6

Stocks and Flows

I'm very good at integral and differential calculus,
I know the scientific names of beings animalculous;
In short, in matters vegetable, animal, and mineral,
I am the very model of a modern Major-General.
—W. S. Gilbert, *The Pirates of Penzance,* Act 1.

This chapter introduces the concept of stocks and flows, a central idea in dynamics. It presents the conceptual and mathematical definitions of stocks and flows, the diagramming tools for mapping networks of stocks and flows, and case studies of the use of stocks and flows in modeling projects including automobile recycling and the construction of pulp and paper mills. Developing facility in identifying, mapping, and interpreting the stock and flow networks of systems is a critical skill for any modern systems modeler.

6.1 STOCKS, FLOWS, AND ACCUMULATION

Causal loop diagrams are wonderfully useful in many situations. They are well suited to represent interdependencies and feedback processes. They are used effectively at the start of a modeling project to capture mental models—both those of a client group and your own. They are also used to communicate the results of a completed modeling effort.

However, causal loop diagrams suffer from a number of limitations and can easily be abused. Some of these are discussed in chapter 5. One of the most important limitations of causal diagrams is their inability to capture the stock and flow structure of systems. Stocks and flows, along with feedback, are the two central concepts of dynamic systems theory.

Stocks are accumulations. They characterize the state of the system and generate the information upon which decisions and actions are based. Stocks give systems inertia and provide them with memory. Stocks create delays by accumulating the difference between the inflow to a process and its outflow. By decoupling rates of flow, stocks are the source of disequilibrium dynamics in systems.

Stocks and flows are familiar to all of us. The inventory of a manufacturing firm is the stock of product in its warehouses. The number of people employed by a business is a stock. The balance in your checking account is a stock. Stocks are altered by inflows and outflows. **A** firm's inventory is increased by the flow of production and decreased by the flow of shipments (and possibly other outflows due to spoilage or shrinkage). The workforce increases via the hiring rate and decreases via the rate of quits, layoffs, and retirements. Your bank balance increases with deposits and decreases as you spend. Yet despite everyday experience of stocks and flows, all too often people fail to distinguish clearly between them. Is the US federal deficit a stock or a flow? Many people, including politicians responsible for fiscal policy, are unclear. Failure to understand the difference between stocks and flows often leads to underestimation of time delays, a short-term focus, and policy resistance.

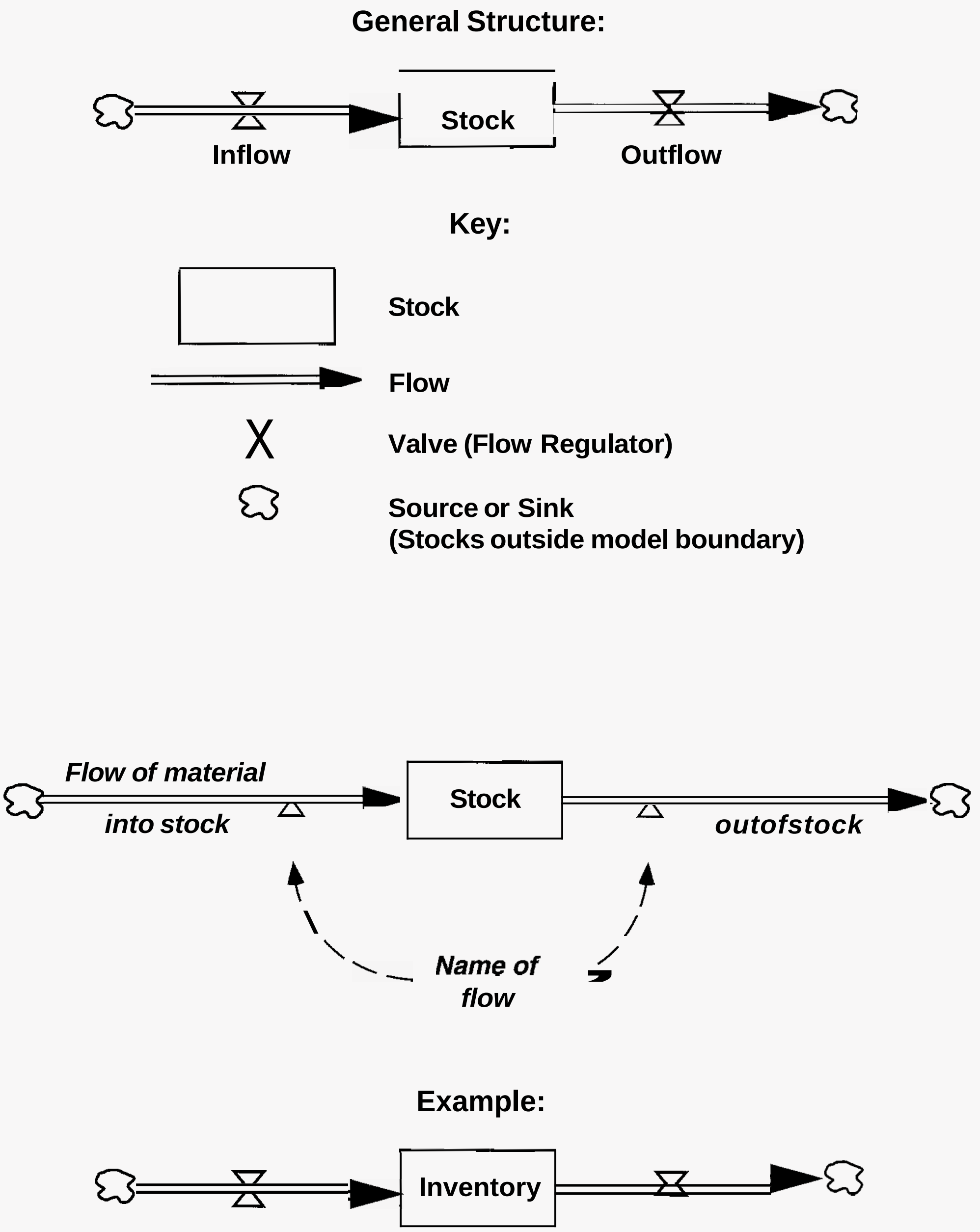
6.1.1 Diagramming Notation for Stocks and Flows

System dynamics uses a particular diagramming notation for stocks and flows (Figure 6-1).

- Stocks are represented by rectangles (suggesting a container holding the contents of the stock).
- Inflows are represented by a pipe (arrow) pointing into (adding to) the stock.
  Outflows are represented by pipes pointing out of (subtracting from) the stock.
- Valves control the flows.
- Clouds represent the sources and sinks for the flows. A source represents the stock from which a flow originating outside the boundary of the model arises; sinks represent the stocks into which flows leaving the model boundary drain. Sources and sinks are assumed to have infinite capacity and can never constrain the flows they support.

The structure of all stock and flow structures is composed of these elements. As the example in the figure shows, a firm's inventory is a stock that accumulates the inflow of production and is reduced by the outflow of shipments. These are the only flows considered in the model: unless explicitly shown, other possible flows into or out of the stock, such as inventory shrinkage or spoilage, are assumed to be zero. The clouds indicate that the stock of raw materials never starves the production rate and the stock of product shipped to customers never grows so high that it blocks the shipment rate.

FIGURE 6-1 Stock and flow diagramming notation



6.1.2 Mathematical Representation of Stocks and Flows

The stock and flow diagramming conventions (originated by Forrester 1961) were based on a hydraulic metaphor—the flow of water into and out of reservoirs. Indeed, it is helpful to think of stocks as bathtubs of water. The quantity of water in your bathtub at any time is the accumulation of the water flowing in through the tap less the water flowing out through the drain (assume no splashing or evaporation). In exactly the same way, the quantity of material in any stock is the accumulation of the flows of material in less the flows of material out. Despite the prosaic metaphor the stock and flow diagram has a precise and unambiguous mathematical meaning. Stocks accumulate or *integrate* their flows; the net flow into the stock is the rate of change of the stock. Hence the structure represented in Figure 6-1 above corresponds exactly to the following integral equation:

$$\text{Stock}(t) = \int_{t_0}^{t} [\text{Inflow}(s) - \text{Outflow}(s)]ds + \text{Stock}(t_0) \tag{6-1}$$

where Inflow(s) represents the value of the inflow at any time s between the initial time t_0 and the current time t. Equivalently, the net rate of change of any stock, its derivative, is the inflow less the outflow, defining the differential equation

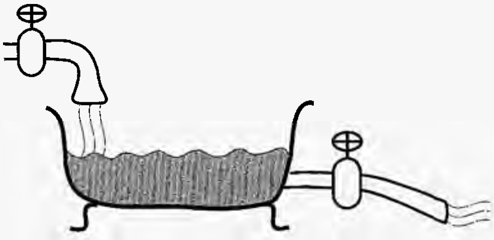
$$d(\text{Stock})/dt = \text{Inflow}(t) - \text{Outflow}(t). \tag{6-2}$$

In general, the flows will be functions of the stock and other state variables and parameters. Figure 6-2 shows four equivalent representations of the general stock and flow structure. The bathtub and stock and flow diagrams may appear to be less rigorous than the integral or differential equation representations, but they are precisely equivalent and contain exactly the same information. From any system of integral or differential equations we can construct the corresponding stock and flow map; from any stock and flow map we can generate the corresponding integral or differential equation system.

FIGURE 6-2 Four equivalent representations of stock and flow structure

Each representation contains precisely the same information.

Hydraulic Metaphor:



Stock and Flow Diagram:

Integral Equation:

$$\text{Stock}(t) = \int_{t_0}^{t} [\text{Inflow}(s) - \text{Outflow}(s)]ds + \text{Stock}(t_0)$$

Differential Equation:

$$d(\text{Stock})/dt = \text{Net Change in Stock} = \text{Inflow}(t) - \text{Outflow}(t)$$

Process Point: Notation for Accumulation

The traditional notation used in calculus and shown in Figure 6-2 is often confusing to many people. In this book, I will generally represent the process of accumulation with the INTEGRAL() function:

$$\text{Stock} = \text{INTEGRAL}(\text{Inflow} - \text{Outflow}, \text{Stock}_{t_0}) \qquad (6\text{-}3)$$

The INTEGRAL() function is exactly equivalent to equation (6-1) and represents the concept that the stock accumulates its inflows less its outflows, beginning with an initial value of Stock_{t_0}.

6.1.3 The Contribution of Stocks to Dynamics

Stocks are critical in generating the dynamics of systems for the following reasons (Mass 1980):

Stocks:
Where you are

Flows:
Where you are
going

1. **Stocks characterize the state of the system and provide the basis for actions.**
   The stocks in a system tell decision makers where they are, providing them with the information needed to act. **A** pilot must know the state of the aircraft including position, heading, altitude, and fuel level. Without knowledge of these states, the pilot is flying blind and won't survive long. Likewise, a firm can't set its production schedule appropriately without knowledge of the order backlog, the stock of inventory, the parts stocks, the labor force, and other stocks. A balance sheet characterizes the financial health of a corporation by reporting the values of stocks such as cash, inventory, payables, and debt. Information about these stocks affects decisions such as issuing new debt, paying dividends, and controlling expenses via layoffs.

2. **Stocks provide systems with inertia and memory.**
   Stocks accumulate past events. The content of a stock can only change through an inflow or outflow. Without changes in these flows, the past accumulation into the stock persists. The stock of lead in the paint of America's inner city housing remains high today even though lead paint was banned in 1978. Once the stock of lead paint accumulated, the only way to eliminate it is through expensive deleading or the eventual demolition of the housing itself. Even then the lead remains, either safely sequestered or more likely dispersed into the environment as dust, chips, or lead leaching from landfills into water supplies. Likewise, the stock of ozone-destroying chlorine generated by CFCs will remain in the atmosphere for decades even after the production rate of CFCs falls to zero because the rate at which chlorine is scrubbed from the stratosphere is very low. Stocks don't have to be tangible. Memories and beliefs are stocks that characterize your mental states. Your beliefs persist over time, generating inertia and continuity in your attitudes and behavior. If you have a bad experience on an airline and never fly on that carrier again, your belief about the low quality of their service remains even if they've improved.

3. **Stocks are the source of delays.**
All delays involve stocks. **A** delay is a process whose output lags behind its input. The difference between the input and output accumulates in a stock of material in process. There is a lag between the time you mail a letter and the time it is received. During this interval, the letter resides in a stock of letters in transit. Even email accumulates in stocks of undelivered packets and messages residing in the memory of various computers between the sender and receiver. There is a lag of several years between the decision to build new office buildings and the time they are ready for occupancy. During this interval there is a supply line of buildings under development, including a stock of proposed projects and a stock of buildings under construction.

   By definition, when the input to a delay changes, the output lags behind and continues at the old rate for some time. During such adjustments, the stock accumulating the difference between input and output changes. If you mail wedding invitations to 1000 of your closest friends all at once, while the rate at which other mail is deposited remains constant, the stock of letters in transit jumps by 1000 and remains at the new level as the letters make their way to their destinations. Only as the invitations begin to arrive does the stock of letters in transit start to fall. The delivery rate exceeds the mailing rate, shrinking the stock of mail in transit, until all the invitations have been delivered, at which point the delivery rate once again equals the rate at which mail is deposited and the stock of letters in transit returns to its original level.

   Perception delays also involve stocks though these stocks do not involve any material flows. For example, the belief of managers in a company's Taiwan headquarters about the shipment rate from their Silicon Valley plant lags behind the true shipment rate due to measurement and reporting delays. Measurement of a rate such as shipments always involves a stock. Due to unpredictable variations in customer orders, product availability, and transportation, shipments can vary significantly from hour to hour, day to day, or over even longer periods. Shipments must be accumulated for some period of time such as a day, week, or month to provide a meaningful measurement of the rate. If shipments are highly volatile, the firm will have to accumulate them over longer intervals to filter out the short-term noise and provide a meaningful average managers can use to make decisions. In addition there are reporting delays involving a stock of shipment information waiting to be uploaded to and downloaded from the firm's computer system. There may be further delays in the adjustment of the executives' beliefs even after they see the latest data. Chapter **11** describes the structure and dynamics of delays in detail.

