3
The Nondurable Goods Index

3.1 Introduction

Governments and other institutions use a host of statistical summaries to track aspects of society across time and space. These range from simple counts of events such as deaths from lung cancer to sophisticated summaries of complex processes. For instance, inflation is monitored by the cost of completing a shopping list carefully designed to reflect the purchases that most citizens would find essential. To give another example, indices such as the Dow Jones summarize stock market performance.

The index of nondurable goods manufacturing for the United States, plotted in Figure 3.1, is a monthly indicator reflecting the producton of goods that wear out within two years, such as food, tobacco, clothing, paper products, fuels, and utilities. Because such items are, in normal times, repeatedly purchased, the index reflects the economy of everyday life. When times are good, people exhibit strong and stable spending patterns, but shocks such as the collapse of the stock market in 1929 and the onset of World War II (1939 in Europe and 1941 in the United States) produce both short-lived transitory effects, and longer-lasting readjustments of lifestyles. Technical innovations such as the development of the personal computer in the early 1980s affect both consumer habits and the production process itself. You can access these data from the Web site for this chapter.

In this and most economic indicators, there is a multilayered structure. There are overall trends that span a century or more, and we see in Figure 3.1 that there is a broad tendency for exponential or geometric increase.

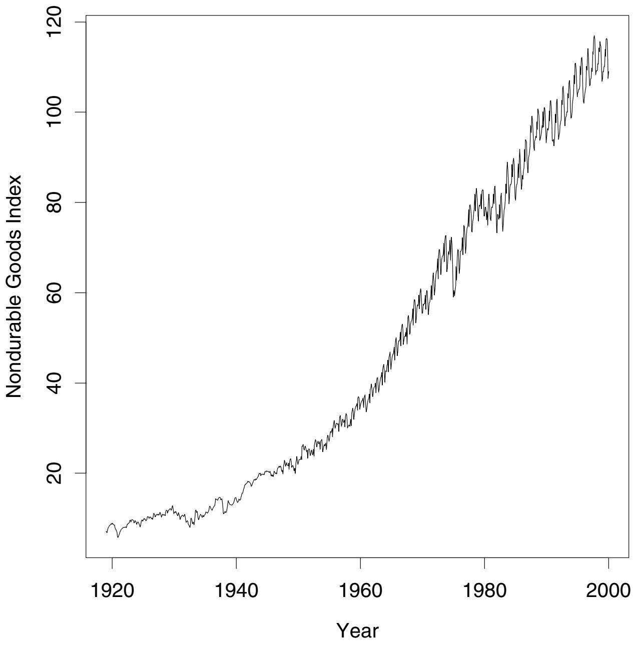

Figure 3.1. Monthly nondurable goods manufacturing for the United States.

Long-term changes last decades, medium-term effects such as recessions last a number of years, and short-term shocks such as the beginning and end of wars are over in a year or two.

We see by the ripples in Figure 3.1 that there is an important seasonal variation in the index. The index includes items often given as gifts, so there is a surge in the index in the last part of each year, followed by a low period in January and February. The beginning of the school year requires new clothes, and we expect to see another surge in the preceding months. On the supply side, though, we need people in the manufacturing process, and vacation periods such as the summer holidays will necessarily have an impact on factory activities.

This seasonal variation is also affected by changes in the economy at various time scales, and so we also want to study how the within-year variation evolves. Perhaps the evolution of seasonal variation can tell us something interesting about how the economy evolves in normal times, and how it reacts to times of crisis and structural change. How did the outbreak of World War II change the seasonal pattern? What about the moving off-shore of a great deal of manufacturing in recent decades? But Figure 3.1 covers too long a time span to reveal much, and we will need to consider some new ways of plotting the seasonal trend.

