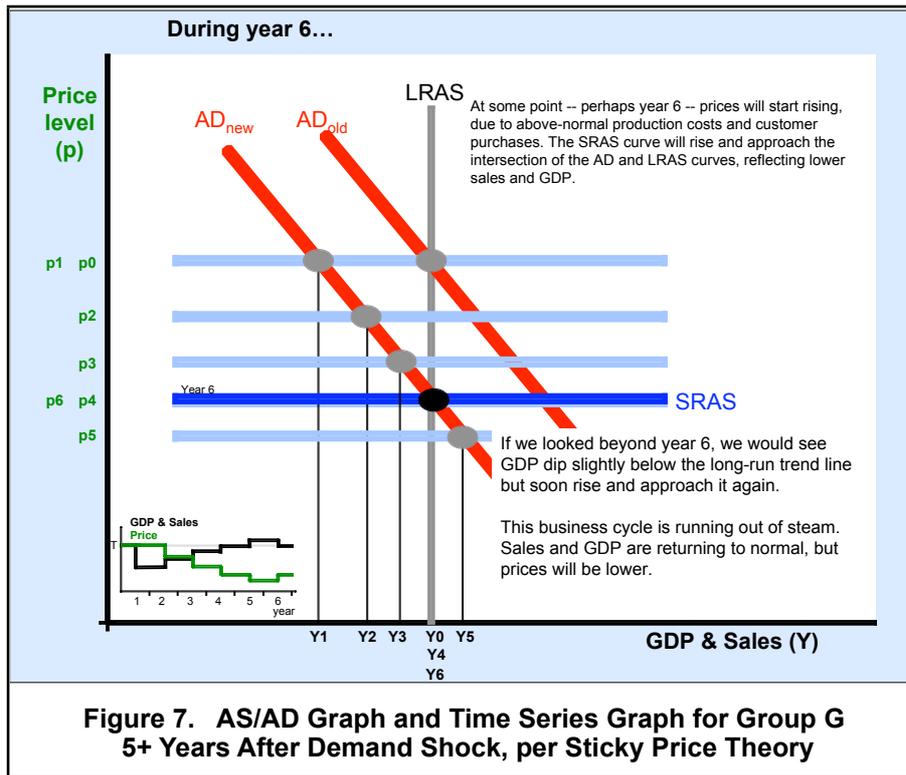
Both instructional methods contained the same number of slides (37), and the mean time for the experiment was about 17 minutes for Group G and 18 minutes for Group L. A t-test confirmed that the difference between observed mean times was not statistically significant. Thus, the results cannot be attributed to one group having more time to learn or having access to more material during the experiment. An effort was made to eliminate any content differences in the experiment that would be relatively easy to eliminate in practice. That is the reason that both methods contain time series graphs, even though they are rare in textbook presentations of the AS/AD model.<sup>8</sup> A graph of simulated behavior over time is an integral by-product of system dynamics models, on which feedback loops are based. In contrast, pure static graphs do not acknowledge the passage of time between equilibria. However, when the static graph format is used in an attempt to convey dynamics (e.g., in the AS/AD model), the subscripted variables denote separate time periods. In that case, a time series graph implicit in the static graph could be a useful pedagogical device. Since a graph of behavior over time is not an essential distinction between the AS/AD model and system dynamics-based feedback loops, it was included in both instructional slide shows. In that way, experimental differences would more likely reflect central rather than peripheral distinctions between graphical and feedback modes of instruction. Since students in both groups received virtually identical visual cues about business cycle patterns over the same time horizon, such cues cannot account for the difference in performance between the two groups.

---

<sup>6</sup> The AS/AD graph is adapted from Hall & Taylor (1997, chs. 8-9), Mankiw (2002, ch. 9), and Mankiw (2004, ch. 20). Mankiw (2002) uses horizontal SRAS curves, while Mankiw (2004) uses upward sloping SRAS curves. Hall & Taylor use horizontal lines that imply a price-determining intersection with the AD curve, but do not explicitly refer to the lines as SRAS curves. The full instructional content is contained in appendix A.

