3 ARIMA Impact Assessment

The time series quasi-experiment was proposed originally by Campbell (1963; Campbell and Stanley, 1966) as a means of assessing the impact of a discrete social intervention (or an "event" as we shall soon call it) on behavioral processes. The reader must not assume that a time series quasi-experiment is always the best method of impact assessment for it requires a rather simple theory of impact. There appear nevertheless to be many situations in which simple theories of impact are justified and, in these situations, the time series quasi-experiment may be the most useful of all designs.

Time series quasi-experiments have been used to measure the impacts of new traffic laws (Campbell and Ross, 1968; Glass, 1968; Ross et al., 1970); the impact of decriminalization (Aaronson et al., 1978); the impact of gun control laws (Zimring, 1975); and the impact of air pollution control laws (Box and Tiao, 1975). The widest use of this design has clearly been in the area of legal impact assessment. Time series quasi-experiments have also been used by political scientists to measure the impacts of political realignments (Caporaso and Pelowski, 1971; Lewis-Beck, 1979; Smoker, 1969), however, and by experimental psychologists to measure the impacts of treatments on behavior (Gottman and McFall, 1972; Hall et al., 1971; Tyler and Brown, 1968). This list is representative (but by no means exhaustive) of the situations in which time series quasi-experiments have been used to assess social impacts.

Our major concern in this chapter is with the analysis of time series quasi-experiments: impact assessment as we call it. Time series analysis cannot be divorced from the design of the time series quasi-experiment, however. According to Cook and Campbell (1979), the design of any quasi-experiment must recognize the threats to four types of validity: internal, external, statistical conclusion, and construct validities. We know of no statistical methods for correcting or controlling flaws in design. Our development of statistical models thus assumes an adequately designed quasi-experiment. Design is the sine qua non. The reader who is unfamiliar with quasi-experimental design is directed to Cook and Campbell for an authoritative treatment.

We will use the term *impact assessment* to refer to the statistical analysis of an adequately designed time series quasi-experiment. More generally, we define this term as "a test of the null hypothesis that a postulated event caused a change in a social process measured as a time series." Acknowledging the faults and limitations of this definition, we must comment on its two key elements.

First, impact assessment is concerned with the effects of a "postulated event." An event for our purposes is a qualitative change in state or, in common terms, "something that happens." Events can be represented as binary variables which indicate the *absence* of the state prior to the event and the *presence* of the state during and (possibly) after the event. In the parlance of experimental psychology, for example, introduction of a treatment is the event associated with a change in state from "no treatment" to "treatment." In legal studies, enactment of a new law is the event associated with a change in state from "no regulation" to "regulation."

Qualitative changes in states (events) are often indistinguishable from quantitative changes in levels (processes). In studying national arms expenditures over time, for example, some social scientists prefer to think of "war" as an *event* which affects expenditures. Other social scientists prefer to think of "the propensity to war" as a continuous *process* which affects expenditures. We will develop multivariate ARIMA models for the case in which the causal agent is a continuous process (as measured by an independent variable time series) in Chapter 5. For now, however, it is important to remember that an impact assessment analysis is concerned with the effect of an *event* on some social or behavioral process.

Because the change agent is an event, it is represented in the impact assessment model as a "dummy" variable or step function such that

$$\begin{aligned} I_t &= 0 \text{ prior to the event} \\ &= 1 \text{ thereafter.} \end{aligned}$$

If the change process is not an event in the technical sense, however, the

impact assessment may lead to invalid conclusions. As a general rule, the independent variable of an impact assessment model should give as accurate a representation of the change agent as possible. In an analysis of a Washington, D.C., gun control law, for example, Zimring (1975) had information on the actual level of enforcement. Using this information, Zimring defined the event in terms of I_t as

$$\begin{aligned} I_t &= 0 \text{ prior to enactment} \\ &= 1/6 \text{ in the first month after enactment} \\ &= 2/6 \text{ in the second month} \\ &\;\;\vdots \\ &= 6/6 = 1 \text{ in the sixth and subsequent months.} \end{aligned}$$

With this definition of I_t, the event corresponding to a change in state from "no regulation" to "regulation" is distributed across a six-month period. The fundamental principle illustrated here is that an impact assessment requires a theory of change. If the change agent is an event, then it can and must be represented by a simple step function. If the change agent is *not* an event in the strictest sense, however, the analysis may lead to invalid conclusions. A more valid analysis can be ensured by modifying the step function, as Zimring did, to accommodate known properties of the change agent.

A second element of the "impact assessment" definition is the a priori specification of the onset of an event. A null hypothesis that an event "caused" a change in some behavior can be tested only because the time of the event is known a priori. It would indeed be possible to search the length of a time series for statistically significant changes but it would be logically impossible to then associate each change with the infinite number of events which might be the causes. An impact analysis based on a blind search (see, e.g., Deutsch, 1978; also, Section 4.3 below) might generously be called "exploratory analysis." Its results are quite uninterpretable. An impact assessment based on an event whose onset is specified a priori, in contrast, is a "confirmatory analysis." It is used only to test theoretically generated hypotheses according to a rigorous set of validity criteria.

These two elements of the definition are so important that we reiterate them. First, impact assessment is concerned only with *events* and, second, impact assessment requires that the *onset* of an event be specified a priori. Lacking these two elements, the results of an impact analysis will be uninterpretable.

Impact assessment (or the time series analysis of impacts) begins with an

ARIMA model for the time series. Since this ARIMA model describes the stochastic behavior of the time series process, we refer to it as the "noise component" of the model. An intervention component is then added to the model. The full impact assessment model may be written as

$$Y_t = f(I_t) + N_t,$$

where N_t denotes the noise component, an ARIMA model, and where $f(I_t)$ denotes a "function of the variable I_t," the intervention component. The intervention component itself describes the deterministic relationship between an event (as represented by the variable I_t) and the time series. The noise component describes the stochastic behavior of the time series around the $Y_t = f(I_t)$ relationship.

The general principles of ARIMA modeling which we developed in Chapter 2 apply as well to impact assessment modeling. Analysis begins with construction of an ARIMA model for the Y_t time series. In some cases, the real impact in the time series may be so large that it overwhelms and distorts the ACF and PACF; this phenomenon is similar to distortions associated with outliers. To avoid problems in identification, the analyst may have to estimate ACFs and PACFs from the preintervention series only.

After an adequate ARIMA model has been identified, its parameters satisfactorily estimated, and its residuals diagnosed, an intervention component is added. The intervention component will ideally be selected on the basis of a theoretically generated null hypothesis. The parameters of the full impact assessment model (both noise and intervention components) are then estimated. If a parameter estimate is not statistically significant or is otherwise unacceptable (if the estimate of a noise parameter lies outside the bounds of stationarity-invertibility, for example), the tentative model must be respecified and its parameters reestimated.

Once a tentative model has been specified and significant, acceptable parameter estimates have been obtained, the impact assessment model must be diagnosed. As in the case of univariate ARIMA models, residuals must not be different than white noise. Although the noise component alone may have white noise residuals, it sometimes happens that the full impact assessment model (noise *and* intervention components) does not. When this happens, a new tentative model must be specified, its parameters estimated, and its residuals diagnosed. The model-building procedure continues iteratively until a parsimonious but statistically adequate impact assessment model is generated.

Impact parameters may then be tested for statistical significance and, more generally, *the model may be interpreted*. We will illustrate the general

model-building strategy for impact assessments with several example analyses. These will not be as definitively instructive as the example analyses of the previous chapter, however. While ARIMA models per se (the noise component of the impact assessment model) are atheoretical and uninterpretable, an impact assessment model (noise *and* intervention components) is built for no reason other than interpretation. The analyst must draw conclusions from the impact assessment analysis and, in every case, these conclusions must be reconciled with the prevailing theory of a substantive area. An impact assessment model may then be the "best" possible model in a statistical sense but not in the substantive sense. Interpretability is everything, and for this reason, impact assessment modeling cannot be reduced to a set of objective, mechanical steps.

3.1 The Zero-Order Transfer Function

Denoting the full impact assessment model as

$$Y_t = f(I_t) + N_t,$$

we are now concerned with the "function of I_t," that is, with the intervention component of the full model. Some writers refer to the intervention component as a "transfer function," a term derived from engineering contexts. We will use both terms synonymously here, although our preference is for the more straightforward "intervention component."