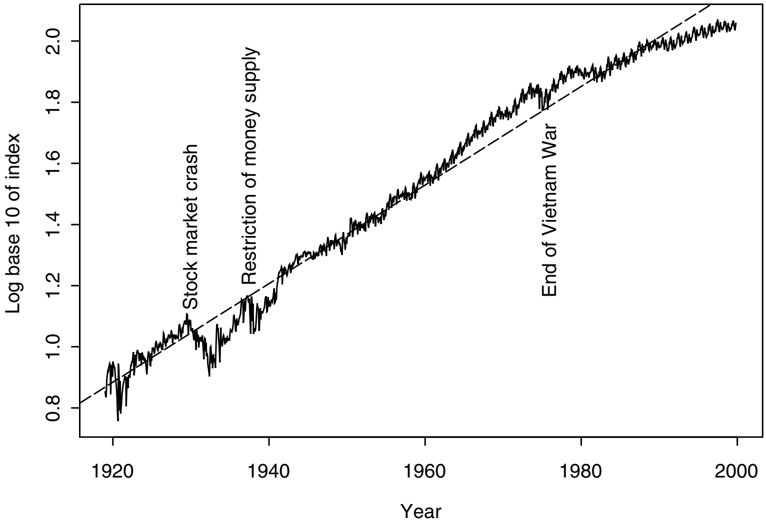

Figure 3.2. The monthly nondurable goods production shown in Figure 3.1 plotted on a logarithmic scale. The dotted straight line is estimated by least squares regression, and has a slope of 0.016, corresponding to a 1.6% increase in the index per year.

3.2 Transformation and smoothing

Like most economic indicators, the nondurable goods index tends to exhibit exponential increase, corresponding to percentage increases over fixed time periods. Moreover, the index tends to increase in size and volatility at the same time, so that the large relative effects surrounding the Second World War seem to be small relative to the large changes in the 1970s and 1980s, and seasonal variation in recent years dwarfs that in early years.

We prefer, therefore, to study the logarithm of this index, displayed in Figure 3.2. The log index has a linear trend with a slope of 0.016, corresponding to an annual rate of increase of 1.6%, and the sizes of the seasonal cycles are also more comparable across time. We now see that the changes in the Great Depression and the war periods are now much more substantial and abrupt than those in recent times. The growth rate is especially high from 1960 to 1975, when the baby boom was in the years of peak consumption; but in subsequent years seems to be substantially lower, perhaps because middle-aged "boomers" consume less, or possibly because the nature of the index itself has changed.

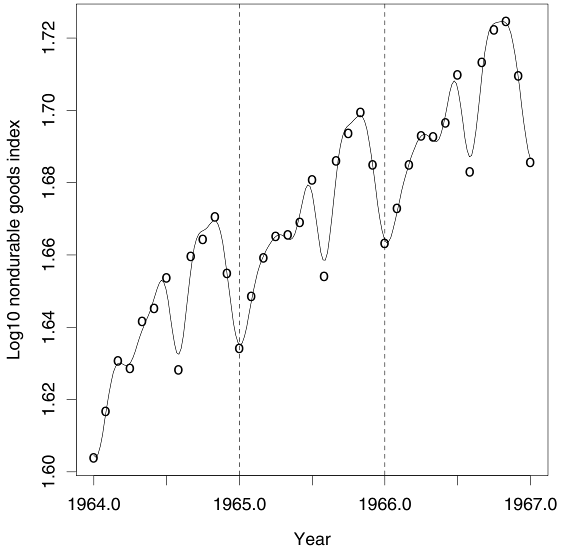

Figure 3.3. The log nondurable goods index for 1964 to 1967, a period of comparative stability. The solid line is a fit to the data using a polynomial smoothing spline. The circles indicate the value of the log index at the first of the month.

A closer look at a comparatively stable period, 1964 to 1967 shown in Figure 3.3, suggests that the index varies fairly smoothly and regularly within each year. The solid line is a smooth of these data using a method described in Section 3.6. We now see that the variation within this year is more complex than Figure 3.2 can possibly reveal. This curve oscillates three times during the year, with the size of the oscillation being smallest in spring, larger in the summer, and largest in the autumn. In fact each year shows smooth variation with a similar amount of detail, and we now consider how we can explore these within-year patterns.

3.3 Phase-plane plots

The rate of change of the index at any point is rather more interesting than its actual size. For example, the increase of 1.6% per year over the twentieth century gives us a reference value or benchmark for the average change of 2.0% from 1963 to 1972 or the smaller 0.8% increase following 1990. The crash of 1929, after all, mattered, not because the index was around 15 at that point, but because it was a change so abrupt that everybody noticed that something had happened.

If, then, it is change that matters, it follows that we need to study whatever alters velocity or the first derivative of the curve. The second derivative of the curve is its acceleration, and is instantaneous curvature in the index. When the index is curving upward, the velocity is increasing. Note the strong positive curvature in the index at the beginning of August, for example.

The smoothing method used to compute the curve in Figure 3.3 was designed to give a good impression of the velocity and acceleration of the log nondurable goods index. The capacity to generate high quality estimates of derivatives as well as curve values is a comparatively recent technical development in statistics and applied mathematics, and more details are provided in Section 3.6.