<sup>7</sup> The full instructional content is contained in appendix B.

<sup>8</sup> An exception is Hall and Taylor (1997, pp. 216 and 219). Even there, however, the AS/AD graph and its implicit times series graph are not on the same page.



The most subjective instructional content decisions concerned the annotations accompanying each slide. When explaining adjustments in the supply curve on the AS/AD graph, an effort was made to justify those adjustments in business and consumer decision-making terms. The slide reproduced in Figure A13, for example, attributes the eventual rise in price to rising consumer demand and rising production costs. Of course, the graph does

not mention costs, and the reference to costs was included in the annotation solely for the purpose of giving additional real-world rationale for the movement of the SRAS curve. It was, in essence, an appeal to a force that was either exogenous or was triggered by unacknowledged feedback from the AS/AD model; either way, it was giving that model the benefit of the doubt. On the other hand, the feedback loop diagram includes several real variables, and the slide annotations were couched in terms of those variables. There was no appeal to explanations outside the model. If the annotations were biased at all—in the sense of relying on exogenous explanations—an argument could be made that they favored the graph and not the loops.

In summary, the design of the instructional content was motivated by a commitment to eliminate any differences between the two methods that were not central to their respective underlying models. The specific content decisions discussed here all cut in the same direction—making sure that students receiving the graphical instruction were *not* disadvantaged by the design of the experiment. Whether that goal was achieved is an open question. However, there was an equal number of slides even though the number of students with economic training vastly outnumbered those with system dynamics training. Moreover, a time series graph was included in both methods, even though that is not common in textbook presentations of the AS/AD model. Finally, the annotations accompanying the slides attempted to give the students in Group G a real-world rationale for the movement of the SRAS curve. Ultimately, of course, the reader studying the slides in appendices A and B will have to decide whether there was a level playing field.

**Test Instrument.** After students completed their learning task, they answered several test questions designed to measure knowledge and understanding (Figure 9). One purpose of the range of questions, in addition to trying to uncover a sense of dynamics, was to ask questions requiring more than mere factual recall. Ostensibly, Q1-Q2 gauge students' *comprehension* (as defined by Bloom, 1956). However, both questions relate to fundamental hypotheses of the sticky price theory and, in fact, could have been recalled from memory—the *knowledge* level in Bloom's taxonomy—since the text-only version of the theory presented the “answer” in slide 10 (Figure 6). Either characterization puts those two questions near the bottom of the hierarchy of cognitive skills. That is not to suggest that such skills should not be valued; arguably, they are foundational to so-called higher order thinking skills.

The next five questions (Q3-Q7) go beyond knowledge and comprehension and require various combinations of higher order cognitive skills. Q3 requires prediction (an *application* skill). Q4-Q6 require inference, which Blooms considers *analysis*, although the tasks require seeing how a system works rather than taking it apart. Inference may not fit Bloom's concept of *synthesis*, but it seems to fit the plain meaning of that term better than it fits *analysis*. Unquestionably, however, Q3-Q6 require skills above and beyond factual recall or even restatement of a theory. Q7 addresses a fundamental point of the sticky price theory—quicker price adjustments restore stability sooner after a shock to the economy. Getting the correct answer (for the right reason) would seem to be an exercise in *application* of a theory, which is above *knowledge* and *comprehension* in Bloom's rankings.

The last question (Q8) was designed to probe student understanding of *why* the implicit economic agents in the model were generating the observed behavior. In Bloom's

terms, answering Q8 was probably an exercise in *comprehension* and/or *analysis*. It requires a generalization, but to be done properly it requires searching for supporting data. However, one could argue that a correct interpretation of the AS/AD model is that searching for equilibrium *is* the motivating influence on behavior. Assessing the responses to Q8 is not so much a comparison of “right” or “wrong” answers as it is merely a comparison of answers.