The simplest possible intervention component is the zero-order transfer function

$$f(I_t) = \omega_0 I_t.$$

This is a *zero*-order transfer function because the highest power of B in the function is zero. Where the variable I_t is defined as a step function such that

$$\begin{aligned} I_t &= 0 \text{ prior to the event} \\ &= 1 \text{ thereafter,} \end{aligned}$$

the impact assessment model is:

$$Y_t = \omega_0 I_t + N_t.$$

Now because the impact assessment model is linear in its components, the noise component, N_t, may be subtracted from the time series:

$$\begin{aligned} Y_t^* &= Y_t - N_t \\ &= \omega_0 I_t. \end{aligned}$$

So long as the N_t component is statistically adequate, that is, so long as it has been built along the lines described in the previous chapter, subtracting it from the Y_t time series results in a deterministic intervention component. Working with the Y_t^* series (instead of the Y_t series), the deterministic effects of the transfer function may be examined.

Prior to the event, when $I_t = 0$, the level of the Y_t^* series is:

$$\begin{aligned} Y_t^* &= \omega_0(0) \\ &= 0. \end{aligned}$$

But with the onset of the event, when $I_t = 1$, the level of the Y_t^* series is:

$$\begin{aligned} Y_t^* &= \omega_0(1) \\ &= \omega_0. \end{aligned}$$

The zero-order transfer function thus determines an *abrupt, permanent* shift in process level from pre- to postintervention, a pattern of impact such as

o o o o o *After*
$\uparrow$
ω_0
$\downarrow$
Before o o o o o

At the onset of the event, the level of the process increases by the quantity ω_0 (or decreases if ω_0 is negative).

A few comments on the concept of "level" may be helpful here. An impact assessment model describes a change in level and/or (sometimes) trend for the generating process of the time series. Some writers use the term *equilibrium* rather than "level," but whichever term is used, it is important to remember that a statistical concept (not a substantive concept) is implied. For a stationary time series process, the parameter ω_0 is an estimate of the difference between the pre- and postintervention process levels.

For a nonstationary series, an analogous interpretation is possible. Nonstationary time series generated by ARIMA(p,d,q) processes can be represented by stationary ARIMA(p,0,q) models after an appropriate differencing. As noted in Chapter 2, the inverse relationship between differencing and summation operators (or filters; see Figure 2.13) allows for a mapping between the stationary model and the nonstationary process. This is also true of impact assessment models. An *abrupt, permanent* pattern of impact, for example, as determined by the zero-order transfer function, in a trending

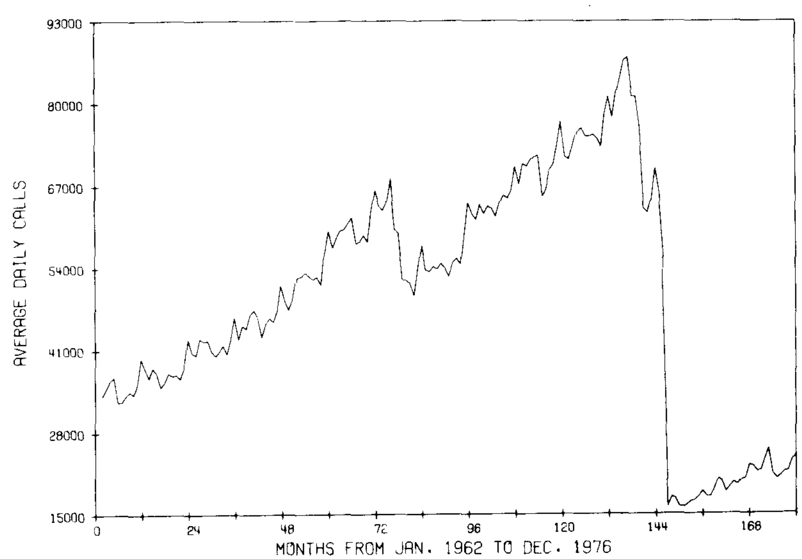

FIGURE 3.1(a) Directory Assistance, Monthly Average Calls per Day

series may appear as

After
o
o
o
o
o
$\uparrow$
ω_0
$\downarrow$
o
o
o
Before o

The interpretation of the parameter ω_0 is more or less the same, then, whether the N_t component is a stationary ARIMA(p,0,q) model or a nonstationary ARIMA(p,d,q) model.

Using the simple zero-order transfer function, we will now demonstrate the impact assessment model-building strategy. Figure 3.1(a) shows a monthly time series of calls to Directory Assistance in Cincinnati, Ohio, as reported by McSweeny (1978). The first observation of this series is Janu-

(text continued on p. 152)

```
SERIES.. TEL146    (NOBS= 146)  DIRECTORY ASST - PRE INTERVENTION
NO. OF VALID OBSERVATIONS  =    146.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25, 146) =  1727.3       SIG =     0.000

 LAG   CORR   SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1    .971  .083                      (    IXXX)XXXXXXXXXXXXXXXXXXXX
  2    .939  .141                    (      IXXXXXX)XXXXXXXXXXXXXXXX
  3    .917  .178                  (        IXXXXXXXX)XXXXXXXXXXXXXX
  4    .894  .208                 (         IXXXXXXXXX)XXXXXXXXXXXX
  5    .869  .233                (          IXXXXXXXXXXX)XXXXXXXXXX
  6    .842  .254               (           IXXXXXXXXXXXX)XXXXXXXX
  7    .814  .273              (            IXXXXXXXXXXXXX)XXXXXX
  8    .790  .289              (            IXXXXXXXXXXXXX)XXXXXX
  9    .761  .303             (             IXXXXXXXXXXXXXX)XXXX
 10    .732  .316            (              IXXXXXXXXXXXXXXX)XX
 11    .713  .328            (              IXXXXXXXXXXXXXXX)XX
 12    .697  .338           (               IXXXXXXXXXXXXXXXX*
 13    .665  .348           (               IXXXXXXXXXXXXXXXX*
 14    .632  .356          (                IXXXXXXXXXXXXXXXX )
 15    .606  .364          (                IXXXXXXXXXXXXXXX  )
 16    .579  .371         (                 IXXXXXXXXXXXXXX    )
 17    .551  .377         (                 IXXXXXXXXXXXXXX    )
 18    .526  .382         (                 IXXXXXXXXXXXXX     )
 19    .502  .387         (                 IXXXXXXXXXXXXX     )
 20    .481  .392        (                  IXXXXXXXXXXXX       )
 21    .459  .396        (                  IXXXXXXXXXXX        )
 22    .435  .399        (                  IXXXXXXXXXXX        )
 23    .421  .403        (                  IXXXXXXXXXXX        )
 24    .412  .406        (                  IXXXXXXXXXX         )
 25    .387  .408        (                  IXXXXXXXXXX         )
                       -2SE                                   +2SE
```

```
PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

 LAG   CORR   SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1    .971  .083                      (    IXXX)XXXXXXXXXXXXXXXXXXXX
  2   -.047  .083                      (  XI   )
  3    .145  .083                      (    IXXX*
  4   -.052  .083                      (  XI   )
  5    .009  .083                      (    I  )
  6   -.082  .083                      ( XXI   )
  7   -.016  .083                      (    I  )
  8    .029  .083                      (    IX )
  9   -.092  .083                      ( XXI   )
 10   -.015  .083                      (    I  )
 11    .156  .083                      (    IXXX*
 12    .039  .083                      (    IX )
 13   -.277  .083                   XXX(XXXI   )
 14    .005  .083                      (    I  )
 15    .055  .083                      (    IX )
 16   -.079  .083                      ( XXI   )
 17   -.012  .083                      (    I  )
 18    .089  .083                      (    IXX )
 19    .001  .083                      (    I  )
 20   -.024  .083                      (   XI  )
 21    .040  .083                      (    IX )
 22   -.017  .083                      (    I  )
 23    .080  .083                      (    IXX )
 24    .010  .083                      (    I  )
 25   -.172  .083                      *XXXI   )
                                     -2SE    +2SE
```