Now that we have derivatives at our disposal, we can learn new things by studying how derivatives relate to each other. Our tool is the *phase-plane plot*, a plot of acceleration against velocity. To see how this might be useful, consider the phase-plane plot of the function $\sin(2\pi t)$, shown in Figure 3.4. This simple function describes a basic *harmonic process*, such as the vertical position of the end of a suspended spring bouncing with a period of one time unit and starting at position zero at time $t = 0$.

The spring oscillates because energy is exchanged between two states: *potential* and *kinetic*. At times $1, 3, \ldots$ the spring is at one or the other end of its trajectory, and the restorative force due to its stretching has brought it to a standstill. At that point, its potential energy is maximized, and so is the force, which is acting either upward (positively) or downward. Since force is proportional to acceleration, the second derivative of the spring position, $-(2\pi)^2 \sin(2\pi t)$, is also at its highest absolute value, in this case about ± 40. On the other hand, when the spring is passing through the position 0, its velocity, $2\pi \cos(2\pi t)$, is at its greatest, about ± 8, but its acceleration is zero. Since kinetic energy is proportional to the square of velocity, this is the point of highest kinetic energy. The phase-plane plot shows this energy exchange nicely, with potential energy being maximized at the extremes of Y and kinetic energy at the extremes of X.

Now harmonic processes and energy exchange are found in many situations besides mechanics. In economics, potential energy corresponds to available capital, human resources, raw material, and other resources that are at hand to bring about some economic activity, in this case the manufacture of nondurable goods. Kinetic energy corresponds to the manufacturing process in full swing, when these resources are moving along the assembly line, and the goods are being shipped out the factory door.

The process moves from strong kinetic to strong potential energy when the rate of change in production goes to zero. We see this, for example, after a period of rapid increase in production when labor supply and raw material stocks become depleted, and consequently potential energy is actually in a negative state. Or it happens when management winds down produc-

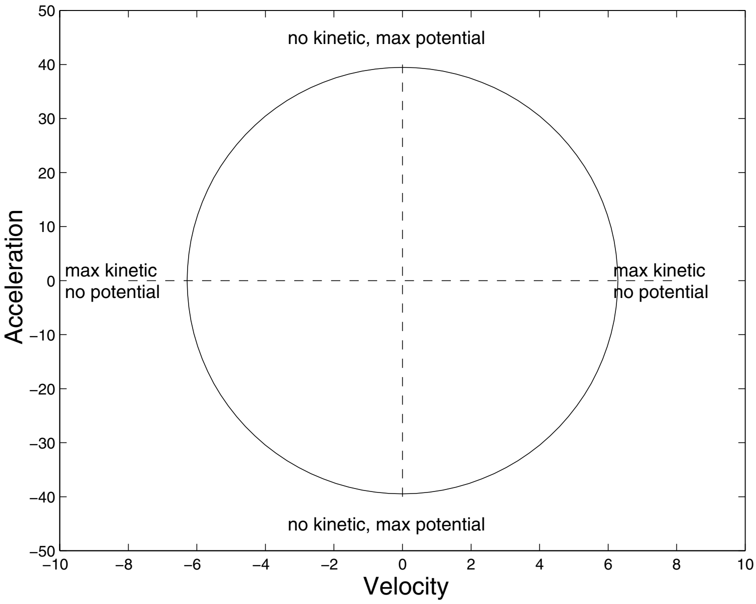

Figure 3.4. A phase-plane plot of the simple harmonic function $\sin(2\pi t)$. Kinetic energy is maximized when acceleration is 0, and potential energy is maximized when velocity is 0.

tion because targets have been achieved, so that personnel and material resources are piling up and waiting to be used anew.

After a period of intense production, or at certain periods of crisis that we examine shortly, we may see that both potential and kinetic energy are low. This corresponds to a period when the phase-plane curve is closer to zero than is otherwise the case.

To summarize, here's what we are looking for:

- a substantial cycle;
- the size of the radius: the larger it is, the more energy transfer there is in the event;
- the horizontal location of the center: if it is to the right, there is net positive velocity, and if to the left, there is net negative velocity;
- the vertical location of the center: if it is above zero, there is net velocity increase; if below zero, there is velocity decrease; and
- changes in the shapes of the cycles from year to year.