Question Stem	Multiple Choices			
Q1. Each year, according to the diagram, a change in	(a) sales affected prices first and GDP later.	(b) sales affected GDP first and prices later.	(c) sales affected GDP and prices at the same time.	(d) prices affected GDP first and sales later.
Q2. After the initial drop in sales, GDP	(a) rose before falling and then fell without stopping.	(b) fell and sales dropped even more before both rose.	(c) fell and then rose as soon as prices started falling.	(d) each of the above is correct.
				(e) none of the above is correct.
Q3. According to the diagram,	(a) a change in sales will affect GDP and prices.	(b) a change in prices will affect sales and GDP.	(c) a change in GDP will affect prices and sales.	(d) each of the above is correct.
				(e) none of the above is correct.
Q4. After the initial drop in sales and GDP, the fall in prices indicates that	(a) GDP was below its long-run trend.	(b) GDP was above its long-run trend.	(c) sales dropped more than GDP.	(d) sales dropped less than GDP.
Q5. Suppose sales dropped suddenly and prices adjusted slowly. The diagram used in this activity would predict the behavior in graph	(a) A	(b) B	(c) C	(d) D
	see appendix for diagrams A - D			
Q6. Assume sales suddenly increased, followed by a rise in GDP. If prices rose, that would indicate	(a) sales increased less than GDP.	(b) sales increased more than GDP.	(c) GDP was above its long-run trend.	(d) GDP was below its long-run trend.
Q7. Suppose each time sales changed, prices adjusted a month later instead of a year later. Then	(a) GDP and employment would rise and fall more.	(b) GDP would rise and fall more, but employment would rise and fall less.	(c) GDP would rise and fall less, but employment would rise and fall more.	(d) GDP and employment would rise and fall less.
Q8. According to the diagram, decisions about employment, production, and pricing are based on knowledge of	(a) random events.	(b) long-run trends.	(c) business conditions.	(d) equilibrium requirements.
<b>Figure 9. Test Questions for Both Groups (correct answers shaded)</b>				

The effectiveness of each instructional method was operationally defined in terms of accurate student responses to post-experiment test questions. For several reasons, a pre-test (for benchmarking) was not used. First, the sticky price theory was considered a sufficiently obscure topic that the subjects in the experiment (even those who had some prior economics education) would have no prior knowledge of the theory; essentially, a zero pre-test baseline was assumed for all students. Second, the learning assessment focused on understanding the

structure and behavior of each method's economic model, which again was assumed to be virgin territory. The questions requiring inference and interpretation would have been meaningless out of context in a pre-test setting. Finally, pre-tests can be problematic if they heighten awareness of important concepts prior to instruction (the "pre-test effect") and, therefore, blur distinctions between the impacts of the instructional treatments.

Let  $P_G$  and  $P_L$  represent the mean percentage of correct answers in groups G and L, respectively, on questions 1-8. The null hypothesis for the full test was

$$H_0: P_G = P_L$$

and

$$H_{0Q}: P_{GQ} = P_{LQ}$$

for individual questions ( $Q = 1...8$ ). The full test means were compared with a paired t-test at the 0.05 level for significance. For each individual question, the proportion of correct answers in each group was compared, using a t-test at the 0.05 significance level.

**Sample Selection and Characteristics.** The experiment was conducted in May 2005 with 117 volunteers at two sites: 33 undergraduate macroeconomics students at Virginia Western Community College in Roanoke, Virginia; and 84 calculus and advanced algebra students at Wilson High School in Portland, Oregon.<sup>9</sup> The results were pooled for analysis.

The students were randomly assigned to two groups—56 in group G that received *graphical* instruction in comparative statics and 61 in Group L that received feedback *loop* instruction based on system dynamics. Based on course grades (in economics, calculus, and advanced algebra), it can be inferred that the two groups were comparable in terms of learning potential. The mean course grades were estimated at 79.8 and 83.4 for the G and L groups, respectively, and a t-test confirmed that the three-point difference in means was not statistically significant.<sup>10</sup> However, twenty-five of the Oregon students were not included in the course grade comparison because they used fictional names when they participated in the experiment, and records of their real names (and grades) were later unavailable. Those students were equally distributed between groups G and L, however, and it was assumed that their course grades would not significantly change the difference between mean course grades for the two groups. Nevertheless, the claim that the two groups had similar learning potential has to be qualified because of the "missing" students during the t-test.