FIGURE 3.1(b) ACF and PACF for the Raw Preintervention Series


```
SERIES.. TEL146    (NOBS= 146)  DIRECTORY ASST - PRE INTERVENTION
DIFFERENCED  1 TIME(S) OF ORDER  1.
NO. OF VALID OBSERVATIONS  =    145.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25, 145) =  63.683       SIG =     .000

 LAG   CORR   SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1    .032  .083                      (    IX  )
  2   -.195  .083                    X(XXXI    )
  3    .053  .086                      (    IX  )
  4    .012  .086                      (    I   )
  5    .006  .086                      (    I   )
  6    .022  .086                      (    IX  )
  7   -.069  .087                      (  XXI   )
  8    .104  .087                      (    IXXX)
  9    .021  .088                      (    IX  )
 10   -.235  .088                    XX(XXXI   )
 11   -.017  .092                     (     I    )
 12    .360  .092                     (     IXXXX)XXXX
 13   -.014  .101                     (     I    )
 14   -.150  .101                     (XXXXI     )
 15    .026  .103                     (     IX   )
 16   -.045  .103                     (    XI    )
 17   -.058  .103                     (    XI    )
 18   -.010  .103                     (     I    )
 19   -.072  .103                     (   XXI    )
 20    .023  .104                     (     IX   )
 21    .045  .104                     (     IX   )
 22   -.239  .104                    X(XXXXI     )
 23   -.017  .107                     (     I    )
 24    .319  .107                     (     IXXXX)XXX
 25   -.002  .114                    (      I     )
                                   -2SE         +2SE
```

```
PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

 LAG   CORR   SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1    .032  .083                      (    IX  )
  2   -.196  .083                    X(XXXI    )
  3    .070  .083                      (    IXX )
  4   -.033  .083                      (   XI   )
  5    .033  .083                      (    IX  )
  6    .012  .083                      (    I   )
  7   -.065  .083                      (  XXI   )
  8    .123  .083                      (    IXXX)
  9   -.023  .083                      (   XI   )
 10   -.193  .083                    X(XXXI    )
 11   -.007  .083                      (    I   )
 12    .311  .083                      (    IXXX)XXXX
 13   -.038  .083                      (   XI   )
 14   -.069  .083                      (  XXI   )
 15    .020  .083                      (    I   )
 16   -.073  .083                      (  XXI   )
 17   -.089  .083                      (  XXI   )
 18   -.005  .083                      (    I   )
 19   -.033  .083                      (   XI   )
 20   -.082  .083                      (  XXI   )
 21    .026  .083                      (    IX  )
 22   -.116  .083                      (XXXI    )
 23    .003  .083                      (    I   )
 24    .179  .083                      (    IXXX*
 25    .016  .083                      (    I   )
                                     -2SE    +2SE
```

FIGURE 3.1(c) ACF and PACF for the Regularly Differenced Preintervention Series

```
SERIES.. TEL146    (NOBS= 146)   DIRECTORY ASST - PRE INTERVENTION
DIFFERENCED  1 TIME(S) OF ORDER  1.
DIFFERENCED  1 TIME(S) OF ORDER 12.
NO. OF VALID OBSERVATIONS  =    133.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25, 133) =  25.266       SIG =      .448

LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1  -.036  .087                       (  XI   )
  2   .033  .087                       (   IX  )
  3   .083  .087                       (   IXX )
  4   .009  .088                       (   I   )
  5   .087  .088                       (   IXX )
  6   .073  .088                       (   IXX )
  7   .021  .089                       (   IX  )
  8   .110  .089                       (   IXXX)
  9  -.034  .090                       (  XI   )
 10  -.023  .090                       (  XI   )
 11   .070  .090                       (   IXX )
 12  -.294  .090                     XX(XXXXI   )
 13   .021  .097                      (    IX   )
 14  -.023  .097                      (   XI    )
 15  -.035  .097                      (   XI    )
 16  -.069  .097                      (  XXI    )
 17   .102  .098                      (    IXXX )
 18  -.121  .098                      ( XXXI    )
 19  -.061  .100                      (  XXI    )
 20  -.038  .100                      (   XI    )
 21   .116  .100                      (    IXXX )
 22  -.095  .101                      (  XXI    )
 23  -.038  .102                      (   XI    )
 24  -.015  .102                      (    I    )
 25   .032  .102                      (    IX   )
                                     -2SE      +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1  -.036  .087                       (  XI   )
  2   .031  .087                       (   IX  )
  3   .086  .087                       (   IXX )
  4   .014  .087                       (   I   )
  5   .083  .087                       (   IXX )
  6   .073  .087                       (   IXX )
  7   .021  .087                       (   IX  )
  8   .096  .087                       (   IXX )
  9  -.040  .087                       (  XI   )
 10  -.045  .087                       (  XI   )
 11   .042  .087                       (   IX  )
 12  -.305  .087                   XXXX(XXXI   )
 13  -.022  .087                       (  XI   )
 14  -.032  .087                       (  XI   )
 15   .004  .087                       (   I   )
 16  -.075  .087                       ( XXI   )
 17   .173  .087                       (   IXXX*
 18  -.069  .087                       ( XXI   )
 19  -.049  .087                       (  XI   )
 20   .025  .087                       (   IX  )
 21   .150  .087                       (   IXXX*
 22  -.133  .087                       (XXXI   )
 23   .002  .087                       (   I   )
 24  -.106  .087                       (XXXI   )
 25   .039  .087                       (   IX  )
                                      -2SE    +2SE
```

FIGURE 3.1(d) ACF and PACF for the Regularly *and* Seasonally Differenced Preintervention Series


```
SERIES.. RESIDUAL (NOBS= 133)  DIRECTORY ASST - PRE IMPACT RESIDUALS
NO. OF VALID OBSERVATIONS  =   133.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 24, 133) =  12.522       SIG =      .973

LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1  -.026  .087                       (  XI   )
  2   .009  .087                       (   I   )
  3   .089  .087                       (   IXX )
  4  -.032  .087                       (  XI   )
  5   .165  .088                       (   IXXX*
  6   .029  .090                       (   IX  )
  7  -.041  .090                       (  XI   )
  8   .111  .090                      (    IXXX )
  9   .023  .091                      (    IX   )
 10  -.061  .091                      (  XXI    )
 11  -.033  .091                      (   XI    )
 12   .012  .092                      (    I    )
 13  -.051  .092                      (   XI    )
 14  -.001  .092                      (    I    )
 15  -.053  .092                      (   XI    )
 16  -.088  .092                      (  XXI    )
 17   .055  .093                      (    IX   )
 18  -.060  .093                      (   XI    )
 19  -.019  .093                      (    I    )
 20  -.014  .093                      (    I    )
 21   .087  .093                      (    IXX  )
 22  -.061  .094                      (  XXI    )
 23  -.064  .094                      (  XXI    )
 24  -.018  .094                      (    I    )
 25  -.029  .094                      (   XI    )
                                     -2SE      +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
  1  -.026  .087                       (  XI   )
  2   .008  .087                       (   I   )
  3   .089  .087                       (   IXX )
  4  -.028  .087                       (  XI   )
  5   .164  .087                       (   IXXX*
  6   .029  .087                       (   IX  )
  7  -.037  .087                       (  XI   )
  8   .083  .087                       (   IXX )
  9   .033  .087                       (   IX  )
 10  -.083  .087                       ( XXI   )
 11  -.066  .087                       ( XXI   )
 12   .024  .087                       (   IX  )
 13  -.072  .087                       ( XXI   )
 14  -.018  .087                       (   I   )
 15  -.030  .087                       (  XI   )
 16  -.072  .087                       ( XXI   )
 17   .038  .087                       (   IX  )
 18  -.023  .087                       (  XI   )
 19   .009  .087                       (   I   )
 20  -.020  .087                       (  XI   )
 21   .139  .087                       (   IXXX)
 22  -.073  .087                       ( XXI   )
 23  -.062  .087                       ( XXI   )
 24  -.015  .087                       (   I   )
 25  -.023  .087                       (  XI   )
                                      -2SE    +2SE
```

FIGURE 3.1(e) *Diagnosis:* ACF and PACF for the Model Residuals (Preintervention Series Only)

```
SERIES.. RESIDUAL (NOBS= 167)  ESTIMATED MODEL RESIDUALS
NO. OF VALID OBSERVATIONS  =   167.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 24, 167) =  17.218      SIG =      .839

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0   .2   .4   .6   .8  +1
                   +----+----+----+----+----+----+----+----+----+----+
   1  -.030  .080                         (  XI   )
   2  -.030  .060                         (  XI   )
   3   .080  .080                         (    IXX )
   4  -.030  .080                         (  XI   )
   5   .110  .080                         (    IXXX)
   6   .040  .080                         (    IX  )
   7  -.060  .080                         ( XXI   )
   8   .070  .080                         (    IXX )
   9   .050  .080                         (    IX  )
  10  -.060  .080                         ( XXI   )
  11  -.090  .080                         ( XXI   )
  12   .040  .080                         (    IX  )
  13  -.080  .080                         ( XXI   )
  14   .010  .080                         (    I   )
  15   .020  .080                         (    IX  )
  16  -.100  .080                         (XXXI   )
  17  -.020  .060                         (  XI   )
  18  -.040  .080                         (  XI   )
  19  -.030  .080                         (  XI   )
  20  -.040  .080                         (  XI   )
  21   .090  .080                         (    IXX )
  22  -.090  .080                         ( XXI   )
  23  -.120  .080                         (XXXI   )
  24   .020  .060                         (    IX  )
  25  -.060  .080                         ( XXI   )
                                        -2SE    +2SE
```

FIGURE 3.1(f) *Diagnosis:* ACF for the Model Residuals

ary 1962 and the 180th and last observation is December 1976. In March 1974, the 147th month, Cincinnati Bell initiated a 20-cent charge for each call to Directory Assistance. Prior to this time, there was no charge for these calls. The impact of this event is visually striking. In the 147th month, the level of this time series drops abruptly and profoundly.