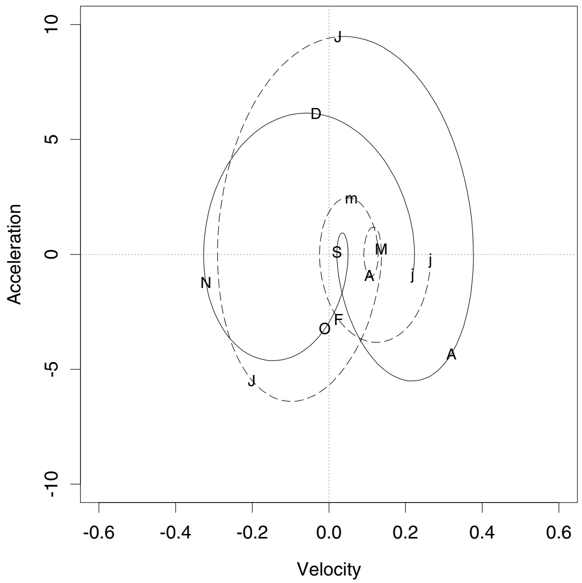

Figure 3.5. A phase-plane plot of the first derivative or velocity and the second derivative or acceleration of the smoothed log nondurable goods index for 1964. Letters indicate midmonths, with lowercase letters used for January and March. For clarity, the first half of the year is plotted as a dashed line, and the second half as a solid line.

3.4 The nondurable goods cycles

We use the phase-plane plot, therefore, to study the energy transfer within the economic system. We can examine the cycle within individual years, and also see more clearly how the structure of the transfer has changed throughout the twentieth century. Figure 3.5 phase-plane plots the year 1964, a year in a relatively stable period for the index. To read the plot, find the lower-case "j" in the middle right of the plot, and move around the diagram clockwise, noting the letters indicating the months as you go. You will see that there are two large cycles surrounding zero, plus some small cycles that are much closer to the origin.

The largest cycle begins in mid-May (M), with positive velocity but near zero acceleration. Production is increasing linearly or steadily at this point. The cycle moves clockwise through June (first J) and passes the horizontal zero acceleration line at the end of the month, when production is now decreasing linearly. By mid-July (second J) kinetic energy or velocity is near zero because vacation season is in full swing. But potential energy or acceleration is high, and production returns to the positive kinetic/zero potential phase in early August (A), and finally concludes with a cusp at summer's end (S). At this point the process looks like it has run out of both potential and kinetic energy.

The cusp, near where both derivatives are zero, corresponds to the start of school in September, and to the beginning of the next big production cycle passing through the autumn months of October through November. Again this large cycle terminates in a small cycle with little potential and kinetic energy. This takes up the months of February and March (F and m). The tiny subcycle during April and May seems to be due to the spring holidays, since the summer and fall cycles, as well as the cusp, don't change much over the next two years, but the spring cycle cusp moves around, reflecting the variability in the timings of Easter and Passover.

To summarize, the production year in the 1960s has two large cycles swinging widely around zero, each terminating in a small cusplike cycle. This suggests that each large cycle is like a balloon that runs out of air, the first at the beginning of school, and the second at the end of winter. At the end of each cycle, it may be that new resources must be marshaled before the next production cycle can begin.

With this basic pattern characterizing the phase-plane plot for a stable year, it can be revealing to examine years in which important events took place. Figure 3.6 shows what happened in 1929 to 1931. Year 1929 has the same features as we saw above for 1964, but we see a bulge to the left in the late autumn, when the stock market crashed. By November of that year production was in a state of freefall. We pick up the story in the middle cycle for 1930, and see that, after a small spring and larger summer cycle, the autumn cycle loses much of its potential energy, and this is even more evident in 1931. Probably this is attributable to the collapse of consumer demand in the holiday period as people restrict spending to the essentials.

Figure 3.7 pictures the events leading to World War II. The first part of 1937 shows only small amounts of energy as the Depression continues. But the cycle is dramatically altered in the fall by the sudden decrease in the money supply imposed by the Treasury Board when it feared that the economy might be overheated and headed for another crash. You can see in Figure 3.2 that this precipitous event is comparable in size to the stock market crash of 1929, but even more sudden. The spring and fall cycles were all but wiped out in 1938.

The bottom plot in Figure 3.7 shows the reduced seasonal variability during the war years, and this is also clearly visible in Figure 3.2. In times of war people don't take holidays, make do with what they have, and spend less at Christmas. Moreover, war production did not exhibit much seasonal variation since the demand for nondurable goods, like the war itself, was steady through the year.