The Virginia students had just enrolled in a summer session and had not received instruction prior to the experiment. Eleven had taken a prior economics course. Forty-two of the Oregon math students had taken a high school course in economics. Overall, therefore, 45.3 percent of the students had some economics education. In contrast, only 7.7 percent (9

---

<sup>9</sup> I am grateful to Diana Fisher for administering the experiment at Wilson and to her students for participating. An award-winning mathematics teacher, she has written two books on modeling, including one on teaching algebra and calculus with system dynamics (Fisher, 2001 and 2005).

<sup>10</sup> Statistical significance test results:  $t = .778$ ,  $d.f. = 91$ , and  $p = .44$ , given the null hypothesis that the means were equal.

students, all from Oregon) had received training in system dynamics. Any advantage due to prior experience should favor the graphical method.

Since 72 percent of the subjects in the experiment were high school students, there is an issue of external validity. Clearly, the sample was not drawn from a population of macro undergraduates. However, almost 30 percent of the students were enrolled in a macro course at the time of the experiment. An equal percentage were college-bound calculus students who had taken a high school course in economics. The remaining 40 percent were younger advanced algebra students, most of whom would be in college within two to three years and, given their math aptitude, might be more likely to take an economics course than the average undergraduate.

#### 4. Results

For both groups combined (117 students), only 40.5 percent of the responses were correct. Such a low accuracy rate is not surprising since the complex instructional content was presented in a format designed for brevity. Most students completed their self-paced instruction within twenty minutes. The performance was not the same in both groups, however. As shown in Figure 10, averaging the results over the eight questions yields a mean accuracy score of 46.5 percent for Group L, compared to 33.9 percent for Group G, a statistically significant difference. Therefore, the full test null hypothesis was rejected.<sup>11</sup>

	P <sub>G</sub>	P <sub>L</sub>	p
Q1-Q8	33.9%	46.5%	<b>.028</b>
<b>Figure 10 Mean Percentage of Correct Answers for Full Test, by Instructional Method</b>			

Figure 11 presents the results at the individual question level.

	P <sub>GQ</sub>	P <sub>LQ</sub>	p
Q1	67.9%	63.9%	.660
Q2	37.5%	36.1%	.870
Q3	42.9%	73.8%	<b>.001</b>
Q4	12.5%	29.5%	<b>.025</b>
Q5	26.8%	52.5%	<b>.004</b>
Q6	25.0%	32.8%	.360
Q7	33.9%	37.7%	.670
Q8	25.0%	45.9%	<b>.018</b>
<b>Figure 11 Percentage of Correct Answers on Individual Questions, by Instructional Method</b>			

<sup>11</sup> The paired t-test yielded a value of 2.76 with 7 degrees of freedom and  $p < 0.028$ .

On four of the eight questions (Q3, Q4, Q5, and Q8), the null hypothesis was rejected, as Group L significantly outperformed Group G. With respect to the remaining four questions, there was no significant difference between the two groups. On none of the eight questions was the performance by Group G significantly better than the performance of Group L.<sup>12</sup>

## 5. Discussion

When assessed for knowledge and understanding of the sticky price theory, students who received the feedback loop instructional method scored significantly higher than those who received graphical instruction. Analysis of individual questions suggests where the feedback loop method was more effective. On the two questions (Q1-Q2) at the low end of the cognitive skill requirement range, there was no significant performance difference between the two instructional methods. However, on four of the six questions requiring somewhat higher order thinking skills (Q3-Q8), the feedback loop method was significantly more effective.