When an impact is as large as the one in this example, the change in process level complicates identification of the noise component. The change in level is a significant proportion of the series variance which tends to overwhelm the ACF and PACF. To avoid biased estimates of the ACFs and PACFs, only the first 146 observations of the series will be used. ACFs and PACFs used in the identification are shown in Figures 3.1(b) to 3.1(f).

Identification

The ACF and PACF estimated from the raw series indicate a nonstationary ARIMA process. The ACF and PACF estimated from the regularly differenced series, shown in Figure 3.1(c), indicate the ARIMA process is

seasonally nonstationary as well. The ACF and PACF estimated from the regularly *and* seasonally differenced series, shown in Figure 3.1(d), suggest an ARIMA$(0,1,0)(0,1,1)_{12}$ model:

$$N_t = \frac{\Theta_0 + (1 - \Theta_{12}B^{12})}{(1 - B)(1 - B^{12})} a_t.$$

This is an interesting model and somewhat rare. The only autocorrelation is at the seasonal lags.

Estimation

Parameter estimates for the N_t model are:

$$\hat{\Theta}_0 = -.70 \text{ with t-statistic} = -\ 1.51$$
$$\hat{\Theta}_{12} = \ .85 \text{ with t-statistic} = \ 15.30.$$

The estimate of Θ_0 is not statistically significant, so it is dropped from the tentative model. The estimate of Θ_{12} is statistically significant and lies within the bounds of invertibility.

Diagnosis

The residual ACF and PACF, shown in Figure 3.1(e), indicate that the residuals of this model are not different than white noise. There is a significant spike at ACF(5) but nothing else: Q = 12.52 with 24 degrees of freedom is not statistically significant, so the tentative model is accepted. As an aside, we note that another analyst might be concerned about the spike at ACF(5) which might indicate the need for a more elaborate model. Also, the estimated value of Θ_0 is "marginally" significant and some other analyst might decide to keep that parameter in the model. The reader is invited to explore these possibilities.

Impact Assessment

The full impact assessment model is tentatively set as

$$Y_t = \omega_0 I_{147} + \frac{1 - \Theta_{12}B^{12}}{(1 - B)(1 - B^{12})} a_t$$

where $I_{147} = 0$ for the first 146 observations
$= 1$ for the 147th and subsequent observations.

Parameter estimates for the tentative model are:

$$\hat{\Theta}_{12} = .81 \text{ with t statistic} = 11.21$$
$$\hat{\omega}_0 = -39{,}931 \text{ with t statistic} = -17.41.$$

Both estimates are statistically significant and the estimate of Θ_{12} lies within the bounds of invertibility. A final diagnosis indicates that the estimated model is statistically adequate. The residual ACF, shown in Figure 3.1(f), has no significant spikes at all; the Q statistic for this ACF is not statistically significant.

Our interpretation of these findings is obvious. In the 147th month, the level of this series dropped by nearly 40,000 average daily calls to Directory Assistance.

The model-building strategy outlined in this example can be followed generally in all analyses. Each analysis will present a unique set of problems, however, which may require a slight adaptation of the strategy. We note finally that, in this example, a test of the null hypothesis was not at all in question. The impact was visually obvious. Impact assessment analysis nonetheless provided a precise estimate of the form and magnitude of the effect.

3.2 The First-Order Transfer Function

When impacts are as abrupt and dramatic as the one in the Directory Assistance example, the zero-order transfer function will adequately model the impact. Such abrupt, dramatic patterns of impact are rare, however. Most social impacts will be realized gradually, so the zero-order transfer function will not adequately reflect the expected impact. Returning to Zimring's reformulation of the step function in an impact assessment of a gun control law as

$$
\begin{aligned}
I_t &= 0 \text{ prior to enactment} \\
&= 1/6 \text{ in the first month after enactment} \\
&= 2/6 \text{ in the second month} \\
&\ \ \vdots \\
&= 6/6 = 1 \text{ in the sixth and subsequent months,}
\end{aligned}
$$

the reformulation of I_t reflects a *gradual* impact of the new law. Zimring could have measured such an impact by using a first-order transfer function rather than a zero-order function with a reformulated I_t variable.

To date, most social science impact assessments have used the zero-order transfer function exclusively. There are two reasons for this. First, there has been little discussion in the methodological literature of any impact patterns other than the *abrupt, constant* pattern associated with the zero-order transfer function; and the computer software required for estimation of higher order transfer functions has not been widely available. Second, social science time series are often reported as annual statistics; an impact realized gradually over an eight-month period, for example, will appear as an abrupt impact if annual (rather than monthly) data are analyzed. If the data are aggregated in a way that obscures the form of an impact (abrupt *versus* gradual, for example), the zero-order transfer function will adequately model the impact. If the data are aggregated so that the form is not obscured, however, and if an impact is not abrupt, the zero-order transfer function will not adequately model the impact.

A *gradual, permanent* change in process level is implied by the first-order transfer function

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B} I_t$$

where the parameter δ_1 is constrained to the interval

$$-1 < \delta_1 < +1.$$

These constraints are called the *bounds of system stability.* If the value of δ_1 lies outside these bounds, the impact assessment model is unstable. It will be demonstrated later that system instability is identical with nonstationarity. When $\delta_1 \geqslant 1$, the postintervention time series is nonstationary and this is generally interpreted to mean that the event has affected a trend in the time series process.

Again, because the impact assessment model is linear in its two components, a Y_t^* time series can be defined as

$$
\begin{aligned}
Y_t^* &= Y_t - N_t \\
&= \frac{\omega_0}{1 - \delta_1 B} I_t.
\end{aligned}
$$

$$
\text{So} \quad
\begin{aligned}
(1 - \delta_1 B) Y_t^* &= \omega_0 I_t \\
Y_t^* &= \delta_1 Y_{t-1}^* + \omega_0 I_t.
\end{aligned}
$$

This formulation of Y_t^* can be used recursively to examine the behavior of this first-order transfer function. Prior to the event, when $I_t = 0$, the Y_t^* series has a zero level:

$$\begin{aligned} Y_t^* &= \delta_1 Y_{t-1}^* + \omega_0 I_t \\ &= \delta_1(0) + \omega_0(0) = 0. \end{aligned}$$

Now if the final preintervention observation of the series is Y_i^*, the event occurs at $t = i+1$ and $I_{i+1} = 1$. The value of Y_{i+1}^* is thus

$$\begin{aligned} Y_{i+1}^* &= \delta_1 Y_i^* + \omega_0 I_{i+1} \\ &= \delta_1(0) + \omega_0(1) = \omega_0. \end{aligned}$$

In the next postintervention observation, $I_{i+2} = 1$, and the value of Y_{i+2}^* is:

$$\begin{aligned} Y_{i+2}^* &= \delta_1 Y_{i+1}^* + \omega_0 I_{i+2} \\ &= \delta_1(\omega_0) + \omega_0(1) = \delta_1\omega_0 + \omega_0. \end{aligned}$$

And in the next postintervention observation, $I_{i+3} = 1$, and the value of Y_{i+3}^* is:

$$\begin{aligned} Y_{i+3}^* &= \delta_1 Y_{i+2}^* + \omega_0 I_{i+3} \\ &= \delta_1(\delta_1\omega_0 + \omega_0) + \omega_0(1) = \delta_1^2\omega_0 + \delta_1\omega_0 + \omega_0. \end{aligned}$$

Continuing this procedure, it can be shown that, in the n[th] postintervention observation, $I_{i+n} = 1$, and the value of Y_{i+n}^* is:

$$\begin{aligned} Y_{i+n}^* &= \delta_1 Y_{i+n-1}^* + \omega_0 I_{i+n} \\ &= \delta_1(\delta_1^{n-1}\omega_0 + \ldots + \delta_1\omega_0 + \omega_0) + \omega_0(1) \\ &= \sum_{k=0}^{n} \delta_1^k\omega_0. \end{aligned}$$

The importance of constraining the value of δ_1 to the bounds of system stability may now be obvious. So long as δ_1 is a fraction,

$$|\delta_1^{n-1}\omega_0| > |\delta_1^n\omega_0|.$$

Each successive term of the series $\sum_{k=0}^{n} \delta_1^k\omega_0$ is smaller than the previous term. As time passes then, the postintervention series level continues to change (increasing or decreasing, depending upon whether ω_0 is positive or negative) but by smaller and smaller increments (or decrements).