Another three years in which important changes occur are 1974 to 1976, plotted in Figure 3.8. The Vietnam War was concluded in this period, and

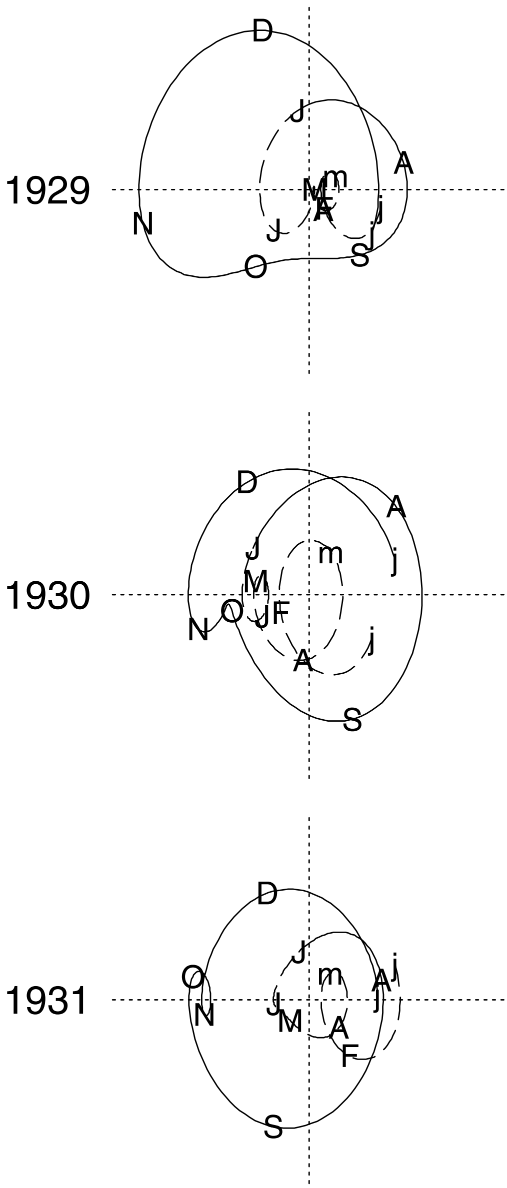

Figure 3.6. Phase-plane plots for the years 1929 to 1931, during the onset of the Great Depression. The horizontal and vertical scale is the same as in Figure 3.5.

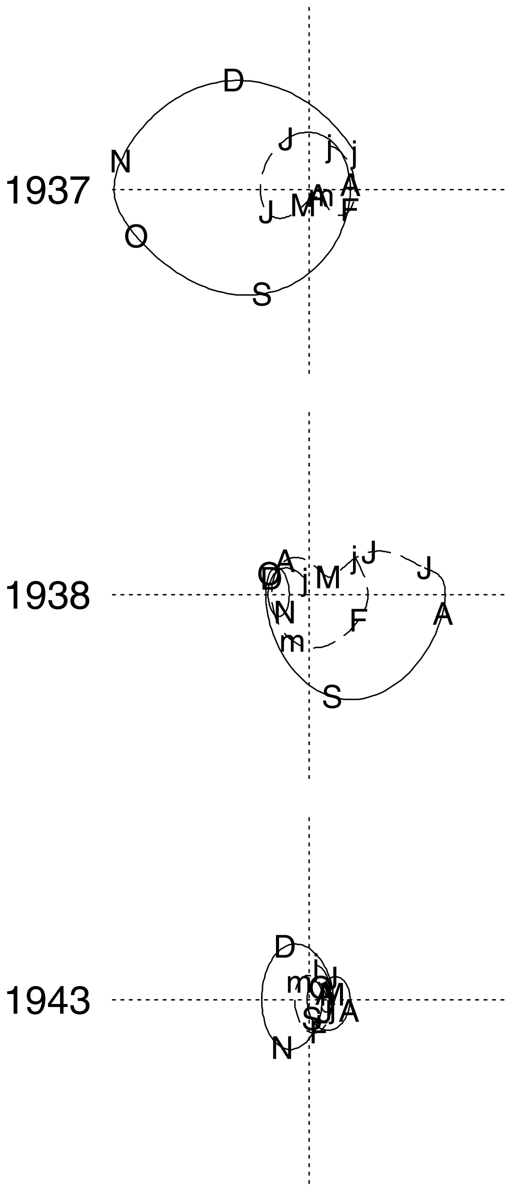

Figure 3.7. Phase-plane plots for two years preceding the Second World War and a typical war year.