4. **Stocks decouple rates of flow and create disequilibrium dynamics.**
Stocks absorb the differences between inflows and outflows, thus permitting the inflows and outflows to a process to differ. In equilibrium, the total inflow to a stock equals its total outflow so the level of the stock is unchanging. However, inflows and outflows usually differ because they are often governed by different decision processes. Disequilibrium is the rule rather than the exception.

The production of grain depends on the yearly cycle of planting and harvest, along with unpredictable natural variations in weather, pest populations, and so on. Consumption of grain depends on how many mouths there are to feed. The difference between grain production and consumption rates accumulates in grain stocks, stored throughout the distribution system from field to grain elevator to processor inventories to market to kitchen cupboard. Without a stock of grain to buffer the differences between production and consumption, consumption would necessarily equal production at all times and people would starve between harvests. Thus Joseph advised Pharaoh to stockpile grain during the 7 good years in anticipation of the 7 lean years during which consumption would exceed harvests. While on average the production of grain balances consumption (and losses) as farmers respond to market prices and inventory conditions in determining how much to plant, and as consumers adjust consumption in response to prices and availability, production and consumption are rarely equal.

Whenever two coupled activities are controlled by different decision makers, involve different resources, and are subject to different random shocks, a buffer or stock between them must exist, accumulating the difference. **As** these stocks vary, information about the size of the buffer will feed back in various ways to influence the inflows and outflows. Often, but not always, these feedbacks will operate to bring the stock into balance. Whether and how equilibrium is achieved cannot be assumed but is an emergent property of the whole system as its many feedback loops interact simultaneously. Understanding the nature and stability of these dynamics is often the purpose of a system dynamics model.

6.2 Identifying Stocks and Flows

The distinction between stocks and flows is recognized in many disciplines. Table 6-1 shows some common terms used to distinguish between stocks and flows in various fields. In mathematics, system dynamics, control theory, and related engineering disciplines, stocks are also known as *integrals* or *state variables.* Flows are also known as *rates* or *derivatives.* Chemists speak of *reactants* and *reaction products* (the stocks) and *reaction rates* (the flows). In manufacturing settings, stocks and flows are also called *buffers* and *throughput.* In economics, stocks are also known as *levels* and flows as *rates.* For example, the capital stock of an economy is its level of wealth (measured in, say, dollars) while the GDP is the aggregate rate of national output (measured in $/year). In accounting, balance sheet items are stocks, such as cash, the book value of inventory, long-term debt, and shareholder equity (all measured in, e.g., dollars). Items appearing on the income statement or flow of funds report are flows which alter the corresponding stocks on the balance sheet, such as net receipts, the cost of goods sold, long-term borrowing, and the change in retained earnings. These flows are measured in $/year. Physiological models often lump different stocks into a small number of *compartments* or boxes connected by diffusion rates (the flows). For example, the stock of glucose in a human can be represented by a three compartment model:

TABLE 6-1 Terminology used to distinguish between stocks and flows in different disciplines

Field	Stocks	Flows
Mathematics, physics and engineering	Integrals, states, state variables, stocks	Derivatives, rates of change, flows
Chemistry	Reactants and reaction products	Reaction rates
Manufacturing	Buffers, inventories	Throughput
Economics	Levels	Rates
Accounting	Stocks, balance sheet items	Flows, cash *flow* or income statement items
Biology, physiology	Compartments	Diffusion rates, flows
Medicine, epidemiology	Prevalence, reservoirs	Incidence, infection, morbidity and mortality rates

glucose in the digestive system, glucose in the bloodstream, and glucose in the intracellular matrix. In epidemiology, *prevalence* measures the number or stock of people who have a particular condition at any given time, while *incidence* is the rate at which people come down with the disease or condition. In December 1998 the prevalence of HIV/AIDS worldwide was estimated by the United Nations AIDS program to be 33.4 million and the incidence of HIV infection was estimated to be 5.8 million/year. That is, a total of 33.4 million people were estimated to be HIV-positive or to have AIDS; the rate of addition to this stock was 5.8 million people per year (16,000 new infections per day). The net change in the population of HIV-positive individuals was estimated to be 3.3 million people per year due to the death rate from AIDS, estimated to be 2.5 million people per year in 1998.

How can you tell which concepts are stocks and which are flows? Stocks are quantities of material or other accumulations. They are the states of the system. The flows are the rates at which these system states change. Imagine a river flowing into a reservoir. The quantity of water in the reservoir is a stock (measured in, say, cubic meters). If you drew an imaginary line across the point where the river enters the reservoir, the flow is the rate at which water passes the line—the rate of flow in cubic meters per second.

6.2.1 Units of Measure in Stock and Flow Networks

The units of measure can help you distinguish stocks from flows. Stocks are usually a quantity such as widgets of inventory, people employed, or Yen in an account. The associated flows must be measured in the same units ***per time period,*** for example, the rate at which widgets are added per week to inventory, the hiring rate in people per month, or the rate of expenditure from an account in ¥/hour. Note that the choice of time period is arbitrary. You are free to select any measurement system you like as long as you remain consistent. You can measure the flow of production into inventory as widgets per week, widgets per day, or widgets per hour. The statement "The current rate of production is 1200 widgets per day" is exactly equivalent to the statement that production is proceeding at a rate of 8400 widgets per week, *50* widgets per hour, *5/6* widgets per minute, or even 43,800,000 widgets per century. All are statements about how many widgets are being produced *right now—at this instant*. Whether the cumulative number of widgets produced in any given interval such as a day, week, or century is equal to 1200,8400, or 43,800,000 depends on whether the current rate stays constant over that interval (or averages out to the current rate). Most likely it won't.

6.2.2 The Snapshot Test

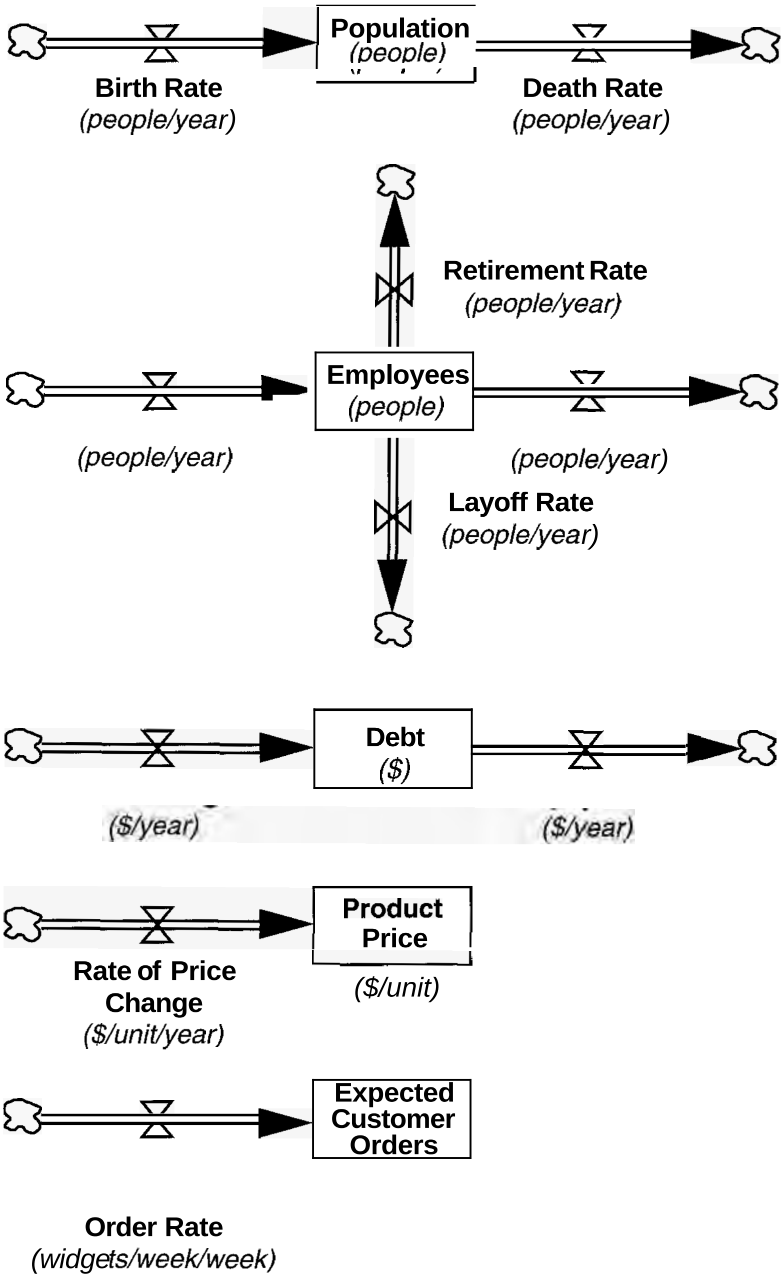
Stocks characterize the state of the system. To identify key stocks in a system, imagine freezing the scene with a snapshot. Stocks would be those things you could count or measure in the picture, including psychological states and other intangible variables. You can estimate the stock of water in a reservoir from a set of satellite images and topographic data, but you cannot determine whether the water level is rising or falling. Your bank statement tells you how much money is in your account but not the rate at which you are spending it now. If time stopped, it would be possible to determine how much inventory a company has or the price of materials but not the net rate of change in inventory or the rate of inflation in materials prices. The snapshot test applies also to less tangible stocks. The plant manager's expectation of the customer order rate at any instant or perception of the size of inventory are stocks, even though they are mental and not physical stocks. A snapshot of people's mental states, however, does not indicate how fast they are revising their beliefs.

Figure *6-3* lists some common concepts and identifies them as stocks or flows, showing their stock and flow structure and units of measure. Population, Employees, and Debt are straightforward. Why is the price of a product a stock? Prices characterize the state of the system, in this case how much you must pay per unit. A price posted on an item remains in effect until it is changed, just as the number of widgets in an inventory remains constant until it is changed by a flow of production or shipments. Even the bids and offers called out in a trading pit at a financial market are stocks, albeit short-lived ones: a bid or offer remains in effect until the trader withdraws or alters it by crying out another.

Why is the expected customer order rate for a product a stock? Clearly, the actual customer order rate is a flow. The flow of customer orders accumulates in a backlog or stock of unfilled orders until the product can be delivered. However, a manager's belief about the rate at which customer orders are booked is a stock—it is a state of the system, in this case a mental state. No one knows the true current or future order rate. The manager's belief about orders can, and usually does, differ from the true order rate (the belief can be wrong). Managers' beliefs about the customer order rate will tend to remain the same until they become aware of new information and update their beliefs. The Change in Expected Order Rate is the rate at which the belief is updated. Note the units of measure for the expected order rate. Like the actual order rate, the expected order rate is measured in widgets per time period (say weeks). The units of measure for the rate at which the belief about customer orders is updated are (widgets/week)/week.

FIGURE 6-3
Examples of stocks and flows with their units of measure

The choice of time unit for the flows (e.g., days, weeks, years) is arbitrary but must be consistent within a single model.



Note that the rate of price change and the change in the expected order rate can be positive or negative (prices and demand forecasts can rise or fall). Any flow into or out of a stock can be either positive or negative. The direction of the arrow (pointing into or out of a stock) defines the sign convention for the flow. An inflow adds to the stock when the flow is positive; if the flow is negative it subtracts from the stock. When the outflow is positive, the flow subtracts from the stock.

CHALLENGE

Identifying Stocks and Flows

Are the following concepts stocks or flows? Draw a stock and flow map for each and give their units of measure.

1. Interest rate (e.g., the prime interest rate or rate on the 30-year US Treasury bond).
2. Unemployment rate.

Hint: What does the word "rate" mean in these settings?

6.2.3 Conservation of Material in Stock and Flow Networks

A major strength of the stock and flow representation is the clear distinction between the physical flows through the stock and flow network and the information feedbacks that couple the stocks to the flows and close the loops in the system. The contents of the stock and flow networks are *conserved* in the sense that items entering a stock remain there until they flow out. When an item flows from one stock to another the first stock loses precisely as much as the second gains. Consider the stock and flow structure representing the accounts receivable of a firm (Figure 6-4). The stock of receivables is increased by billings and decreased by payments received and by defaults. The flow of billings is conserved in the sense that once a customer is billed, the invoice remains in the stock of receivables until it explicitly flows out when the receivables department records the customer's payment or acknowledges that the customer has defaulted and writes off the account. In contrast, information about the stock of receivables is not conserved. The corporate accounting system makes the value of the receivables stock available to many throughout the organization. Accessing and using this information does not use it up or make it unavailable to others.

Note also that while the units of accounts payable are dollars and the billing, payment, and default flows are measured in dollars per time period, the contents of the stock are not actually dollars. Rather, the content of the receivables stock is information, specifically, a ledger or database consisting of records of invoices outstanding. To see why, imagine trying to exchange your firm's stock of receivables for cash—you can sell them to a collection agency, but only for much less than 100 cents on the dollar. Though the contents of the stock of receivables is information and not a material quantity, it is nevertheless conserved—you cannot sell a given stock of receivables more than once (not legally, anyway). Stocks can represent information as well as more tangible quantities such as people, money, and materials. Stocks can also represent intangible variables including psychological states, perceptions, and expectations such as employee morale, the expected rate of inflation, or perceived inventory.