Figure 3.2(a) shows the change expected in the Y_t^* time series with each observation. At the moment of intervention, the series level changes from zero to ω_0, and in the next moment, from ω_0 to $(\omega_0 + \delta_1\omega_0)$. The change in level of Y_t^* from the n − 1[st] to the n[th] postintervention moment is:

$$\left|\sum_{k=0}^{n} \delta_1^k\omega_0\right| - \left|\sum_{k=0}^{n-1} \delta_1^k\omega_0\right| = |\delta_1^n\omega_0|.$$

This will be a very small number, approaching zero as a limit.

The *asymptotic* or eventual change in the level of Y_t^* can be calculated by summing the infinite series

$$\sum_{k=0}^{\infty} \delta_1^k\omega_0 = \text{asymptotic change in level.}$$

Because δ_1 is smaller than unity in absolute value, this infinite series can be evaluated as

$$\sum_{k=0}^{\infty} \delta_1^k\omega_0 = \frac{\omega_0}{1 - \delta_1} = \text{asymptotic change in level.}$$

The asymptotic impact is generally realized at a rate determined by the value of δ_1. Figure 3.2(b) shows the expected patterns of impact for various values of this parameter. When δ_1 is small, near zero for example, the

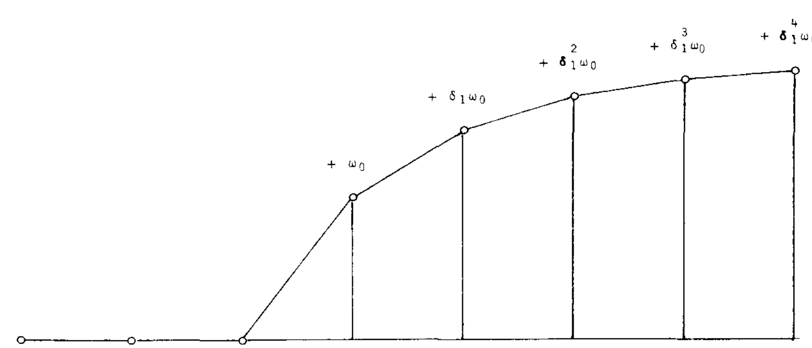

FIGURE 3.2(a) Pattern of Impact Expected of a First-Order Transfer Function

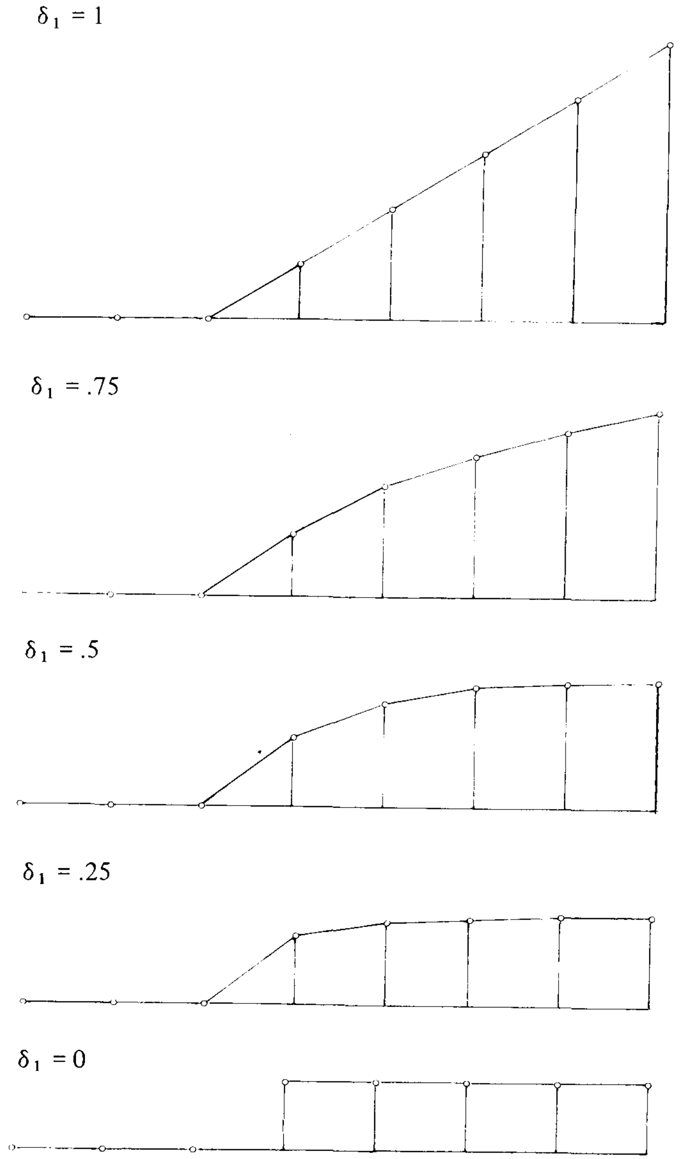

FIGURE 3.2(b) Pattern of Impact Expected for Several Values of δ_1

asymptotic impact is realized quickly. When δ_1 is larger, however, near unity for example, the asymptotic impact is realized slowly. The parameter δ_1 is thus interpreted as a *rate* parameter.

It is of some interest to note the behavior of this first-order transfer function at two values of δ_1. First, when $\delta_1 = 1$, the level of the Y_t^* series changes by the quantity ω_0 in each postintervention moment. This is so because when $\delta_1 = 1$, the first-order transfer function becomes:

$$Y_t^* = \frac{\omega_0}{1 - B} I_t.$$

Prior to the event, when $I_t = 0$, Y_t^* is an ARIMA (0,0,0) process, but when $I_t = 1$, Y_t^* becomes an ARIMA(0,1,0) process. The interpretation here is that, prior to intervention, the series is trendless, whereas postintervention, the series follows a trend with the parameter ω_0 interpreted as the slope.

Intuitively, the case where $\delta_1 = 1$ may seem useful. When examined more closely, however, the substantive implications of this model would seem to limit its utility. In its simplest form, the model describes a fixed-level (or stationary) process which, at the moment of intervention, begins to grow at a constant rate. Such a radical change (from a state of equilibrium to a state of growth) would rarely be observed in the social sciences in our opinion; and if observed, it is unlikely that this change would be associated with a manipulable social intervention.

Nevertheless, it is possible to observe an impact of this sort *when the postintervention time series segment is too short to encompass the equilibrium state of the process*. For example, the intervention may be such that the postintervention process reaches its equilibrium level slowly (the value of δ_1 may be quite large, that is, near unity). If the postintervention time series is too short, however, the postintervention change in level may have the appearance of a change in trend and the analyst may mistakenly conclude that the postintervention process is nonstationary and trending.

The only real solution to this dilemma is to wait for more postintervention data to become available. As these data become available, the analyst will be better able to decide whether the value of δ_1 is unity (and thus, that the postintervention process is trending) or slightly less than unity (in which case the postintervention process is *not* trending). Lacking these data, the analyst must depend upon informed substantive knowledge of the social process under analysis. If a change in slope seems to be a substantively reasonable impact, the parameter ω_0 is interpreted as the postintervention slope.

A second case of interest occurs when $\delta_1 = 0$. As shown in Figure 3.2(b), asymptotic impact is realized instantaneously in this case. This is so because when $\delta_1 = 0$, the first-order transfer function reduces to

$$Y_t^* = \frac{\omega_0}{1-(0)B} = \omega_0 I_t,$$

the zero-order transfer function which we developed in the preceding section. We will make use of this relationship at a later point.

3.2.1 Chlorpromazine Impacts on Perceptual Speed

We are now concerned with the problem of selecting an appropriate intervention component. What are the consequences of using a zero-order transfer function to measure an impact that is not abrupt? Figure 3.2.1(a) shows a time series of 120 daily "perceptual speed" scores for a single schizophrenic patient as reported by Holtzman (1963; see also, Glass et al., 1975). On the 61st day, the patient was placed on a chlorpromazine regimen. Chlorpromazine is a radical tranquilizer, so one might expect a drop in perceptual speed for this patient coincident with the regimen. But is the impact abrupt or gradual? Figure 3.2.1(a), in our opinion, shows a gradual impact with perceptual speed dropping for several days before a new level is realized.