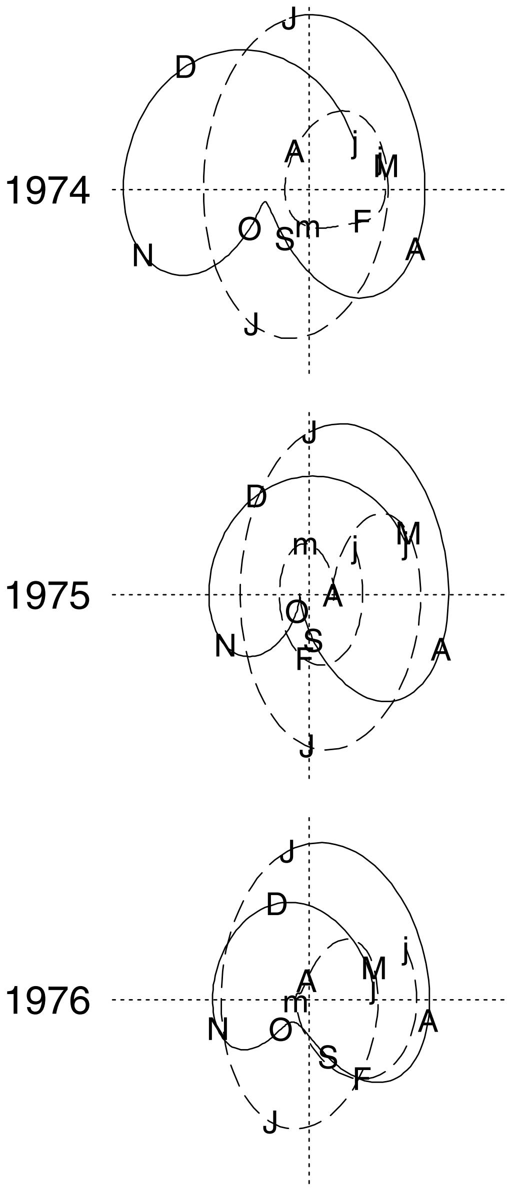

Figure 3.8. Phase-plane plots for 1974 to 1976, when the production cycles are changing rapidly.

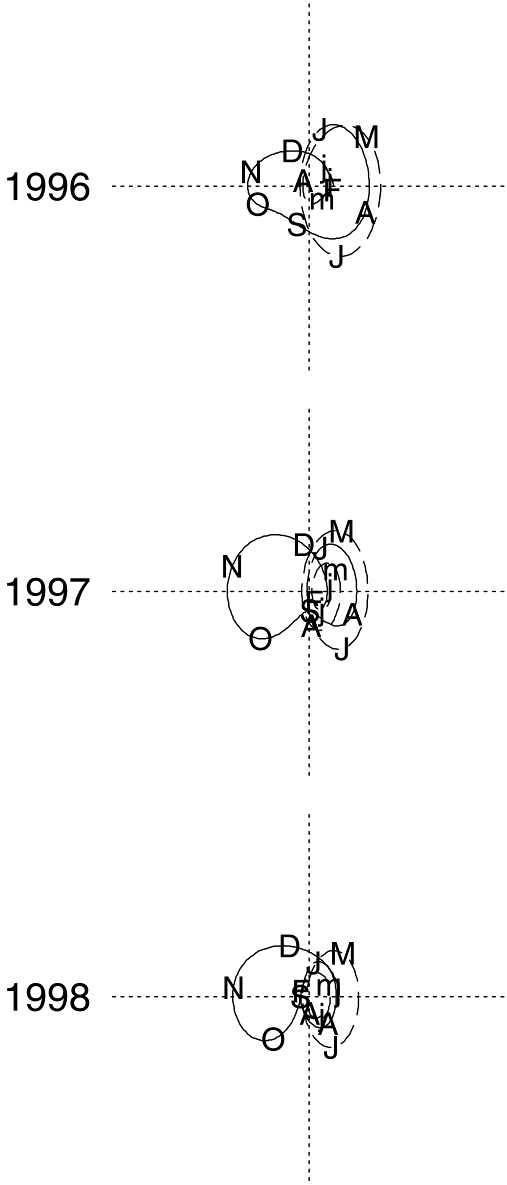

Figure 3.9. Phase-plane plots for 1996 to 1998, showing the greatly reduced variability of current production cycles.