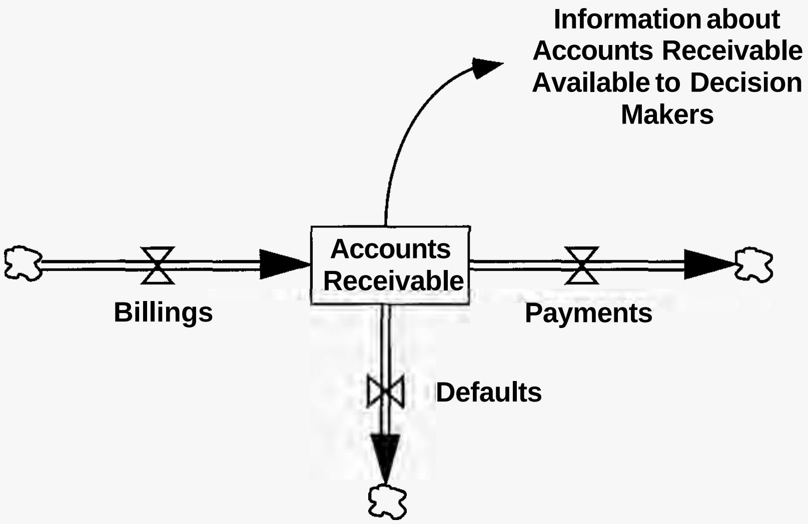


FIGURE 6-4
Stock and flow structure of accounts receivable

The material flowing through the network is actually information about customers and the amounts they owe. This information is conserved—the only way a receivable, once billed, is removed from the stock is if the customer pays or defaults. Information about the size and composition *of* accounts payable, however, can be made available throughout the system and is not depleted by usage.

6.2.4 State-Determined Systems

The theory of dynamic systems takes a state-determined system or state variable approach. The only way a stock can change is via its inflows and outflows. In turn, the stocks determine the flows (Figure 6-5).

Systems therefore consist of networks of stocks and flows linked by information feedbacks from the stocks to the rates (Figure 6-6). As shown in the figure, the determinants of rates include any constants and exogenous variables. These too are stocks. Constants are state variables that change so slowly they are considered to be constant over the time horizon of interest in the model. Exogenous variables are stocks you have chosen not to model explicitly and are therefore outside the model boundary. For example, in a model of the demand for a new video game, the size of the potential market might depend on the population between, say, ages 4 and 20. The product life cycle will last a few years at most. Over this time horizon the population between 4 and 20 years of age is not likely to change significantly and can reasonably be assumed constant. Alternatively, you could model the stock of children in the target age group as an exogenous variable, using census data and projections to estimate its values. Making population constant or exogenous is acceptable in this case since there are no significant feedbacks between sales of video games and birth, death, or migration rates.

6.2.5 Auxiliary Variables

As illustrated in Figure 6-6, mathematical description of a system requires only the stocks and their rates of change. For ease of communication and clarity, however, it is often helpful to define intermediate or *auxiliary variables.* Auxiliaries consist of functions of stocks (and constants or exogenous inputs). For example, a population model might represent the net birth rate as depending on population and the fractional birth rate; fractional birth rate in turn can be modeled as a function of food per capita. The left side of Figure 6-7 shows the structure and equations for

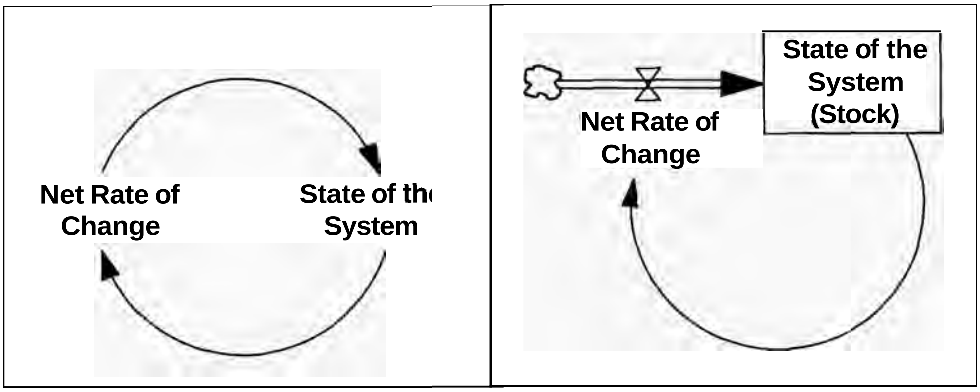


FIGURE 6-5
State-determined systems

Systems evolve by feedback of information from the state *of* the system to the flows that alter the states.

Left: Causal loop representation in which the stock and flow structure is not explicit.

Right: Explicit stock and flow structure for the same feedback loop.

The equations correspond to the stock and flow map. The net rate of change of the stock is a function of the stock itself, closing the feedback loop.

State of the System = INTEGRAL(Net Rate of Change, State of the System$_{t_0}$)

Net Rate of Change = *f* (State of the System)

the model. The Net Birth Rate accumulates in the Population stock. The auxiliary variables Fractional Birth Rate and Food per Capita are neither stocks nor flows. They are functions of the stocks (and exogenous inputs, in this case Food). Population participates in two feedback loops: a positive loop (more people, more births, more people) and a negative loop (more people, less food per person, lower fractional net birth rate, fewer births). The inclusion of the auxiliary variables distinguishes the two loops and allows unambiguous assignment of link and loop polarities.

The auxiliaries can always be eliminated and the model reduced to a set of equations consisting only of stocks and their flows. By substituting the equation for Food per Capita into the equation for Fractional Birth Rate and then substituting the result into the equation for Net Birth Rate, you can eliminate the auxiliaries, reducing the model to one with only Net Birth Rate and Population. The right side of Figure 6-7 shows this model and its equations. Though the model is mathematically equivalent to the model with auxiliaries, it is harder to explain, understand, and modify. Note that in the reduced form model population enters the equation for the rate of change of population in both the numerator and denominator. The polarity of the causal link between Population and Net Births is now ambiguous, and it is not possible to distinguish the two feedback loops involving population and births.

The process of creating the reduced form model by substitution of intermediate variables into their rates is a general one and can be carried out on any model. However, the use of auxiliary variables is critical to effective modeling. Ideally, each equation in your models should represent one main idea. Don't try to economize on the number of equations by writing long ones that embed multiple concepts. These long equations will be hard for others to read and understand. They will be hard for *you* to understand. Finally, equations with multiple components and ideas are hard to change if your client disagrees with one of the ideas.

FIGURE 6-6 Networks of stocks and flows are coupled by information feedback.

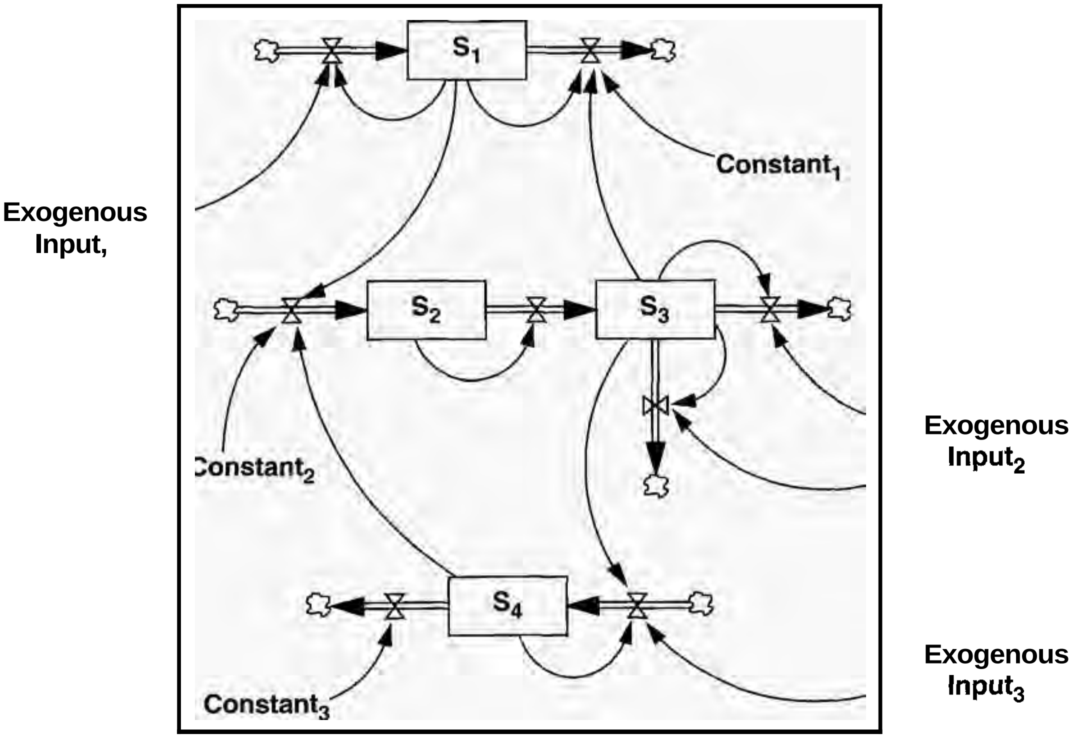
Stocks accumulate their rates of flow; information about the stocks feeds back to alter the rates, closing the loops in the system. Constants are stocks changing too slowly to be modeled explicitly; exogenous variables are stocks outside the model boundary (shown by the rectangle with rounded corners).

Equation representation: The derivatives of the stocks in dynamic systems are, in general, nonlinear functions of the stocks, the exogenous variables, and any constants. In matrix notation, the rates of change $d\mathbf{S}/dt$ are a function $\boldsymbol{f}()$ of the state vector $\mathbf{S}$, the exogenous variables $\mathbf{U}$ and the constants $\mathbf{C}$:

$$d\mathbf{S}/dt = \boldsymbol{f}(\mathbf{S}, \mathbf{U}, \mathbf{C}) \tag{6-4}$$

For the diagram below, the equation for the rate of change of S_4 is

$$dS_4/dt = f_4(S_3, S_4, U_3, C_3) \tag{6-5}$$



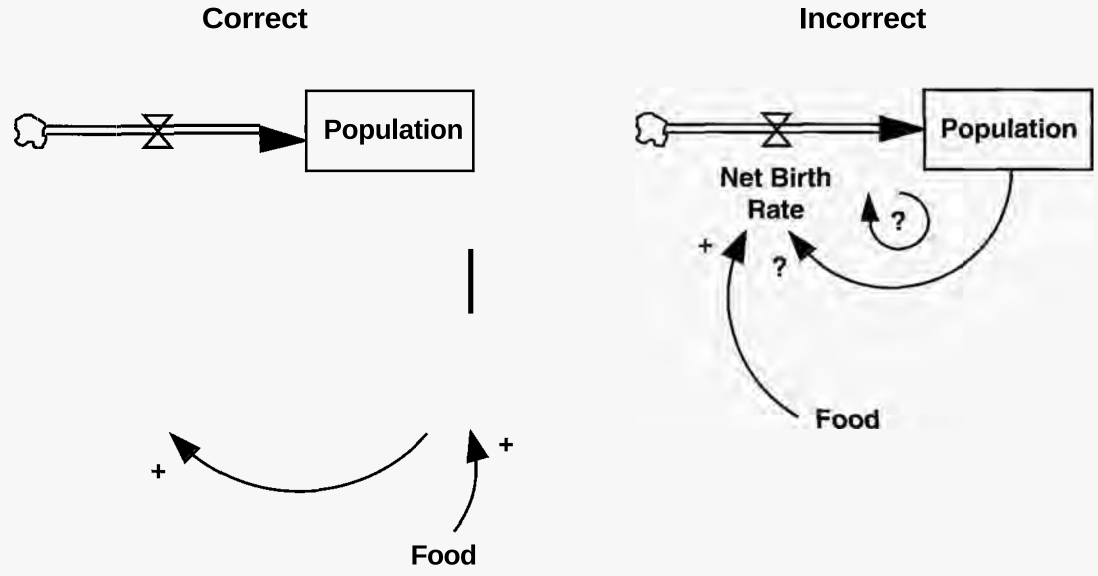
6.2.6 Stocks Change Only Through Their Rates

Stocks change only through their rates of flow. There can be no causal link directly into a stock. Consider a model for customer service. Customers arrive at some rate and accumulate in a queue of Customers Awaiting Service. The queue could be a line at a fast food restaurant, cars awaiting repair at a body shop, or people on hold calling for airline reservations. When the service is completed customers depart from the queue, decreasing the stock of customers waiting for service. The rate at which customers can be processed depends on the number of service personnel, their productivity (in customers processed per hour per person), and the number of hours they work (the workweek). If the number of people waiting for service increases, employees increase their workweek as they stay an extra shift, skip lunch, or cut down on breaks.

FIGURE 6-7 Auxiliary variables

Left: A simple population model with auxiliary variables. Fractional Birth Rate and Food per Capita are neither stocks nor flows, but intermediate concepts added to the model to aid clarity.

Right: The same model with the auxiliary variables eliminated by substitution into the rate equation. The link from Population to Net Birth Rate now has an ambiguous sign, a poor practice.



Population = INTEGRAL(Net Birth Rate, Population_{t_0})

Net Birth Rate = Population * Fractional Birth Rate

Fractional Birth Rate = *f*(Food per Capita)

Food per Capita = Food/Population

Population = INTEGRAL(Net Birth Rate, Population_{t_0})

Net Birth Rate = Population * *f*(Food/Population)

I have often seen people in workshops draw the diagram shown in the top of Figure 6-8. They correctly recognize that the rate at which customers are processed is the product of Service Staff, Productivity, and Workweek and that higher queues of waiting customers lead to longer hours and hiring of additional staff, forming two balancing feedbacks. But often people draw information feedbacks directly from the workweek and service staff to the stock of Customers Awaiting Service, assigning them a negative polarity. They reason that an increase in the workweek or staff level decreases the number of customers remaining in the queue, thus closing the negative feedback loops.