Using only the first 60 observations, we have estimated the ACF and PACF shown in Figure 3.2.1(b). These statistics would seem to indicate an ARIMA (1,0,1) model for the noise component. The basis of this identification is decay in both the ACF and the PACF. As noted in Section 2.9, ARIMA(p,0,q) processes are rarely encountered in social science time series. The ACF and PACF shown in Figure 3.2.1(b) nevertheless support identification of a mixed process.

Although the visual evidence supports a first-order transfer function for the intervention component, we will use a zero-order transfer function to demonstrate our point. The impact assessment model is:

$$Y_t = \omega_0 I_{61} + \frac{1-\Theta_1 B}{1-\phi_1 B} a_t.$$

Parameter estimates for this model are:

$$\begin{aligned}\hat{\phi}_1 &= .96 \text{ with t statistic} = 32.43\\ \hat{\Theta}_1 &= .75 \text{ with t statistic} = 9.75\\ \hat{\omega}_0 &= -27.09 \text{ with t statistic} = -\ 4.40.\end{aligned}$$

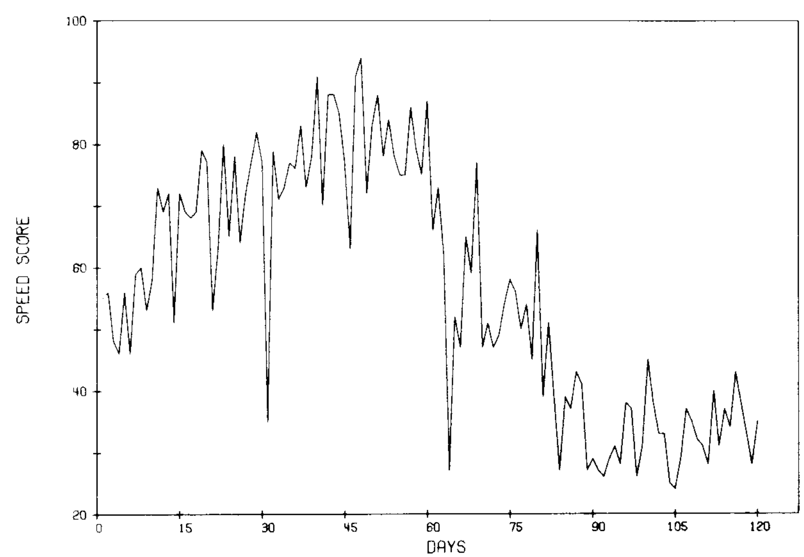

FIGURE 3.2.1(a) Daily Perceptual Speed Scores for a Schizophrenic Patient

While all three parameter estimates are statistically significant, the estimate of ϕ_1 is dangerously close to the bounds of stationarity. This indicates that an ARIMA(1,0,1) model is inappropriate. The 95% confidence interval around this estimate includes the value of $\hat{\phi}_1 = 1$, but, more important, because the estimates of ϕ_1 and Θ_1 are both large and positive, a problem with parameter redundancy is indicated.

In fact, the villain here is an outlier. Returning to Figure 3.2.1(a), we note that the 31st observation of the series is an order of magnitude smaller than neighboring observations. This outlier explains the aberrant ACF and PACF shown in Figure 3.2.1(b). The outlier has exaggerated the variance estimate used in these statistics and thus has biased the estimates of the ACF and PACF downward.

On the basis of the initial estimation, the noise component can be respecified as ARIMA(0,1,1). This leads to the impact assessment model

$$Y_t = \omega_0 I_{61} + \frac{\Theta_0 + (1-\Theta_1 B)}{1-B} a_t.$$

```
SERIES.. SPEED      (NOBS=  60,  PERCEPTUAL SPEED - PRE INTERVENTION
NO. OF VALID OBSERVATIONS  =      60.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25,   60) =   84.676       SIG =     .000

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0   .2   .4   .6   .8  +
                    +----+----+----+----+----+----+----+----+----+----
   1   .427  .129                      (    IXXXXX)XXXXX
   2   .386  .151                    (      IXXXXXXX)XX
   3   .471  .167                    (      IXXXXXXX)XXXX
   4   .404  .187                   (       IXXXXXXXX)X
   5   .428  .201                  (        IXXXXXXXXX)X
   6   .239  .216                 (         IXXXXXX     )
   7   .347  .220                 (         IXXXXXXXXX  )
   8   .267  .229                 (         IXXXXXXX    )
   9   .166  .234                (          IXXXX         )
  10   .324  .236                (          IXXXXXXXXX    )
  11   .109  .244                (          IXXX          )
  12   .155  .244                (          IXXXX         )
  13   .150  .246                (          IXXXX         )
  14   .090  .248                (          IXX           )
  15   .180  .248                (          IXXXXX        )
  16  -.000  .250               (           I               )
  17   .128  .250               (           IXXX            )
  18   .090  .251               (           IXX             )
  19   .006  .252               (           I               )
  20   .039  .252               (           IX              )
  21   .008  .252               (           I               )
  22  -.018  .252               (           I               )
  23  -.016  .252               (           I               )
  24   .011  .252               (           I               )
  25   .065  .252               (           IXX             )
                              -2SE                        +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0   .2   .4   .6   .8  +
                    +----+----+----+----+----+----+----+----+----+----
   1   .427  .129                      (    IXXXXX)XXXXX
   2   .249  .129                      (    IXXXXX*
   3   .314  .129                      (    IXXXXX)XX
   4   .142  .129                      (    IXXXX )
   5   .175  .129                      (    IXXXX )
   6  -.159  .129                      ( XXXXI      )
   7   .111  .129                      (    IXXX   )
   8  -.080  .129                      (   XXI      )
   9  -.074  .129                      (   XXI      )
  10   .166  .129                      (    IXXXX )
  11  -.158  .129                      ( XXXXI      )
  12   .012  .129                      (    I       )
  13  -.008  .129                      (    I       )
  14  -.027  .129                      (   XI       )
  15   .086  .129                      (    IXX     )
  16  -.089  .129                      (   XXI      )
  17   .068  .129                      (    IXX     )
  18  -.004  .129                      (    I       )
  19  -.032  .129                      (   XI       )
  20  -.100  .129                      (  XXXI      )
  21   .074  .129                      (    IXX     )
  22  -.157  .129                      ( XXXXI      )
  23   .069  .129                      (    IXX     )
  24   .073  .129                      (    IXX     )
  25   .053  .129                      (    IX      )
                                     -2SE        +2SE
```

FIGURE 3.2.1(b) ACF and PACF for the Raw Preintervention Series

Parameter estimates for this model are:

$$\hat{\Theta}_0 = .01 \text{ with t statistic } = .06$$
$$\hat{\Theta}_1 = .77 \text{ with t statistic } = 12.93$$
$$\hat{\omega}_0 = -22.13 \text{ with t statistic } = -3.37.$$

As the estimate of Θ_0 is not statistically significant, it must be dropped from the model. All other parameters are statistically significant and otherwise acceptable. Diagnostic checks of the model residuals indicate that they are not different than white noise, so this model is accepted.

The results of this impact assessment lead to the conclusion that the chlorpromazine regimen affected a drop of over 22 units in perceptual speed. This result is consistent at least with what is known about the physiological effects of chlorpromazine. However, the model implies an *abrupt* or instantaneous drop to the new level and this may be inconsistent with what is known about treatments in general.

Using the same ARIMA(0,1,1) model for a noise component, we now specify a first-order transfer function for the intervention component. This leads to the impact assessment model

$$Y_t = \frac{\omega_0}{1 - \delta_1 B} I_{61} + \frac{1 - \Theta_1 B}{1 - B} a_t.$$

Parameter estimates for this model are:

$$\hat{\Theta}_1 = .77 \text{ with t statistic } = 13.01$$
$$\hat{\delta}_1 = .53 \text{ with t statistic } = 2.37$$
$$\hat{\omega}_0 = -13.39 \text{ with t statistic } = -2.23.$$

All parameter estimates are statistically significant and otherwise acceptable. Diagnostic checks of the model residuals indicate that they are not different than white noise, so this model is accepted.