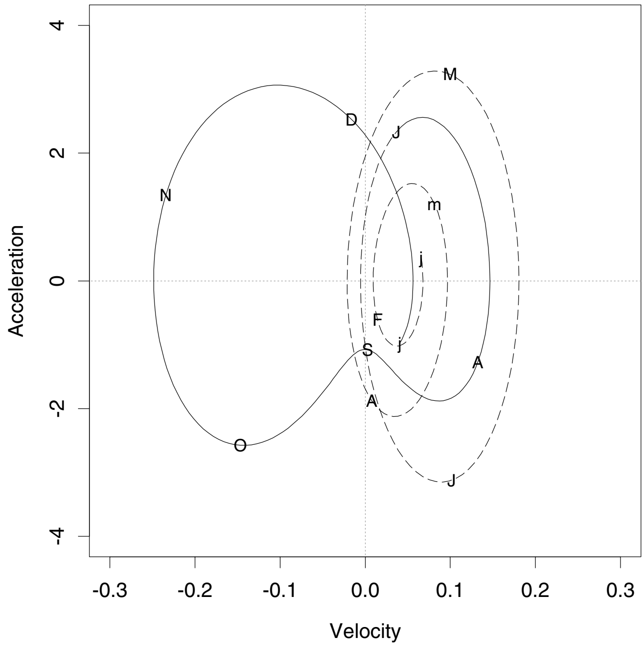

Figure 3.10. The phase-plane plot for 1997 on a larger scale, showing the structural changes in current production cycles.

the OPEC oil crisis also contributed to a change in economic patterns. One consequence was the decrease in the size of the fall loop. What we cannot see in this small time window, though, is that fundamental changes initiated in the mid-1970s persist to the present day.

What is happening now? Figure 3.9 shows that the production cycles are now much smaller than they once were. We still see fairly large seasonal oscillations, but they are now much smoother, and hence show less variation in velocity and acceleration. Also, if we look at Figure 3.10 showing the 1997 cycles on a larger scale, we see that there are now four cycles rather than three, and that the final winter cycle has a strongly negative net velocity. Are this loss of dynamism and these structural changes due to the fact that production is no longer so dependent on manpower? Or, perhaps, that it is more tightly controlled by information technology? On the other hand, it may be simply that far more nondurable goods are now manufactured outside the United States.

A further clue to recent changes is that in the early 1990s, personal computers and other electronic goods were classified as durable. Consequently, one sees in the comparable index for durable goods a strong increase in its typical slope at that point. Although it is true that electronic goods usually last more than two years, the pace of technological development in this sector has meant that, effectively, consumers have tended to discard these items because they are obsolete. This loss of electronic goods in the nondurable goods index has surely diminished its energy.

3.5 What have we seen?

Phase-plane plotting is revealing because it focuses our attention on the *dynamics* of the seasonal component of variation in the goods index. We plot velocity on the horizontal axis, representing the rate of change of the process; and plot acceleration on the vertical axis, indicating the input or withdrawal of whatever resources or forces produce this change. Because seasonal components tend to exhibit oscillatory or harmonic behavior, we can interpret what we see as a transition between two types of energy: kinetic associated with velocity, and potential associated with acceleration. Harmonic behavior, in which the system moves between these two states, shows up as a loop surrounding the origin. The bigger the radius of the loop, the more energy the system has, and the smaller or closer it is to zero, the less the energy.

We saw that the typical year shows three such loops, associated with the spring, summer, and fall. The summer loop typically has the largest associated energy. But the fall loop seems to be most affected by shocks such as the stock market crash of 1929, the shutting down of the money supply in 1937, and the end of the Vietnam War in 1974. This is probably due to the fact that the fall production loop is associated with buying for the Christmas holidays, and therefore is something consumers can turn on and off according to whether times are good or tough, respectively.

We also saw the seasonal dynamics reflecting longer-term changes. There is much less energy in the system now than in the 1960s, as reflected in the smallness of loops in recent times.

The dynamics of a process typically show more variation than the statics or position of the process, and we could see things happening in the phase-plane plots that would be hard to spot in the plot of the process itself, such as in Figure 3.2.

This focus on dynamics leads to the question of whether we can model these dynamic features directly, rather than putting all of our statistical energy into reproducing the curve itself. This leads us naturally to the idea of using a *differential equation* to describe the process, a type of modeling that will allow us to model the dynamic behavior seen in the phase-plane plot as well as the curve itself. We use differential equations in models in Chapters 11 and 12.