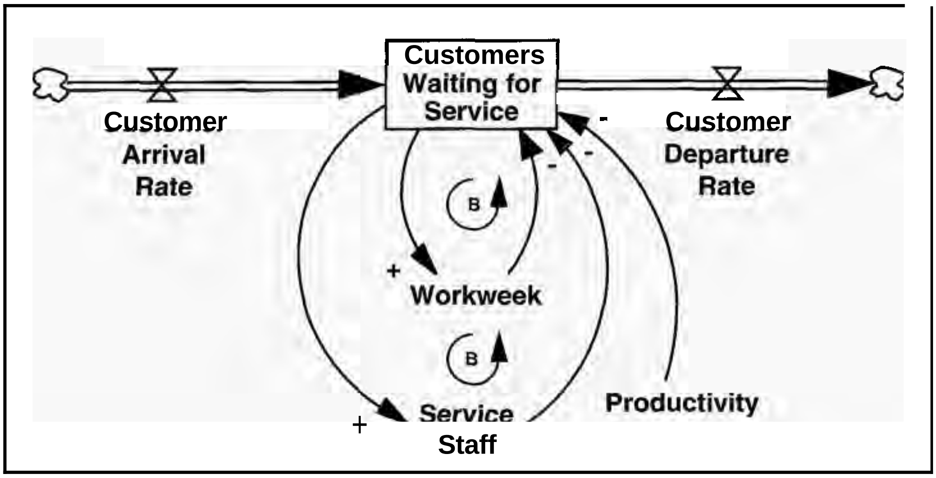
The correct diagram is shown in the lower panel of Figure 6-8. The only way customers can exit the stock is via the departure rate. The departure rate is the product of the number of staff, their workweek, and their productivity. An increase in any of these inputs boosts the rate at which customers are processed and leave the queue. The balancing feedbacks are still present: A longer queue of waiting customers leads to longer hours and more staff and an increase in the processing rate. The valve controlling the outflow from the stock of waiting customers opens wider, and customers depart the queue at a higher rate. The polarities of the information links in the feedback loop are all positive, but an increase in the customer departure rate causes a reduction in the stock of waiting customers because the departure rate is an outflow from the stock.

FIGURE 6-8 Stocks change only through their rates.

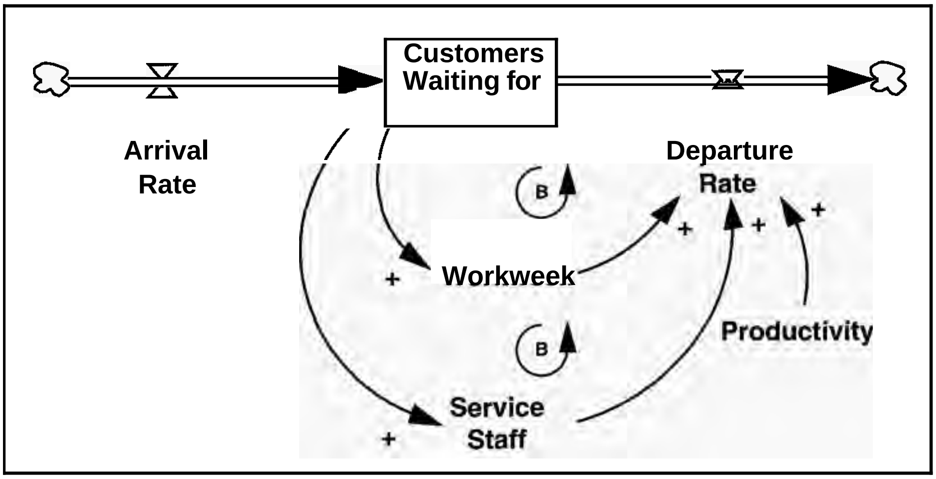
Top: Incorrect stock and flow map of a service operation. Workweek, Service Staff, and other variables cannot directly alter the stock of Customers Awaiting Service.

Bottom: Corrected diagram. The Workweek, number of Service Staff, and Productivity drive the Customer Departure Rate, which decreases the stock of Customers Awaiting Service.

Incorrect



Correct



6.2.7 Continuous Time and Instantaneous Flows

The stock and flow perspective (and its equivalent integral or differential equation structure) represents time as unfolding continuously. That is, as our experience suggests, time progresses smoothly and continuously. In system dynamics we almost always represent time as continuous. Events can happen at any time; change can occur continuously; and time can be divided into intervals as fine as one desires.]

---

[1] In numerical simulation time is divided into discrete intervals. However, these intervals must be small enough that the numerical solution is a good approximation of the underlying continuous dynamics, and model dynamics cannot depend on the length of the solution interval (cutting it in half, e.g., should not affect any of your conclusions). In discrete time or difference equation systems the time interval is an irreducible minimum time delay in every feedback loop and often has a large impact on the dynamics. Appendix A discusses numerical integration and the selection of an appropriate time step for your simulations.

A flow at any time is defined to be its instantaneous value—the rate at which water is flowing into your bathtub *right* now. Mathematically, the net flow to a stock (inflows less outflows) is the instantaneous rate of change of the stock—its derivative (this is the meaning of equation 6-2). No one can measure the instantaneous value of any flow. The government does not and cannot report the GDP at a particular moment but instead reports the average rate of production over some prior, finite interval of time (typically a quarter of a year). Likewise, quarterly reports of a firm's sales are the cumulative sales during the quarter, not the instantaneous sales rate at the end of the quarter. During the quarter sales likely varied substantially. Sales reports at more frequent intervals such as monthly or even daily are better approximations of the instantaneous sales rate but still represent averages taken over some prior, finite interval. Similarly, the speedometer of a car does not measure its instantaneous velocity. Because the components of the velocity sensor and instrumentation have inertia, the speedometer indicates an average of the velocity over a (short) prior interval.

As the length of the measurement interval shrinks, the reported average rate becomes a better approximation of the instantaneous rate. Most speedometers respond quickly relative to the rate of change of the car's true velocity, so for practical purposes the average velocity reported on the instrument panel is the same as the actual, current velocity. On the other hand, the delays in reporting the state of the economy or the profits of a company are often long relative to their rates of change and dramatically influence the stability of the system. Though we might develop instruments for our physical and social systems that shrink the delays in measuring and reporting rates of flow, we can never measure the instantaneous value of the flows affecting any stock.

6.2.8 Continuously Divisible versus Quantized Flows

Just as time can be represented as unfolding continuously or in discrete intervals, so too the units flowing into and out of stocks can be thought of either as continuously divisible or as a discrete numbers of items. Most flows are actually quantized, meaning they consist of collections of individual items which cannot be divided into arbitrarily small units. Oil tankers are commissioned one at a time—it is not meaningful to speak of launching half a tanker. Company hiring consists of a whole number of people. Even the flow in a river consists of an (astronomically large) integer number of water molecules. The stock and flow concept and the four equivalent notations shown in Figure 6-2 apply whether the flow is conceived to be infinitely divisible or quantized. The metaphor of the flow of water into a bathtub emphasizes our everyday experience of water as a continuously divisible substance—we aren't concerned with the identity of the individual water molecules. However, if it were important to our purpose, we could just as easily imagine that the tub is filled by a lumpy flow of individual ice cubes. Whether continuous or quantized, the quantity in the stock is always the accumulation of the inflows to the stock less its outflows.

In many models it is appropriate and useful to approximate the flow of individual items as a continuous stream. In modeling the cash flow of a large organization you usually do not need to track individual payments; it is a perfectly acceptable approximation to consider the flow of revenue and expenditures as continuous in time and continuously divisible (though of course the accounting department must track individual payments). Similarly, though organizations hire discrete, whole individuals, it is usually acceptable to assume the flows of people are continuously divisible. Some clients are troubled by the fractional people your model will generate, but almost always the error is insignificant compared to the measurement error in parameters including the number of employees the firm actually has. Since people can be hired part time, work in job-sharing situations, or be assigned to multiple projects, it is quite meaningful to speak of fractional employees, measured in FTE (Full-Time Equivalent) people.

When the purpose of the model requires tracking the individual people, for example modeling the behavior of people entering the line at the supermarket to determine the optimal number of checkout counters, then people can be modeled as discrete individuals arriving at discrete points; this is a classic modeling paradigm in queuing theory (Prabhu 1997; Gross and Harris 1998; Papadopoulos 1993). Yet even in many queuing applications, the continuous time, continuous flow approximation works extremely well and the errors it introduces are often small compared to measurement error and parameter uncertainty in the real system. The decision to represent stocks and flows as continuous or discrete always depends on the purpose of the model. For example, if your purpose is to understand the dynamics of price and the origin of cycles in the oil tanker market (see chapter 20), it is fine to assume that the rates of tanker ordering, construction, and scrappage are continuous in time and continuously divisible. In contrast, if your purpose were to model the arrival and offloading of tankers to optimize port facilities you would have *to* model the ships as discrete entities.

6.2.9 Which Modeling Approach Should You Use?

The choice of modeling technique and software will depend on which assumptions about the stocks and flows in your system are appropriate to your purpose. In all cases make sure your modeling software and method can include the feedback processes you consider important. In modeling the behavior of people in line at the supermarket, for example, you might choose to use a discrete time, quantized flow representation and select a stochastic modeling package, or even use a spreadsheet. Be sure, however, that your tools allow you to capture behavioral feedbacks such as the feedback from the length of the line to the rate at which people join the line. Some models and a great many of the theorems in queuing theory assume that the rate of arrivals to a queue such as the checkout line is exogenous. People actually choose to enter a line based on its length (more precisely, their estimate of expected waiting time). **A** long line will cause people to switch *to* another, defer their shopping to a less crowded time of day, or even go to a different store. Such ***balking*** creates a powerful negative feedback loop: The longer the line, the smaller the arrival rate. Omitting such feedback processes from your model in the interests of analytical tractability or programming convenience will often lead to a fatal flaw in your analysis and policy conclusions.

6.2.10 Process Point: Portraying Stocks and Flows in Practice

Each of the stock and flow representations in Figure 6-2 (the bathtub, stock and flow diagram, integral equation, and differential equation) contains precisely the same information. They are exactly equivalent. Which should you use to develop and present your models, especially when you are working in a team?

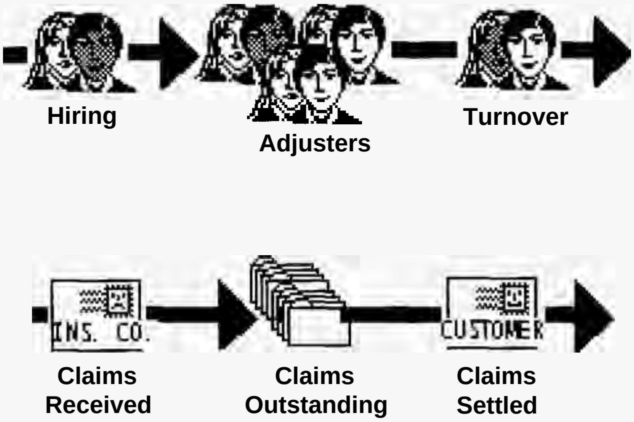
The answer depends on the context of the modeling project you are doing and the background of your client team. While many mathematically sophisticated modelers scoff at the idea of explaining a complex model using bathtubs and pipes, I have many times seen otherwise brilliant modeling efforts founder because the analyst tried to explain a model using differential equations and mathematical notation—or the simulation code—to a client team with little technical background. One of the worst things a consultant can do is humiliate the client. Showing off your mathematical knowledge by using differential equations, lots of Greek letters, and other notation the client never studied or forgot a long time ago is a sure-fire way to convince your clients you care more for the elegance of your equations than for helping them solve their problem.

Stock and flow diagrams contain the same information as the more mathematically formal notation but are easier to understand and to modify on the fly. Still, some team members consider even the stock and flow diagram format to be too abstract. I have often seen clever graphics of tanks, pipes, and valves used to excellent effect with client teams. For example, a consulting project with a multinational chemicals firm represented the flows of production, inventories, shipments, and customer stocks—along with capacity, cash, and even equipment defects—as a series of pipes, valves, and tanks. The team members were able to grasp the stock and flow structure readily since they were all familiar with the tanks and pipes carrying materials in their plants. In fact, most of the client team were engineers by training and had plenty of background in mathematics. Yet several commented that they never really understood how the business worked until they saw the chart showing its stock and flow structure as tanks and pipes.

What if your clients have even less technical training than these chemical company executives? The bathtub metaphor is often used to good effect, as illustrated by the case of automobile leasing (see Figure 2.4). What if the stocks and flows in your model aren't as tangible as barrels of oil or automobiles? Get creative. In a management flight simulator of insurance claims processing (Kim 1989; Diehl 1994), a flow of letters arriving to an inbox represented the addition of new claims to the stock of unresolved claims. Letters containing checks flowed out to the customers as claims were settled. Icons of people represented the stock and flow structure of claims adjusters (Figure 6-9). Participants in workshops using the model were able to understand the system structure much better than if the more abstract symbols had been used.

I am not recommending that you keep the equations or stock and flow diagrams hidden from your client. Never hide your model from a curious client. You should always look for and create opportunities for client team members to learn more about the modeling process; you should always be prepared to explain the workings of your model.

FIGURE 6-9 Stocks and flows of claims and claims adjusters in an insurance company



Source: Kim 1989.

And while I caution the mathematically sophisticated modeler against overly technical presentation, the opposite problem can also arise: some clients are offended by what they consider to be simplistic cartoon diagrams and prefer what they view as the more professional presentation of stock and flow diagrams or even equations. As always, you must get to know your client deeply and early in the modeling process.

Finally, a caution for those with less technical training and mathematical background: Clients may not need to understand the deep relationship between their bathtub and the mathematics underlying stocks and flows, but you do. While you don't need to be able to solve differential equations to be a successful modeler, you do need to understand the structure and dynamics of stocks and flows thoroughly and rigorously.

6.3 MAPPING STOCKS AND FLOWS

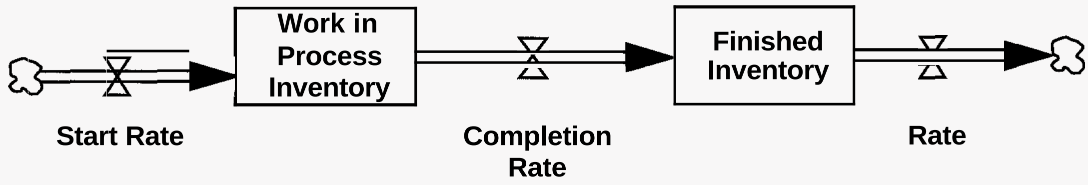
6.3.1 When Should Causal Loop Diagrams Show Stock and Flow Structure?