Perhaps most noteworthy were the answers on Q8, where students using the feedback loops were more likely to retain a real-world perspective on business decision-making. Students using the AS/AD graphs were more likely to conclude that business managers are guided by abstract criteria (e.g., the search for equilibrium or long-run trends). This misperception is significant because it illustrates how—even when students seem to “get it”—the AS/AD graph can mislead. When students are asked to focus on the choreographed movement of lines on a graph, the learning challenge can become “seeing the dance” rather than thinking about real people making real decisions in a real economy. Students may think they know what is happening in the economy when they have learned to read the graph. In Colander’s (1991, p. 106) terms, they “understand incorrectly.”

There is little doubt that substantial improvement could be made in the design and administration of the experiment. The most urgent need in this area of research, however, is a definitive set of measures of the dependent variable: an individual’s *sense* of macroeconomic dynamics. One approach would require development of two short question sets, the answers to which would enable dispersion of respondents on ordinal scales—one for behavior and another for structure. An economic behavior scale should indicate respondents’ ability to predict economic behavior over time, given shocks to a range of generic structures (behavioral assumptions and institutional arrangements). An economic structure scale should rank respondents according to their ability to infer structural relationships, given a range of behaviors. Until a consensus develops around such measures, questions such as those used in this experiment can be considered *ad hoc* and the experimental results given less weight than prior beliefs, regardless of the levels of statistical significance associated with the findings. At a minimum, therefore, it is hoped that this experiment reinforces the commitment to develop dynamic thinking skills among economics students and encourages creative efforts to develop a yardstick for measuring attainment of such skills.

---

<sup>12</sup> The t values were 0.44, 0.16, 3.55, 2.27, 2.91, .092, 0.42, and 2.39 for Q1-Q8, respectively (df = 115 for each question).

## References

- Blinder, A. S. (1997). On Sticky Prices: Academic Theories Meet the Real World. In N. G. Mankiw (Ed.), *Monetary Policy*. (pp. 117-154). Chicago: The University of Chicago Press.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*. New York: David McKay Co. Inc.
- Cohn, E. S., Cohn, S., Balch, D., & Bradley, J. (2001). Do Graphs Promote Learning in Principles of Economics? *Journal of Economic Education*, Fall, 299-310.
- Colander, D. C. (1991). *Why Aren't Economists as Important as Garbagemen? Essays on the State of Economics*. Armonk, NY: M.E. Sharpe, Inc.
- \_\_\_\_\_ (1995). The Stories We Tell: A Reconsideration of AS/AD Analysis. *The Journal of Economic Perspectives*, 9(3), 169-188.
- DeLong, J. B. (2002). *Macroeconomics* (International ed.). New York: McGraw-Hill Irwin.
- Fisher, D. M. (2001). *Lessons in Mathematics: A Dynamic Approach with Applications Across the Sciences*. Lebanon, NH: isee systems, inc.
- \_\_\_\_\_ (2005). *Modeling Dynamic Systems: Lessons for a First Course*. Lebanon, NH: isee systems, inc.
- Ford, A. (1999). *Modeling the Environment* (paperback ed.). Washington, D.C.: Island Press.
- Forrester, J. W. (1968). Market Growth as Influenced by Capital Investment. *Industrial Management Review* (now *Sloan Management Review*), 9(2), 83-105.
- \_\_\_\_\_ (1976). Business Structure, Economic Cycles, and National Policy. *Futures*, June.
- \_\_\_\_\_ (1979). An Alternative Approach to Economic Policy: Macrobehavior from Microstructure. In Kamrany & Day (Eds.), *Economic Issues of the Eighties*. Baltimore: The Johns Hopkins University Press.
- Forrester, J. W., Mass, N. J., & Ryan, C. (1980). The System Dynamics National Model: Understanding Socio-economic Behavior and Policy Alternatives. *Technology Forecasting and Social Change*, 9, 51-68.
- Forrester, N. B. (1982). *A Dynamic Synthesis of Basic Macroeconomic Theory: Implications for Stabilization Policy Analysis*. Unpublished PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Hall, R. E. & Taylor, J. B. (1997). *Macroeconomics* (5th ed.). New York, NY: W.W. Norton & Company.
- Kennedy, P. E. (2000). *Macroeconomic Essentials: Understanding Economics in the News*. Cambridge MA: The MIT Press.
- Low, G. (1980). The Multiplier-Accelerator Model of Business Cycles Interpreted from a System Dynamics Perspective. In J. Randers (Ed.), *Elements of the System Dynamics Method*. Cambridge, MA: MIT Press.
- Mankiw, N. G. (2002, 2004). *Macroeconomics* (5th, 7th ed.). New York: Worth Publishers.