As a first step in interpreting the results of this impact assessment analysis, the asymptotic impact is estimated as

$$\text{asymptotic change } = \frac{\hat{\omega}_0}{1 - \hat{\delta}_1} = \frac{-13.39}{1 - .53}$$
$$= -28.49 \text{ units.}$$

The model implies an impact amounting to a drop of over 28 perceptual speed score units. This impact is realized gradually, however. On the sixth day of the regimen, for example, the change from preintervention is:

$$\begin{aligned} \text{change} &= \hat{\omega}_0(1 + \hat{\delta}_1 + \hat{\delta}_1^2 + \hat{\delta}_1^3 + \hat{\delta}_1^4 + \hat{\delta}_1^5) \\ &= -13.39(1 + .53 + .281 + .149 + .079 + .042) \\ &= -27.86 \text{ units,} \end{aligned}$$

which is 98% of the asymptotic change. On the seventh and successive days, this patient's perceptual speed continues to drop but by negligible amounts.

Now having conducted two impact assessment analyses of the same time series data, we must note that the analyses lead to slightly different conclusions. To be sure, both models imply a drop in perceptual speed as a result of the chlorpromazine regimen. Assuming an *abrupt* response to treatment, however, the estimated reduction is 22.13 units while assuming a *gradual* response to treatment, the estimated reduction is 28.49 units. This is a substantial difference which must be reconciled. One of these estimated impacts must be judged more correct than the other.

If there were no substantive issues involved, the analyst could decide between these two estimates by statistical criteria alone. But whereas purely statistical criteria (such as the RMS statistics) can be used to compare two ARIMA noise models, these same criteria are less important in comparing two impact assessment models. If a model makes the best substantive sense, the analyst may judge it the "best" model regardless of its relative statistical properties. In this case, the first-order transfer function model is "better" than the zero-order transfer function model in both the statistical sense (its RMS statistic is the lower of the two; it fits the time series better) *and* the substantive sense.

At a later point, we will develop a strategy for comparing various low-order transfer function components by statistical criteria. For the time being, however, the analyst should understand that an estimate of impact will vary in quality according to how well the model represents the substantive process.

3.2.2 Sutter County Workforce: Temporary Impacts

A useful model of impact can be generated by applying the first-order transfer function to a *differenced* step function. For the step function I_t,

$$\ldots 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1, \ldots$$

differencing results in

$$\begin{array}{llllll} \ldots(0-0), & (0-0), & (0-0), & (1-0), & (1-1), & (1-1), \ldots \\ \ldots\ \ 0, & 0, & 0, & 1, & 0, & 0, \ldots \end{array}$$

a pulse function, $(1 - B)I_t$, defined such that

$$\begin{aligned} (1 - B)I_t &= 0 \text{ prior to the event} \\ &= 1 \text{ at the onset of the event} \\ &= 0 \text{ thereafter.} \end{aligned}$$

Applying the first-order transfer function to $(1 - B)I_t$, the impact assessment model is:

$$Y_t = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_t + N_t$$

or

$$Y_t^* = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_t$$

$$(1 - \delta_1 B)Y_t^* = \omega_0(1 - B)I_t$$

$$Y_t^* = \delta_1 Y_{t-1}^* + \omega_0(1 - B)I_t.$$

This formulation may now be used to examine the behavior of the first-order transfer function applied to the differenced step function. Prior to the intervention, the step function and the Y_t^* series are both zero. If Y_i^* is the last preintervention observation, then $I_i = 0$ and $I_{i+1} = 1$. Hence,

$$(1 - B)I_{i+1} = I_{i+1} - I_i = 1 - 0 = 1$$

and the value of Y_{i+1}^* is:

$$\begin{aligned} Y_{i+1}^* &= \delta_1 Y_i^* + \omega_0(1 - B)I_{i+1} \\ &= \delta_1(0) + \omega_0(1) = \omega_0. \end{aligned}$$

In the next postintervention moment, the differenced step function is equal to zero:

$$(1 - B)I_{i+2} = I_{i+2} - I_{i+1} = 1 - 1 = 0$$

and the value of Y_{i+2}^* is:

$$\begin{aligned} Y_{i+2}^* &= \delta_1 Y_{i+1}^* + \omega_0(1 - B)I_{i+2} \\ &= \delta_1(\omega_0) + \omega_0(0) = \delta_1\omega_0. \end{aligned}$$

And in the next postintervention moment, the differenced step function is zero again and

$$\begin{aligned} Y_{i+3}^* &= \delta_1 Y_{i+2}^* + \omega_0(1 - B)I_{i+3} \\ &= \delta_1(\delta_1\omega_0) + \omega_0(0) = \delta_1^2\omega_0. \end{aligned}$$

A progression begins to emerge. Continuing this procedure, it can be shown that the nth postintervention observation of the series, Y^*_{i+n}, is:

$$\begin{aligned} Y^*_{i+n} &= \delta_1 Y^*_{i+n-1} + \omega_0(1 - B)I_{i+n} \\ &= \delta_1(\delta_1^{n-2}\omega_0) + \omega_0(0) \\ &= \delta_1^{n-1}\omega_0. \end{aligned}$$

And as the value of δ_1 is constrained to the bounds of system stability, this term will be very small, nearly zero.

Figure 3.2.2 shows the expected impacts for various values of δ_1. The pulse function is distributed across the postintervention time series as a decaying spike. The value of the parameter δ_1 determines the rate at which the process returns to its preintervention equilibrium level. When δ_1 is large, near unity, return to the preintervention equilibrium level is slow. When δ_1 is small, return is rapid.

Friesema et al. (1979) used this abrupt, temporary impact model to assess the economic recovery of small communities from natural disasters. Like the pulse function, disasters are abrupt in onset and short in duration. Even though a natural disaster is short-lived, however, its impact remains for some time afterward. Using a temporary impact model, the parameter δ_1 can be interpreted as the *rate* of recovery during the disaster aftermath.

In Section 2.12.1, we built an ARIMA(0,1,1) $(0,1,1)_{12}$ model for the Sutter County Workforce time series. In December 1955 the 120th month of this series, a flood forced the evacuation of Sutter County. To assess the impact of the flood on the Workforce time series, we can use the ARIMA(0,1,1) $(0,1,1)_{12}$ model as the noise component. The impact assessment model is thus

$$Y_t = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_{121} + \frac{\Theta_0 + (1 - \Theta_1 B)(1 - \Theta_{12}B^{12})}{(1 - B)(1 - B^{12})}a_t.$$

Parameter estimates for this model are:

$$\begin{aligned} \hat{\Theta}_0 &= -\ \ .52 \text{ with t statistic } = -\ \ .22 \\ \hat{\Theta}_1 &= \ \ .60 \text{ with t statistic } = 11.38 \\ \hat{\Theta}_{12} &= \ \ .68 \text{ with t statistic } = 13.33 \\ \hat{\delta}_1 &= \ \ .84 \text{ with t statistic } = 2.64 \\ \hat{\omega}_0 &= -276.44 \text{ with t statistic } = -\ 1.36. \end{aligned}$$

The estimates of Θ_0 and ω_0 are not statistically significant. All parameter estimates are otherwise acceptable and a diagnostic check of the residuals indicates that they are not different than white noise.

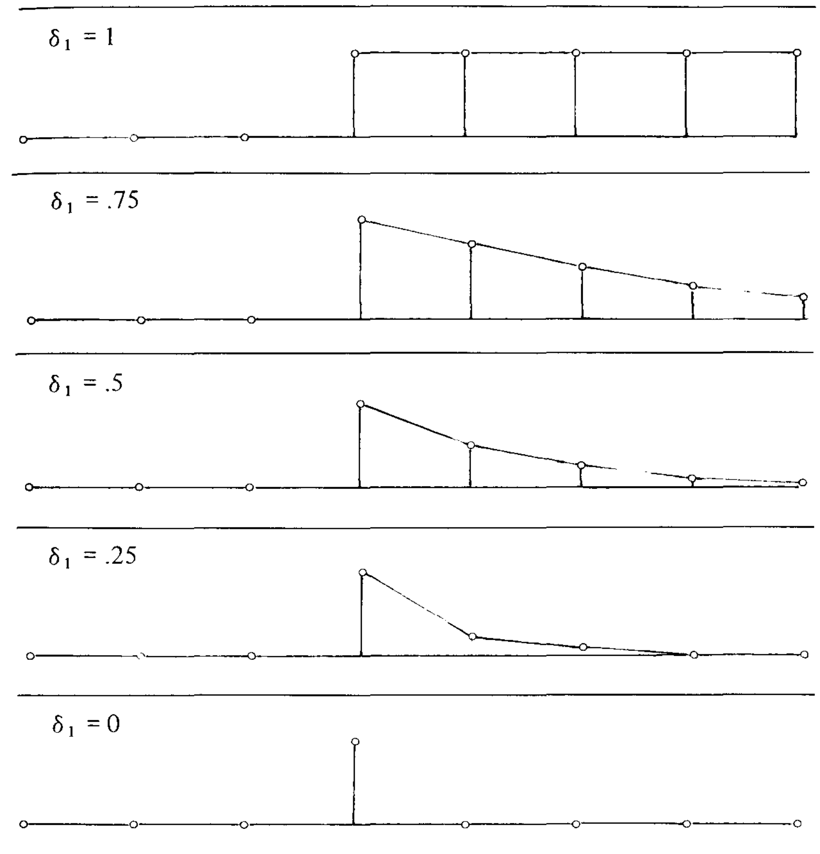

FIGURE 3.2.2 Pattern of Impact Expected for Several Values of δ_1

The parameter Θ_0 must be dropped from the model. A decision on the parameter ω_0 is not so easily made. Because the estimate of this parameter is not statistically different than zero, one might conclude that the flood had no effect whatsoever on the economy of Sutter County. This conclusion would be unsatisfactory, however. It is known in this case that there was an interruption in the local economy and, given this, the analysis must come up with a "best estimate" of the interruption.