Causal diagrams can be drawn without showing the stock and flow structure of a system. Or, as shown in Figure 6-8, they can include the stock and flow structure explicitly. When should you include the stock and flow structure, and when can you omit it? Generally, you should include stock and flow structures representing physical processes, delays, or stocks whose behavior is important in the dynamics you seek to explain. For example, consider the flow of a product through a supply chain from producer to consumer. The product travels through a network of stocks (inventories) and flows (shipment and delivery rates). The stock and flow representation for this process is shown in the top panel of Figure 6-10.

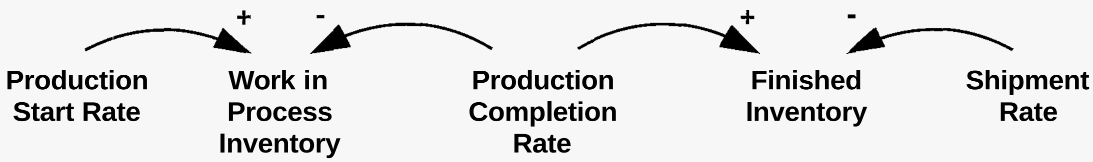
Production starts add to the stock of work in process (WIP) Inventory. The Production Completion Rate reduces the stock of WIP and increases the stock of

FIGURE 6-10 Stock and flow vs. causal diagram representations

Stock and Flow Representation of a Manufacturing Process



Causal Loop Diagram Representation of the Manufacturing Process



Finished Inventory. Shipments to customers deplete Finished Inventory. Equivalently, the stock of WIP accumulates production starts less completions, and the stock of finished inventory accumulates production completions less shipments.

The causal diagram representation is shown in the bottom panel of Figure 6-10. While technically correct, the causal diagram makes it hard to see the physical flow of product through the system and the conservation of material in the stock and flow chain. It is often confusing to interpret the polarities of the causal links when they involve stocks and flows. An increase in the Production Completion Rate causes Finished Inventory to rise above what it would have been otherwise (it rises at a faster rate), hence the polarity of the link is positive. A decrease in production completions, however, does not cause finished inventory to fall. Rather, a decrease in the production completion rate causes finished inventory to be less than it would have been. You cannot say whether finished inventory will be rising or falling based on the behavior of the production rate alone. Inventory will rise only when production completions exceed the shipment rate; that is, inventory rises only when we add to it faster than we remove units from it. You need to know the values of all the flows affecting a stock to determine its behavior. Richardson (1986a, 1997) carefully documents the pitfalls of causal diagrams, most of which involve the failure of causal diagrams to show the stock/flow distinction.

CHALLENGE Adding Stock and Flow Structure to Causal Diagrams

Consider the causal loop diagrams in chapter 5. For each, redraw the diagram showing the important stock and flow structure along with the feedback structure shown in the diagram. In particular, identify the main stocks and flows in the following conceptualization case studies presented in chapter 5:

1. The workload management example.
2. The oil industry and horse racing examples.
3. The traffic congestion example.

In each case, consider whether the explicit representation of the main stocks and flows enhances your ability to understand the dynamics of the system or merely clutters the diagram.

CHALLENGE

Linking Stock and Flow Structure with Feedback

Often understanding the dynamics of a system requires linking the feedback loop structure with the stock and flow structure. As an example, consider the gasoline shortages of the 1970s. In 1979 the United States (and some other industrialized nations) experienced a severe gasoline shortage. Iran's exports of oil dropped in the wake of the revolution there, and petroleum prices on the world market increased sharply. Within weeks, a shortage of gasoline began. Some service stations found their tanks emptied before the next delivery. Drivers, remembering the first oil embargo in 1973 and worried that they wouldn't be able to get gas, began to top off their tanks, instead of filling up only when the gas gauge fell toward empty. Soon, long lines of cars were seen idling in front of gas stations, and "Sorry—No Gas" signs sprouted along the highways of America as station after station found its underground tanks pumped dry. The shortage was the top story on the evening news—aerial footage of cars lined up around the block, close-ups of "No Gas" signs, and interviews with anxious drivers dominated the news. In some states, mandatory rationing was imposed, including limiting purchases to, for example, no more than $10 worth of gas. California imposed odd/even purchase rules: Drivers were allowed to buy gas only every other day, based on whether their license plate number was odd or even. It seemed that the supply of gasoline had been slashed.

Curiously, the impact of the Iranian revolution on the flow of oil to the US was small. True, US oil imports from the Persian Gulf (including Iran) fell by 500,000 barrels per day between 1978 and 1979, about 3% of US consumption, but imports from other nations increased by 640,000 barrels per day, so imports in 1979 actually *increased* by 140,000 barrels per day. Domestic production fell by 150,000 barrels per day, so total supply was essentially constant, while consumption fell by about 330,000 barrels per day, a drop of 2% from 1978. Plainly, for the year as a whole, there was no shortage. But if the flow of oil into the US was essentially constant, what caused the shortage? Where did the gas go?

First, develop a stock and flow map for the gasoline distribution system. You need not consider the entire supply chain for gasoline but can focus on retail distribution. Your diagram should begin with the flow of gasoline to service stations, then represent the stock and flow structure for its subsequent storage, sale, and eventual combustion.

Once you've mapped the stock and flow structure, identify the information inputs to the rates of flow in your diagram. Assume that the rate at which gasoline is delivered to service stations is exogenous. By identifying the information inputs to the flows in your stock and flow map, you will be closing some feedback loops, loops which should help explain why the shortage occurred and answer the question, Where did the gas go? Be sure to ask how individual drivers would learn about the shortage and what their behavior would then be.

Finally, using your diagram, assess the likely effectiveness of the maximum purchase and odd/even policies. Do policies of this type help ease the shortage or make it worse? Why? What policy would you recommend to ease the shortage? Explain why you think your policy would be effective in terms of the stock/flow and feedback structure of the system.

6.3.2 Aggregation in Stock and Flow Mapping

The ability to map the stocks and flows in a system is critical to effective modeling. Usually it is wise to identify the main stocks in a system and then the flows that alter those stocks. You must select an appropriate level of aggregation and boundary for these stock and flow maps. The level of aggregation refers to the number of internal categories or stocks represented. The boundary refers to how far upstream and downstream one chooses to represent the flows of materials and other quantities in the model.

To illustrate, consider the manufacturing process discussed above in which material flows from production starts through WIP inventory to finished inventory and finally shipment to the customer. All the various parts, components, and subassemblies are aggregated together into a single stock of WIP. And though the firm may carry tens of thousands of SKUs (stock keeping units), these individual items are all aggregated into a single stock of finished inventory. For many purposes the aggregate picture is sufficient. However, the model purpose might require more detail. If the purpose involved a closer look at the manufacturing process, you could disaggregate the stock of work in process serially to represent the different stages, such as part fabrication, assembly, and testing (Figure 6-11).

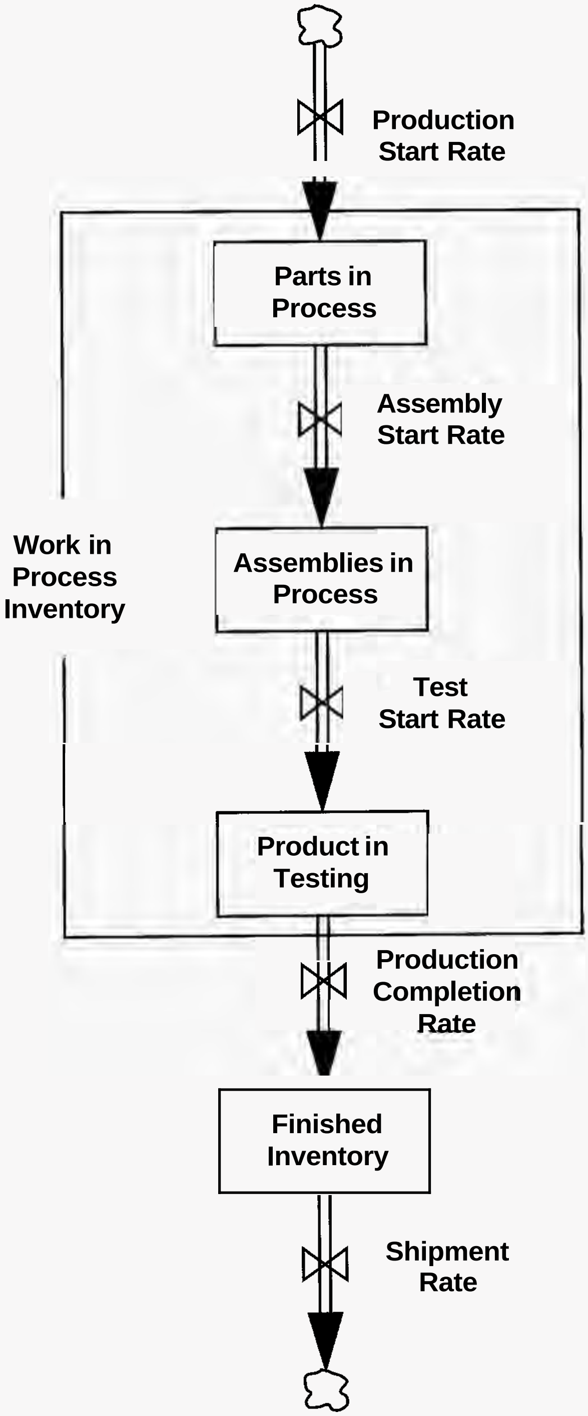
The sum of the three intermediate stocks is the total work in process inventory, but now the model tracks throughput at a finer level of resolution and can represent more potential bottlenecks in the production process. Note that in both the original, aggregate diagram and in this more detailed diagram there is no provision for rework or scrap. All units started are eventually completed—the flow of widgets through the system is conserved. Note also that as material flows through the system it is transformed from parts to finished product. To maintain consistent units of measure we might measure parts in widget equivalents—that is, a widget's worth of parts. If necessary for the purpose, you can further disaggregate the stock and flow structure.

CHALLENGE

Modifying Stock and Flow Maps

1. Modify the diagram in Figure 6-11 to represent the case where units that fail testing are scrapped.
2. Modify your diagram to represent the case where items failing testing are returned to assembly for rework.

FIGURE 6-11 Disaggregated stock and flow map for a manufacturing process



CHALLENGE

Disaggregation

Each of the three stages of WIP identified in Figure 6-11 consists of other steps, each with its own stock and flow structure. Suppose you learn that part fabrication at the plant you are modeling actually requires several operations: welding, grinding, and painting. Observation of the grinding operation reveals that workers draw parts ready for grinding from a buffer generated by the welding operation. When grinding is completed, the parts are placed in a bin which then goes on to the next operation (painting). The welding and paint shops are similar. Draw the disaggregated stock and flow map for the part fabrication step to show the welding, grinding, and painting operations explicitly.

Up to now the discussion has focused on ***serial*** disaggregation: how finely to break down the stages of processing. Throughout, the many different parts and products produced by a typical firm are aggregated into a single chain of stocks and flows. In many situations the process occurs not only in series but also involves parallel activities. You could of course replicate the main stock and flow chain for each product (many simulation software packages support array structures for this purpose). When there are multiple, parallel activities you must make a decision not only about the number of stages of the process to represent but also how much to aggregate the different parallel processes together. For example, the assembly process for automobiles involves integrating the chassis and engine. Each subassembly is built on a separate line, often in plants far from the final assembly point. Suppose the client argues that you can't aggregate all subcomponents into a single flow of parts, but must separate chassis and engine fabrication (omit the body for simplicity). The stock and flow map for the assembly process might be shown as in Figure 6-12.

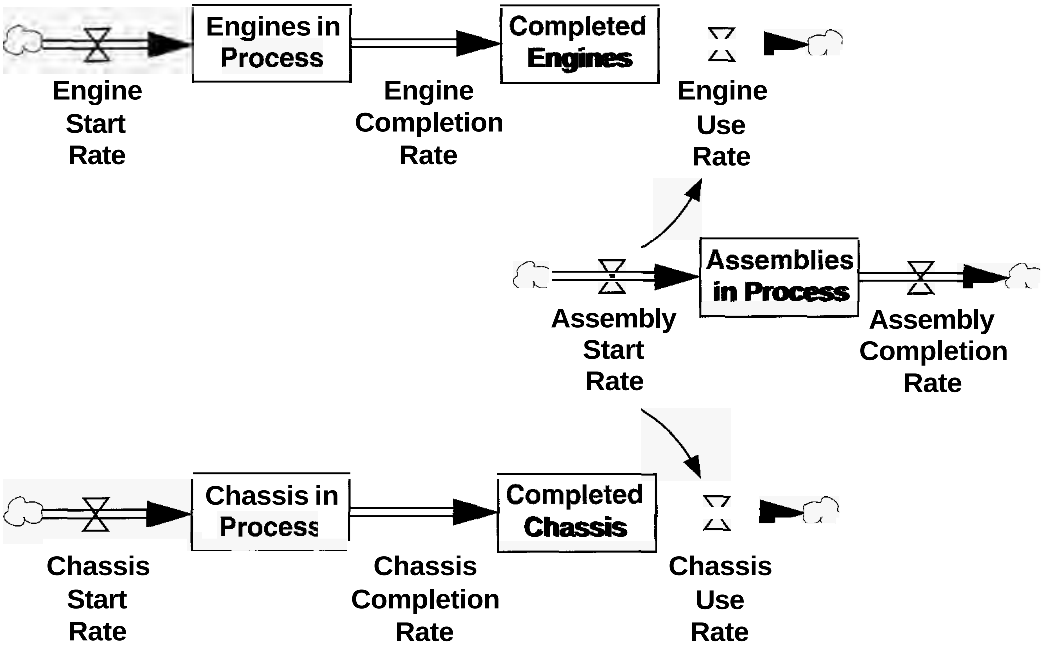
There are now three distinct stock and flow chains, one each for engines, chassis, and assembled cars. Because the three chains are separate, each can be measured in different units: engines, chassis, and cars. The three chains are linked because each car beginning the final assembly process requires one engine from the stock of completed engines and one chassis from the stock of completed chassis. The information arrows from the assembly rate to the engine and chassis use rates show these links. The number of engines and chassis available also determine the maximum assembly start rate, which in turn constrains actual assembly starts: If either component buffer falls to zero, assembly must cease.[2] These links (not shown) define two balancing feedbacks that regulate the outflows from the stocks of components and couple the stock and flow networks. The diagram does not represent many other information flows that must exist in the system (such as the determinants of the chassis start and completion rates); try adding these to the map.