- Mass, N. J. (1975). *Economic Cycles: An Analysis of Underlying Causes*. Cambridge, MA: Wright-Allen Press, Inc.
- \_\_\_\_\_. (1980). Stock and Flow Variables and the Dynamics of Supply and Demand. In J. Randers (Ed.), *Elements of the System Dynamics Method*. (pp. 95-112). Cambridge, MA: MIT Press.
- Moxnes, E. (1998). Overexploitation of Renewable Resources: The Role of Misperceptions. *Journal of Economic Behavior and Organization*, 37(1), 107-127.
- Nowaczyk, R. H., Patton, C., & Santos, L. T. (1998). Student Perception of Multimedia in the Undergraduate Classroom. *International Journal of Instructional Media*, 25.
- Radzicki, M. (1993). A System Dynamics Approach to Macroeconomics (Guest lecture at the Department of Information Science, University of Bergen.).
- Richardson, G. P. (1991). *Feedback Thought in Social Science and Systems Theory*. Waltham, MA: Pegasus Communications, Inc.
- Richardson, G. P. & Pugh, A. L. (1989). *Introduction to System Dynamics Modeling*. Waltham, MA: Pegasus Communications.
- Sankaran, S. R., Sankaran, D., & Bui, T. X. (2000). Effect of Student Attitude to Course Format on Learning Performance: An Empirical Study in Web v. Lecture Instruction. *Journal of Instructional Psychology*, 27(1), 66+.
- Sexton, R. L. (2002). *Exploring Macroeconomics* (2nd ed.). Stamford, CT: Thomson Learning.
- Shaw, G. (2000). *Keeping Mozart in Mind*. San Diego, CA: Academic Press.
- Sterman, J. D. (1985). A Behavioral Model of the Economic Long Wave. *Journal of Economic Behavior and Organization*, 6, 17-53.
- \_\_\_\_\_. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston, MA: McGraw-Hill Companies.
- Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of Mathematics Self-Efficacy and Motivation in Mathematics Performance across Ethnicity. *The Journal of Educational Research*, 97(4), 208+.
- Terry, M. (2001). Translating Learning Style Theory into University Teaching Practices: An Article Based on Kolb's Experiential Learning Model. *Journal of College Reading and Learning*, 32(1), 68+.
- Wheat, I. D., Jr. (2007a). Feedback Loops in the Macro Instructor's Toolkit. *The Feedback Method: A System Dynamics Approach to Teaching Macroeconomics*. Unpublished PhD dissertation, University of Bergen, Bergen, Norway. Presented to the research session of the National Economics Education Association at the Allied Social Sciences Association Conference. Chicago, January 2007.
- \_\_\_\_\_. (2007b). Student Preferences when Explaining Dynamics. *The Feedback Method: A System Dynamics Approach to Teaching Macroeconomics*. Unpublished PhD dissertation, University of Bergen, Bergen, Norway.

\_\_\_\_\_ (2007c). Student Preferences when *Learning Dynamics*. *The Feedback Method: A System Dynamics Approach to Teaching Macroeconomics*. Unpublished PhD dissertation, University of Bergen, Bergen, Norway.

\_\_\_\_\_ (2007d). Do Feedback Diagrams Promote Learning in Macroeconomics? *The Feedback Method: A System Dynamics Approach to Teaching Macroeconomics*. Unpublished PhD dissertation, University of Bergen, Bergen, Norway.