3.6 Smoothing data for phase-plane plots

3.6.1 Fourth derivative roughness penalties

We can imagine that the economic forces generating the log index values are reasonably smooth. In practice, this means that a curve giving a satisfactory picture of these processes has a certain number of derivatives. For phase-plane plotting, in particular, we need to use two derivatives in addition to the curve values themselves. We estimate these derivatives by smoothing the data, using a method that will give useful estimates of velocity and acceleration as well as of the underlying curve itself.

Therefore we choose to fit a smooth curve $h(t)$ to log index values $y_i, i = 1, \ldots, 973$, by using the following criterion

$$\text{PENSSE}_\lambda(h) = \sum_{i=1}^{973}[y_i - h(t_i)]^2 + \lambda \int_{1919}^{2000} [h^{(iv)}(t)]^2 \, dt. \tag{3.1}$$

The criterion has two terms. The first assesses the fidelity of the curve to the observed data in the sense of the sum of squared errors.

But fitting the data is not our only concern, and the second term, the penalty term, measures the extent to which the fitting function $h(t)$ is smooth. The notation $h^{(iv)}(t)$ in (3.1) means the fourth derivative of h evaluated at time t. The penalty term captures the overall size of this fourth derivative by integrating its square over the interval of interest. Why the fourth derivative? Because it is sensitive to the curvature of the second derivative, or acceleration. Recall that curvature is indicated by the second derivative, so the curvature of the acceleration function is its second derivative, or the fourth derivative of the actual curve $h(t)$.

We cannot have smoothness and a nearly perfect fit to the data at the same time, especially when we have this many observations. The smoothing parameter λ controls the relative emphasis on fitting the data and smoothness. As λ increases, smoothness is accentuated more and more, until finally the integrated square of $h^{(iv)}(t)$ will be driven to zero. Only polynomials of degree three or fewer have zero fourth derivatives, and clearly a function this simple would not fit these data at all well. On the other hand, as λ goes to zero, smoothness matters less and less, and hence fitting the data more and more. Finally we will arrive at a function that fits the data exactly. Unfortunately, it will not be at all smooth, and its second derivative will be too wildly varying to be at all useful. The challenge, then, is to find a value for λ that works for us.

3.6.2 Choosing the smoothing parameter

We discussed this problem in Chapter 2. There a data-driven technique, cross-validation, was described that could be used to guide this choice. However, we were not shy to say that our final choice depended on inspection of the results, and that we used a value rather different than that suggested by this purely data-driven method. We now continue to discuss concerns that might govern the amount of smoothness in a curve that smooths data.

Our goal in this chapter was to use phase-plane plots to reveal something about seasonal trend, and how it evolves over time. Of course the technique is not going to be helpful if the curve misses obviously important features in the data. Our first move, therefore, was to carefully study how well the curve tracks the data by using close-up plots such as Figure 3.3. We actually plotted the data and the fit separately for each of the 51 years of interest, and noted where the curve seemed to miss the data repeatedly. We observed, for example, that the curve was too smooth if it underestimated peak values such as that of June year after year, or if it consistently overestimated low values such as July. We also learned a lot by looking at the residuals from the fit, computed by subtracting the fitted from the actual value. If there was some trend running over several months, this was a sign that we had oversmoothed the data. At this stage, one may say, it is rather easier to detect oversmoothing than undersmoothing. These investigations gave us a fairly firm idea of an upper limit on λ, but less intuition about a lower limit.

Next we looked at what we wanted to work with, namely the phase-plane plot. Here smoothness matters a great deal. We wanted to see important and consistent patterns, and too much wiggliness in the plot makes this difficult. In general, high derivatives are rather more unstable than lower ones, so at this point it was primarily smoothness in acceleration that mattered; if acceleration was smooth, so was velocity. So we started with a smallish value of λ, and moved it upward bit by bit until the phase-plane plot seemed stable from year to year over periods when it should be, such as the 1960s, and, of course, to the point where we could make sense out of the structure of the plot. This process gave us a desirable lower limit on λ.

We have to admit that this lower limit is often larger than the upper limit identified by looking at the data fit. However, at this point some fit just has to be sacrificed in order to see what we are looking for in the data—hence the systematic misfitting of the July log index in Figure 3.3. Perhaps we will return to the data someday to have a look at what we missed this time, but for the moment we are satisfied with what we learned. Our final choice for λ was $10^{-9.5}$.

In summary, our philosophy, and, we believe, the perspective of most practitioners of smoothing, is that choosing a level of smoothing is a matter of balancing off fitting the data against getting a stable and interpretable estimate of what interests us. We see the choice of λ as very much driven by the needs of the investigator, and are content to see other analyses of the same data employ a different value.