You could of course continue to disaggregate. The process can be broken down into more steps: The paint process, for example, actually consists of multiple activities separated by buffers such as part preparation (solvent bath), drying, spraying the first coat, drying in the oven, spraying the second coat, drying, and so on, with various inspections along the way. You could also continue the parallel disaggregation by splitting the engine or chassis assembly processes into their

---

[2] Firms can sometimes build incomplete units and add the missing components later. When a 1997 strike shut down a critical supplier, Ford continued to assemble its popular Explorer, storing the nearly completed cars until the missing parts became available and could be retrofitted (try modifying the diagram in Figure 6-12 to accommodate such retrofitting).

FIGURE 6-12 Disaggregating parallel activities



subassemblies. In the limit each and every part and operation would be represented separately. Obviously such a model would be just as complex as the real system, at least as hard to understand, and quite useless.

Where should you stop? How much detail is necessary? This is always a matter of judgment to be made by considering the model purpose and the needs of the client. If the purpose is representing the lag in the response of the manufacturing system to changes in demand as part of a larger model of firm strategy, the simpler representation is probably fine. If you seek to reengineer the flow of material through the production line, a more detailed representation is required. It is better to start with a high-level, aggregate representation and add detail if needed to address the purpose. Beginning with detailed process maps often leads to paralysis due to their complexity, data requirements, and rapid obsolescence rates. The aggregate map showing only production starts, WIP, production, and finished inventory is quite stable and remains appropriate even as the details of the production process change, while a detailed map may become obsolete as new products, tooling, or process technologies are introduced.

6.3.3 Guidelines for Aggregation

When is it appropriate to aggregate different activities together? To determine whether activities taking place serially can be aggregated, consider the average residence time of items in each stock (the average time between entering and exiting the stock). Stocks with short residence times relative to the time scale for the dynamics of interest generally do not need to be represented explicitly and can either be omitted or lumped into adjacent stocks. For example, in the long-term planning models of the US energy system developed by the US Department of Energy (Naill 1992), various stocks of undiscovered petroleum and known reserves are explicitly represented because their lifetimes range from years to decades (at current production rates). However, stocks of crude oil and refined products in the petroleum supply chain represent only a few months of consumption. In a long-term model these stocks are too short-lived to require explicit treatment. They fluctuate around their equilibrium values as producers, refiners, and distributors react to changes in inventory. In a model of short-term movements in spot petroleum prices, however, these stocks are critically important. **A** good model would represent the key stocks in the petroleum supply chain explicitly, perhaps even including a separate stock for the inventory of gasoline at retail service stations and the inventory of gasoline in the tanks of cars on the road. On the other hand, a short-term spot price model need not include petroleum reserves or undiscovered resources, as these stocks change too slowly to influence the spot market over a time horizon of a year.

Parallel activities can legitimately be aggregated together if the individual flows are governed by similar decision rules and if the time the different items spend in the individual stocks is similar. For example, it is often appropriate to aggregate the many parts required to manufacture a product into a small number of categories since they are usually ordered using the same procedures and their delivery lead times and residence times in parts inventories generally don't differ too much. In contrast, plant and equipment sometimes must be disaggregated. Their lifetimes are very different, and the decision rules for new green-field facilities differ substantially from those used to order equipment for existing facilities due to differences in lead times, costs, financing, permitting, and regulatory issues.

As a rule of thumb, clients generally want to see more detail in a model than the modeler thinks is needed, and modelers, in turn, generally overestimate the detail necessary to capture the dynamics of interest. Of course, the amount of detail needed to capture the dynamics relevant to the client's purpose and the amount of detail needed to give the client confidence in the results are two different things. Roberts (1977/1978) estimated that clients often require twice as much detail as the modeler feels is needed to feel comfortable with and accept a model as a basis for action, and in my experience this is often an underestimate. Success requires you to include the detail necessary to satisfy the client. But this does not mean you should acquiesce to all client demands for more detail—you will end up with an expensive and useless black box. You must work with the client to understand why excessive detail is often unnecessary. Often, models end up with too much detail, but as the client gains confidence and understanding of the important feedbacks driving the dynamics, the excess structure can be eliminated, resulting in a simpler, more easily maintained, and more useful model (Randers 1980). Still, Roberts is correct: "You must provide the level of detail that causes [the client] to be persuaded that you have properly taken into account his issues, his questions, his level of concerns. Otherwise he will not believe the model you have built, he will not accept it, and he will not use it" (p. SO).

6.3.4 System Dynamics in Action: Modeling Large-Scale Construction Projects

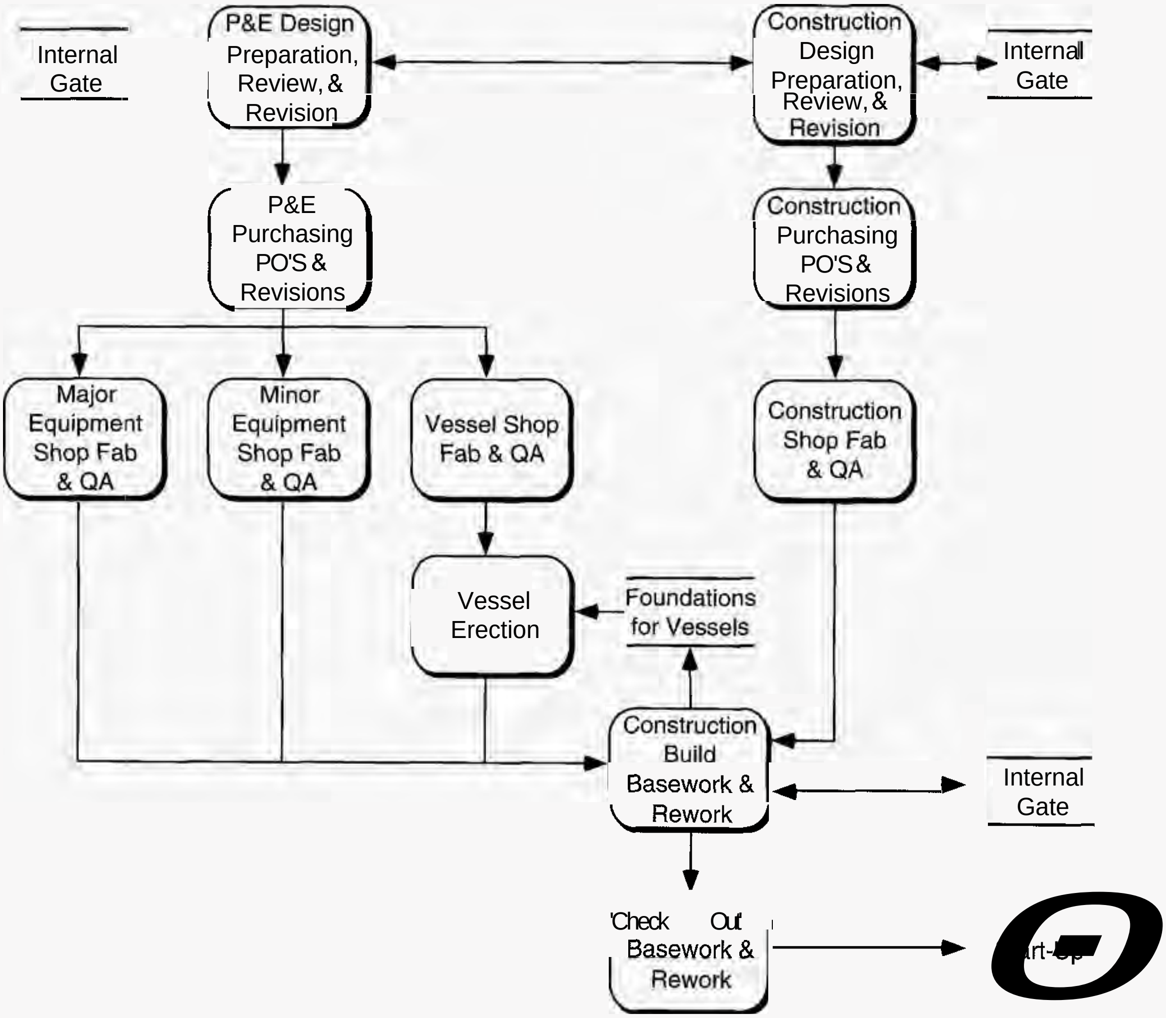
Aggregation of multiple serial and parallel activities is well illustrated in a model of large construction projects developed by Jack Homer (Homer et al. 1993). The client was a multinational forest products company, specifically the division of the company that designs and builds pulp and paper mills. Competition for the small number of mills built each year was intensifying as the industry globalized, and the firm, already a leader, saw that to remain strong they had to dramatically reduce the time required to design and build mills. Their goal was to reduce significantly the total cycle time, from the handshake with a customer to the handoff of a working mill, without increasing costs. They knew traditional project management techniques were not adequate: the design and construction process is exceedingly complex, with tight couplings among the phases, and they had already done all the easy things. They decided to develop a system dynamics model of the entire engineering, procurement, and construction (EPC) process.

Early meetings of the project team focused on the model boundary and aggregation, in particular, descriptions of the stock and flow structure of a typical project. Many issues were raised: Is there a typical project? How much detail is needed? What activities could be aggregated together? One member of the client team argued that the model couldn't be useful if it didn't represent every engineering drawing, every purchase order, and every component installed at the site. Obviously, such a model could never be built or made useful. Other members of the client team argued for a simpler approach. They already had highly disaggregate scheduling and planning models based on traditional project management tools to manage the detail complexity of the projects. They lacked a tool to manage the dynamic complexity and interdependencies among the phases and activities of the projects.

After extensive discussion, an initial model boundary and level of aggregation were set (Figure 6-13). The figure is a high-level subsystem diagram showing how projects were aggregated into a reasonable number of phases. The overall project was divided into two main stock and flow chains representing P&E (process and equipment) and construction. Each activity goes through design preparation, review, and design revisions. Next suppliers are selected and purchase orders are issued. The suppliers then fabricate the materials needed for each activity. On the construction side, the client felt it was acceptable to aggregate all construction materials (e.g., structural steel, concrete forms, rebar) into a single category. The process and equipment side, however, was divided into three categories: reactor vessels, major equipment (e.g., large tanks, pipelines, and conveyors), and minor equipment (e.g., pumps, motors, valves, and instrumentation). The design, procurement, and construction of these types of equipment are sufficiently different in scope, duration, and cost that they could not reasonably be lumped together. The reactor vessels, in particular, had to be modeled in more detail as they are the largest subassembly, almost always fall on the critical path, and are frequently a bottleneck constraining construction progress. During construction, reactor vessels, other equipment, and site preparation such as foundations and grading all

FIGURE 6-13 Building a pulp and paper mill

Subsystem diagram showing flows of engineering, procurement, and construction work in a model of a pulp mill construction project. The diagram illustrates the sector boundaries and level of aggregation without showing all details. Each block represents a project phase, modeled with a generic module with roughly the same structure. An internal gate captures the constraints on work available within a phase as a function of the work completed. For example, foundations cannot be completed until surveying and site preparation are done; see section 14.5.



Source: Homer et al. (1993).

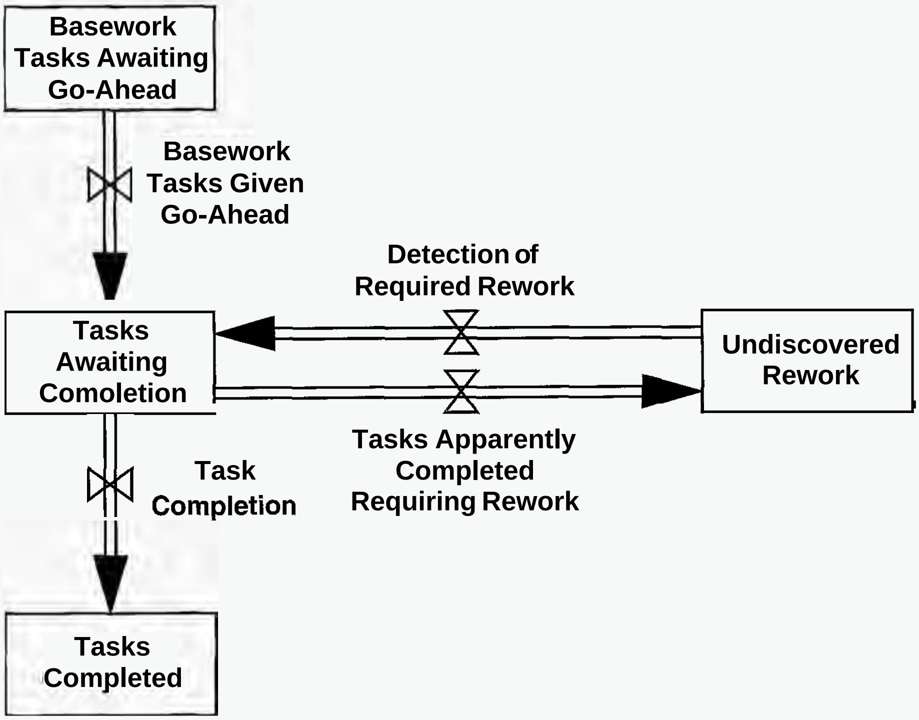
must come together, followed by a functionality check out, start-up and, finally, handoff to the customer.

Each block in Figure 6-13 represents a project phase. The model consisted of a generic project phase module, replicated for each block and linked as shown. Each module contained a stock and flow structure including the flows of tasks within the phase along with scheduled deadlines, the labor force dedicated to the phase, worker productivity, fatigue levels, error rates, and costs. The stock and flow structure for tasks within a phase models the progression of tasks from basework through completion (Figure 6-14). In general the tasks to be completed in a phase can only be done as upstream tasks upon which they depend are completed (the rate at which basework tasks become available). For example, the reactor vessels cannot be erected on site until their foundations are completed. Likewise, not all tasks within a given phase can be done concurrently. For example, the detailed design of conveyors and pipelines between the chippers, reactor vessels, and paper machines cannot be done until the high-level physical design of the plant is completed. These within- and between-phase dependencies were modeled explicitly. The flow of work from the stock of tasks awaiting completion to the stock of tasks requiring rework represents those tasks completed incorrectly or rendered obsolete by changes in other subsystems. Generally, errors are not detected immediately and the delay in the discovery of rework can be substantial, as when a design error is not detected until construction in the field is underway. The discovery of rework moves tasks thought to be complete back into the stock of tasks awaiting completion (see Ford and Sterman 1998b for a detailed and fully documented model of a multiphase project similar to the one used here; see also the shipbuilding project model described in section 2.3).

The model was substantially simpler than the client firm's detailed project planning model, which included literally thousands of individual activities (high detail complexity) but no feedback loops (no dynamic complexity). It was disaggregated enough to capture important interdependencies among design, procurement, and construction activities and between construction and the various types of

FIGURE 6-14
Stock and flow structure of tasks in a project phase

Simplified representation of the stock and flow structure of a phase in the pulp mill project model. Determinants of the flows and couplings among the different phases are not shown.



equipment. The model could capture shifts in the critical path that might result from policies accelerating the fabrication of the reactor vessels, a policy favored by some client team members.

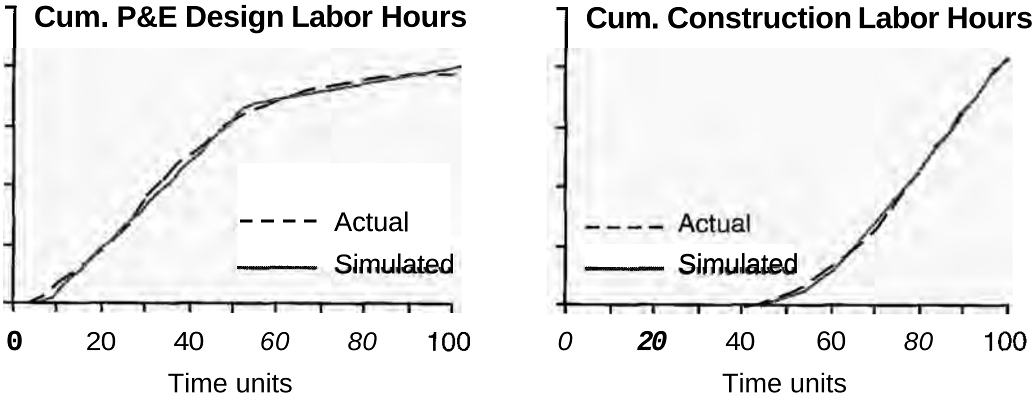
It is important to note that the process of developing the final level of aggregation involved a number of iterations and revisions. And though the model represents the project at a high level of aggregation, the modeling team developed many more detailed diagrams. These more detailed maps helped the modeler and client team discover flaws in their thinking, estimate parameters better, and deepen their understanding of the process. And they developed confidence that the more aggregate representation in the simulation model was acceptable for their purpose so these more detailed stock and flow structures did not have to be incorporated into the model.

The level of detail selected also permitted the model to be calibrated against a wide range of data collected on one of the company's current EPC projects. The model successfully (that is, to the satisfaction of the client) reproduced all relevant project activities, including the various workforces and labor hours, overtime and rework rates, purchase order volumes and revision rates, vendor shipments, and the progress of vessel erection and construction (Figure 6-15 shows an example).

While the clients prefer not to disclose the details of policy recommendations, they viewed the model as credible and useful and developed confidence, shared among the team, that the model did a good job of representing their EPC projects. They used the model to analyze many policies and identified several which, while previously appearing to be desirable, in fact generated harmful side effects. The model also helped identify policies that reduced project delivery times by at least 30% within a few years. Several of the policies were not apparent to the client team or were hotly debated prior to the modeling effort. The modeling process helped build understanding of and consensus around these controversial initiatives, helping the firm successfully implement many of the recommendations.

FIGURE 6-15 Sample comparison of historical and simulated behavior of the pulp mill model

P&E = process and equipment.



Note: Time is expressed as time units to protect client confidential information.
Source: Homer et al. (1993).

6.3.5 Setting the Model Boundary: "Challenging the Clouds"

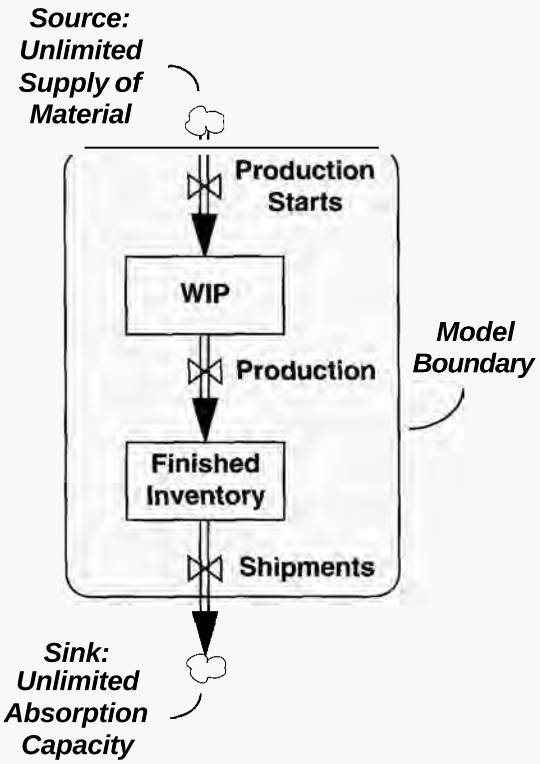
Mapping the stock and flow structure of a system involves important decisions about the boundary of the model. In reality, flows of material, people, and money into a stock have to come from somewhere; the flows out have to go somewhere. To keep your models manageable, you must truncate these chains using sources and sinks, represented in the stock and flow maps by ‘‘clouds’’; see Figure 6-1. Sources and sinks represent the stocks supplying material to or absorbing material from the modeled system. Sources and sinks are assumed to have infinite capacity and can never constrain the flows they support. In the real world, the stocks supplying or absorbing flows have finite capacity and do influence the flows. When you truncate a stock and flow chain with a cloud you are setting the boundary of the model—stocks and flows beyond this point are ignored; you exclude all possible feedbacks from or interactions with the stocks outside the boundary.

As a modeler you must critically examine these boundary assumptions; you must, in the words of Barry Richmond (1993, p. 132), "challenge the clouds." Is it appropriate for your purpose to exclude the stocks outside the boundary of the model? What feedbacks ignored by your model might exist in the real world, and might they affect your policy recommendations? Can the sources for the flows be depleted and constrain the inflow? Can the sinks be filled and block the outflows, backing up the system like a clogged drain?

Consider the automobile industry. A stock and flow map for automobile production might begin with production starts, WIP inventory, production, finished inventory, and shipments (Figure 6-16). Drawing the map with a source for the production start flow presumes that the supply of parts is unlimited and can never constrain the production start rate. Likewise, because shipments flow to a sink, the modeler has assumed stocks of product in the hands of dealers and customers have no effect on shipments. In challenging the clouds you ask whether these assumptions are reasonable. For the auto industry they are not. Production starts require the automaker to have an adequate stock of parts. Yet parts stocks may easily be depleted. Suppliers cannot respond instantly to changes in parts orders. Large orders may outstrip supplier capacity, leading to shortages. A strike at a supplier may interrupt the flow of parts to the firm. At the other end, shipments of new cars to dealers depend on the size of dealer stocks. Dealers generally try to maintain about 40 to 60 days of inventory on their lots; this is enough to provide good selection for consumers without carrying excessive and costly inventory. If stocks are low relative to their targets, dealers order more from the manufacturers; if stocks are high, they cut back. Figure 6-17 expands the model boundary to capture these effects. The model now represents three distinct organizational entities—suppliers, manufacturers, and dealers. The inventory of parts held by the manufacturer is now explicit. The supplier has the same basic structure as the automaker: a stock of finished inventory and a stock of work in process. At the shipment end, manufacturer shipments no longer disappear into a sink but flow into dealer stocks, allowing you to model the purchase rate as a function of the dealer inventory and sales to customers.

FIGURE 6-16 initial stock and flow map for the automobile industry, showing the model boundary

The sources arid sinks for the flows through the system are assumed to be infinite and can have no impact on the dynamics.



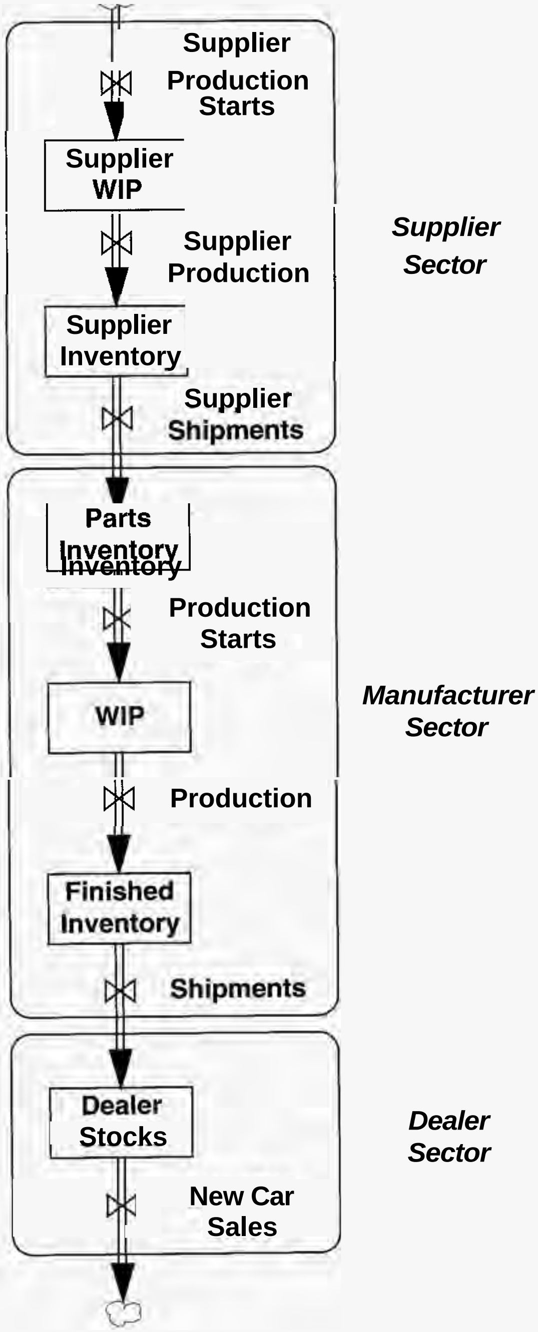
You could and should continue to challenge the boundary of the model. The model now allows you to represent supplier order processing, inventory management, and delivery, including the possibility that suppliers can become a bottleneck and starve automobile production. But now the suppliers are assumed to have unlimited parts and raw materials availability. Is this appropriate? It depends on the model purpose. You could continue to expand the model boundary by adding the suppliers to the suppliers, and their suppliers, and so on, until you reached the point where it is acceptable to assume that the supply of materials to the farthest upstream supplier is unlimited. Alternatively, you could represent the entire upstream supply chain by a single aggregate supplier stage.

The map shown in Figure 6-17 also assumes that dealer sales flow into a sink so there is no feedback from the stock of cars on the road to purchases of new cars. This is obviously a bad assumption: Sales of new cars depend on the number and age of the cars people already have relative to their needs. People who have just acquired a new car are unlikely to buy another for several years, until their loan is paid off, their lease expires, or their car is involved in an accident and must be replaced (see section 2.2). Figure 6-18 expands the downstream end of the stock and flow map to include the stock of cars on the road.

You can continue to challenge the model boundary. What happens to the cars when they are scrapped? In the current map, they simply disappear. In reality, they don't. In North America some 10 to 12 million vehicles are scrapped per year. Roughly 94% are shredded and the steel and some nonferrous metals are

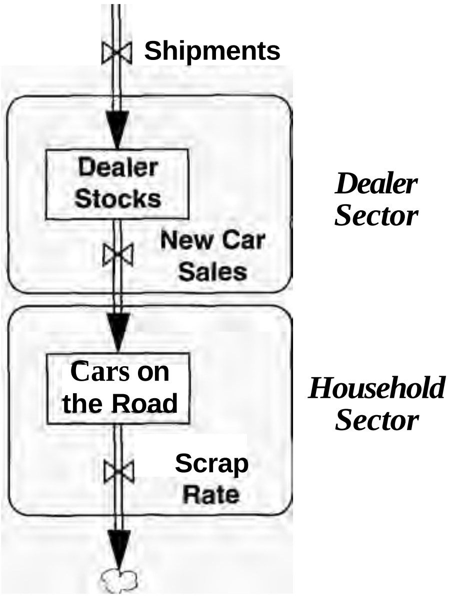
FIGURE 6-17
Challenging the clouds

Adding a supplier and dealer sector to the stock and flow chain for automobile production. Rectangles with rounded corners denote the boundaries between different organizational entities and decision-making units.



recovered, one of the highest recycling fractions of any industry. However, some cars end up abandoned as dangerous eyesores on the side of the road. And much of the plastic, glass, and other nonmetal materials end up in landfills, constituting a significant source of pollution (more than two billion discarded tires, most sitting in huge piles across the country, have already accumulated in the US).

FIGURE 6-18
Expanded automobile model
The boundary now includes the stock of cars on the road, which feeds back to influence sales of new cars.



6.3.6 System Dynamics in Action: Automobile Recycling

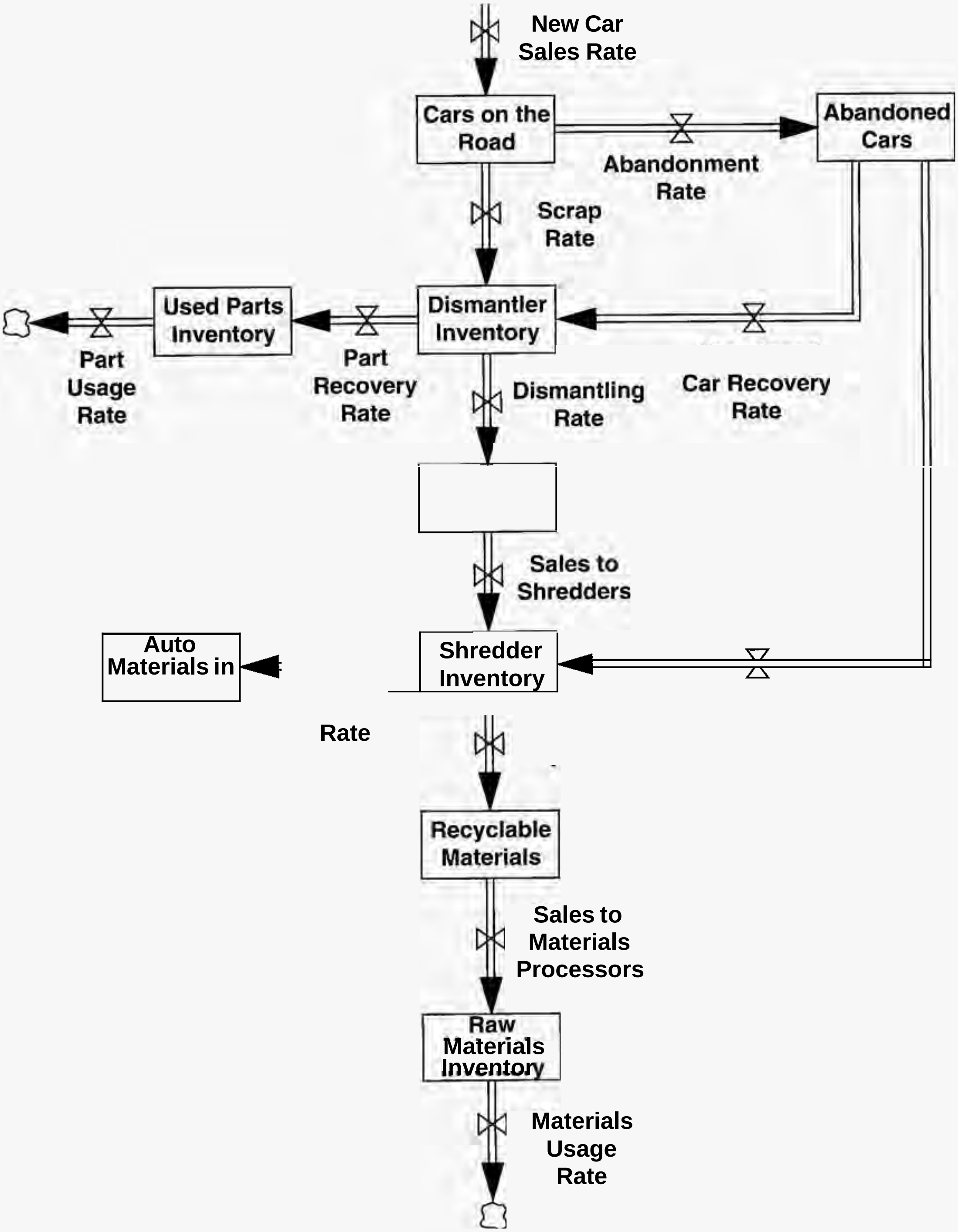
By the mid 1990s, as landfills filled and environmental awareness grew, pressure built to recycle more of the material in cars. Germany debated a law that would require auto manufacturers to take back their old cars when people deregistered them. Pushed by these forces, the auto industry, first in Europe and then in the US, began to study ways to increase the recovery of parts and the recycling of materials from cars.

Pavel Zamudio-Ramirez (1996) modeled part recovery and the materials recycling in the US auto industry to help the industry think about a future of enhanced auto recycling. Figure 6-19 shows a simplified stock and flow structure adapted from the model. Old or wrecked cars can either be scrapped legally (sold to a junkyard or dismantler) or illegally abandoned. The stock of abandoned, often burned-out, cars is a blight on the landscape and significant source of pollution. There are two outflows from the stock of illegally abandoned cars: Dismantlers will process them if the value of the recoverable parts and materials is high enough. Alternatively, illegally dumped cars can be collected (say by local governments) and taken to shredders for proper disposal. Both these flows are relatively small, so the stock of abandoned cars can build up to high levels even if the abandonment rate is low.

Cars held in the dismantlers' inventories are stripped of those parts whose value exceeds the cost of recovery. These parts enter a used parts stock and are then sold to repair shops and used to replace worn or damaged parts on operating cars. In this map, the part usage rate flows into a sink. In actuality, these parts are installed in cars still on the road and eventually flow again through the scrap or abandonment rate. Since the number of recovered parts is very small relative to the total flow of materials through the system, this omission is probably reasonable.

FIGURE 6-19 Stock and flow map for a model of automobile recycling

The stock and flow structure for the development of new vehicle platforms, defining the mass and materials composition of cars and level of design for disassembly, is not shown. The model includes a parallel stock and flow structure (co-flow) tracking each of these properties as vehicles age and are eventually retired, dismantled, and shredded. See chapter 12.



After all parts worth recovering are removed, the gutted car, now called a hulk, is sold to a shredder. In the mid 1990s there were about 200 shredders in the US who processed roughly 94% of all deregistered cars. After shredding, the valuable materials (principally steel and some nonferrous metals) are separated out for recycling. If the prices of the recovered materials don't justify the cost, shredders can take hulks directly to a landfill and cut their purchases from dismantlers. What remains after shredding and separation is a mixture of plastics, glass, elastomers, and some unrecovered metal called automotive shredder residue (ASR) or "fluff," which is then landfilled. ASR is one of the major environmental concerns generated by the disposal of old cars.

The recyclable materials accumulate in an inventory and are eventually sold to materials processors such as steel mills. The inventory of raw materials is then used to manufacture new products, including automobiles, thus helping to create a closed material flow and cutting the use of nonrenewable resources. As in the case of parts, the materials usage rate flows into a sink since the flow of recovered materials relative to the total flow of virgin materials is small.

Zamudio-Ramirez's model included a rich feedback structure representing the behavior of the various actors in the system, including the automakers, car owners, dismantlers, and shredders. Markets for recovered materials were explicit. The stock and flow structure for autos began at the design stage for new models and platforms and tracked key properties of the cars including their mass, materials composition (ferrous, nonferrous, plastics), and the level of design for disassembly built into the design. These attributes were tracked as the cars embodying them moved from the design stage to market, age, and are then retired, dismantled, and shredded.

To gather the required data, Zamudio-Ramirez conducted interviews with various actors, including carmakers, dismantlers, shredders, and industry analysts and made extensive use of various auto and recycling industry databases. Some of the data required, such as age-dependent scrap rates for cars, were relatively easy to gather. Other key parameters were not. Two critical relationships in the model are the supply curves for recovered parts and recovered materials. That is, how will the number of parts recovered by dismantlers vary as the price they can get and the costs of recovery vary?

Estimating the parts supply curve is a daunting problem. The principal cost of recovery is the labor time required to remove a part. But the time required to remove a given part depends on how many other parts must be removed first. These precedence relationships depend on the design of the car and the value of the intervening parts (can the seat be ripped out quickly to get at a valuable part under it or must it be removed carefully? Should workers get at a part from in front or behind?). To estimate these relationships Zamudio-Ramirez worked at the Vehicle Recycling Partnership, a consortium of the Big Three US automakers, dismantlers, and the recycling industry. The Vehicle Recycling Partnership assembled a comprehensive database of part removal times by completely disassembling a variety of late model cars. Zamudio-Ramirez and his colleague Andrew Spicer then developed an optimization model to estimate the supply curve for parts recovery as functions of part and materials prices, labor costs, and the design of the vehicles. The optimization model determined the number of parts worth recovering and the optimal dismantling order for any set of prices, labor costs, and design parameters—the supply curve for recovered parts. The results of the optimization model were then embedded in the simulation model. As the design parameters for cars change and the removal time for key parts falls, the estimated supply curve responds by realistically increasing the number and types of parts recovered.

Though the stock and flow structure in Figure 6-19 is simplified and does not show any of the feedback structure determining the various flows from the full model, it illustrates the response of the automobile and materials markets to policies designed to increase recycling of cars.

First consider the effect of a design for disassembly (DFD) program designed to increase the part recovery rate and reduce the amount of fluff ending up in landfills. DFD can reduce the labor cost of part recovery through better design, different choice of part fasteners, improved selection and labeling of materials, and other techniques. The first effect is...nothing. There is a lag of at least several years between the time an automaker starts a DFD program and the time the first cars designed to those specs roll off the assembly line. The average car in the United States stays on the road for about a decade, and new cars have very low scrap rates (most of these are wrecks declared total losses by insurance companies). Only after a delay of many years will the stock of recycling-ready cars be large enough and old enough for them to constitute a significant fraction of the scrapped cars purchased by dismantlers.

What then happens? Manufacturers expected DFD would eventually cause part and material recovery to rise, permanently reducing the flow of materials to landfills. Instead, the model suggests the next effect will be a glut of used parts, as the part recovery rate rises above the used parts usage rate. As parts inventories build, the price dismantlers can get for used parts falls. The number of parts that can be economically recovered drops, and the dismantling rate drops back. Prices continue to fall until the number of parts recovered falls enough to balance the used parts usage rate. The part usage rate may rise, stimulated by lower prices, but unless the demand for used parts is highly price elastic, the part recovery rate will drop back close to its original rate prior to DFD. The demand for used parts is likely to be rather insensitive to price. Automakers and third-party producers of replacement parts will be reluctant to lose the lucrative parts market and may be able to prohibit the use of recovered parts by authorized service centers or for warranty repairs or compete on price. If the demand for used parts is inelastic, the principal effect of DFD might simply be to depress the price of used parts, offsetting most of the benefit of improved design.

Now consider the effect of a trend toward smaller, lighter cars with significantly higher plastic content and less steel and metal. Such changes are promoted to improve fuel economy, increase part recoverability, and decrease the quantity of fluff ending up in landfills. However, the stock and flow structure may cause the impact of such policies to be counter to their intent. The auto industry is a significant consumer of steel. When new cars begin to use less, the recovery of steel from shredding of old hulks continues at the prior rate. The price of scrap metal will fall, reducing shredder profitability. The number of hulks shredded and the quantity of metals recovered may fall, and the volume of fluff disposed in landfills may actually rise. Further, once the scrap rate of cars with reduced steel content increases, shredder profit can fall further. With less steel and nonferrous content, shredder revenue per hulk falls, while the fixed costs of shredding remain the same. Zamudio-Ramirez found that a sustained increase in the plastic content of cars, as expected, would increase the fraction of materials recovered by dismantlers. But cars with less recyclable metal could also depress hulk prices enough to cut shredder profit, decrease the shredding rate, and actually increase the number of abandoned cars and the amount of fluff buried in landfills.

The stock and flow map helps illustrate the long delays between a change in the design of cars and the flows of old cars to landfills. By making the stocks of recovered parts and materials explicit, it is easier to see that there is imperfect coordination between inflows and outflows, leading to potential imbalances and changes in prices that invalidate the assumptions behind recycling programs. Institutional structures such as requirements that service centers use new replacement parts can overwhelm the logic of the market. Market mechanisms, even when present, are not likely to work smoothly, possibly leading to instability and inefficiency. Similar dynamics have already been observed in the market for recycled paper (Taylor 1999). Supply side steps to increase recyclability alone are not likely to be effective unless matched by policies to increase the usage of recovered parts and materials. The collection of recyclable materials and the actual recycling of those materials aren’t the same thing.

6.4 SUMMARY

This chapter introduced the stock and flow concept. Stocks accumulate their inflows less their outflows. Stocks are the states of the system upon which decisions and actions are based, are the source of inertia and memory in systems, create delays, and generate disequilibrium dynamics by decoupling rates of flow. The diagramming notation for stocks and flows can be used with a wide range of audiences and makes it easier to relate a causal diagram to the dynamics of the system. Stocks accumulate (integrate) their inflows less their outflows. Equivalently, the rate of change of a stock is the total inflow less the total outflow. Thus a stock and flow map corresponds exactly to a system of integral or differential equations. However, stock and flow maps are much easier to work with and explain.

There are several ways to identify the stocks in systems. In the snapshot test you imagine freezing the system at a moment of time—the measurable quantities (physical, informational, and psychological) are the stocks, while flows are not instantaneously observable or measurable. Units of measure can also help identify stocks and flows. If a stock is measured in units, its flows must be measured in units per time period.

Stocks existing in series in a network can be aggregated together if they are short-lived relative to the time horizon and dynamics of interest. Multiple parallel activities can be aggregated into a single stock and flow network if the activities are governed by similar decision processes and utilize similar resources and if the residence times of the items in the stocks is similar enough for the purpose of your model.

Sources and sinks for the flows in a system have infinite capacity, unlike stocks in the real world, and thus represent the boundary of the model. Modelers should always challenge these boundary assumptions, asking if the assumption of infinite supply for sources and infinite absorption capacity for sinks is appropriate relative to the model purpose.