According to Friesema et al., the economy of Sutter County is largely agricultural, and as the flood struck in December, after the normal growing season, there was little disruption. By the time of the next growing season, there was little unrepaired damage to local farmlands. Using the estimated values of δ_1 and ω_0 from this analysis, a "best estimate" of the flood's impact is:

$$
\begin{array}{lll}
\text{January 1956:} & \text{displacement} = \hat{\omega}_0 & = -276.44 \\
\text{February 1956:} & \text{displacement} = \hat{\delta}_1\hat{\omega}_0 & = -232.21 \\
\text{March 1956:} & \text{displacement} = \hat{\delta}_1^2\hat{\omega}_0 & = -195.06 \\
\text{April 1956:} & \text{displacement} = \hat{\delta}_1^3\hat{\omega}_0 & = -163.85 \\
\cdot & & \\
\cdot & & \\
\cdot & & \\
\text{December 1956:} & \text{displacement} = \hat{\delta}_1^{11}\hat{\omega}_0 & = -\ 40.61
\end{array}
$$

and so forth. A year after the flood, the economy of Sutter County had returned to its normal condition. The displacement figures are given in worker-months. To estimate the total number of worker-months lost due to the flood, the infinite series

$$\sum_{k=0}^{\infty} \delta_1^k \omega_0$$

can be evaluated with the formula

$$\text{total displacement} = \frac{\hat{\omega}_0}{1 - \hat{\delta}_1} = -1727.75 \text{ worker-months.}$$

This total is interpreted geometrically as the area under the decaying spike. As there are approximately 36,000 worker-months in an average year, the impact of the flood on the Sutter County Workforce time series is substantively trivial.

Note finally that, because the interruption in this time series was relatively small, the noise component was identified with an ACF estimated from the entire time series.

3.2.3 Testing Rival Impact Hypotheses

In preceding sections, we developed three intervention components, each associated with a distinct pattern of impact. These include (1) an *abrupt, constant* pattern of impact determined by the zero-order transfer function

$$f(I_t) = \omega_0 I_t;$$

(2) a *gradual, constant* pattern of impact determined by the first-order transfer function

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B} I_t;$$

and (3) an *abrupt, temporary* pattern of impact determined by applying the first-order transfer function to a differenced I_t

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_t.$$

In an ideal situation, the analyst works from a body of theory which points to one of these three patterns of impact, and hence, to one of these three intervention components. In many cases, for example, theory will define the impact as abrupt *or* gradual, permanent *or* temporary, and so forth. When theory is lacking, however, logical relationships between these three intervention component models (and between the three patterns of impact) will permit a test of rival hypotheses. This is a crucial aspect of impact modeling because, as demonstrated in the Perceptual Speed example of Section 3.2.1, two different intervention component models may often lead to two substantially different estimates of impact.

To illustrate the logical relationships between the three patterns of impact, consider the behavior of the abrupt, temporary impact pattern at the bounds of system stability. Referring to Figure 3.2.2, recovery is instantaneous when $\delta_1 = 0$. When $\delta_1 = 1$, however, there is no recovery at all. Writing out the first-order transfer function associated with this pattern of impact,

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_t = \frac{1 - B}{1 - \delta_1 B}\omega_0 I_t.$$

Whenever $\delta_1 = 1$, the operator terms cancel out and the first-order transfer function reduces to

$$f(I_t) = \omega_0 I_t,$$

the zero-order transfer function.

This relationship suggests a rather simple method for checking the appropriateness of an intervention component. First, if the analyst has no a priori notions about the expected impact, an abrupt, temporary pattern of impact is hypothesized. If the estimated value of δ_1 is too large, near unity, a temporary impact is ruled out.

Next, the analyst hypothesizes a *permanent* but *gradual* pattern of impact based on the first-order transfer function

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B} I_t.$$

If the estimated value of δ_1 in this model is too small, near zero, a *gradual* pattern of impact is ruled out. When $\delta_1 = 0$, in fact, the first-order transfer function reduces to the zero-order transfer function associated with an *abrupt, constant* impact pattern.

To illustrate this procedure, we return to the Directory Assistance time series analyzed in Section 3.1. Eyeballing the plotted time series, Figure 3.1(a), there is no question but that an *abrupt, constant* pattern of impact is appropriate. But suppose now that the pattern of impact is unknown. The "blind" analysis begins with the model

$$Y_t = \frac{\omega_0}{1 - \delta_1 B}(1 - B)I_t + \frac{1 - \Theta_{12}B^{12}}{(1 - B)(1 - B^{12})} a_t.$$

The noise component is the one identified in Section 3.1. The intervention component hypothesizes an *abrupt, temporary* pattern of impact. Estimates for the transfer function parameters are:

$$\hat{\delta}_1 = .99295 \text{ with t statistic } = 70.64$$
$$\hat{\omega}_0 = -38{,}034 \text{ with t statistic } = -13.47.$$

The 95% confidence intervals about this estimate of δ_1 lie well outside the bounds of system stability. The parameter estimate is clearly "too large" to support the temporary effect hypothesis.

As a second step in the "blind" analysis, a permanent but gradual pattern of impact is hypothesized based on the model

$$Y_t = \frac{\omega_0}{1 - \delta_1 B} I_t + \frac{1 - \Theta_{12}B^{12}}{(1 - B)(1 - B^{12})} a_t.$$

Estimates for the transfer function parameters are:

$$\hat{\delta}_1 = -\ .0396 \text{ with t statistic } = -\ .56$$
$$\hat{\omega}_0 = -37{,}900 \text{ with t statistic } = -13.38.$$

The estimated value of δ_1 is clearly "too small" to support the hypothesis of a gradual effect. The only alternative remaining is the abrupt, constant pattern of impact associated with the zero-order transfer function.

The procedure by which competing patterns of impact are ruled out requires a theory in which only a limited number of distinct impact patterns are plausible. In almost all situations, the analyst can invoke the rule of parsimony, "Occam's razor," to make such a theory plausible. In the general case, there are infinitely many possible patterns of impact. There is no logical reason, for example, why second-, third-, and n^{th}-order transfer functions (and their associated patterns of impact) should not be considered by the analyst. The differences among these many patterns of impact are small, however, so for the purposes of ruling out alternative hypotheses, it will be helpful to collapse the infinitely many *possible* effects into two or three distinct classes of impact.

Figure 3.2.3 shows a scheme which we have found useful in practice. An impact is assumed to be *either* permanent *or* temporary, *either* gradual *or* abrupt. With this simple theory of impact, four distinct effects are possible. Three of these four are associated with zero- and first-order transfer functions. If our experiences are typical, almost all social interventions will have impacts reasonably well represented as one of these three models. Because the zero- and first-order transfer functions are related in the extreme, two of the three models can almost always be ruled out through analysis.

The fourth pattern of impact in Figure 3.2.3 is a *gradual, temporary* effect. The impact pattern cannot be easily modeled with a low-order transfer function applied to a step or pulse.[1] This model would seem to be the least useful of the four, so we will not develop it here.

When theory demands it, of course, the impact analysis should not be restricted to the lower order transfer functions. At a later point, we will develop compound lower order and higher order transfer functions which enable the analyst to model virtually any pattern of impact. We will also demonstrate the techniques of impact *fitting* (rather than modeling) which lead to uninterpretable impact assessments. In almost all situations, however, it will be possible to restrict the impact assessment analysis to the patterns of impact shown in Figure 3.2.3 and this restriction will generally pay off in terms of interpretability.

3.3 Interpreting Impact Parameters in the Natural Log Metric

In Chapter 2, we noted that a log transformation of a time series was appropriate when the series variance was proportional to change in the

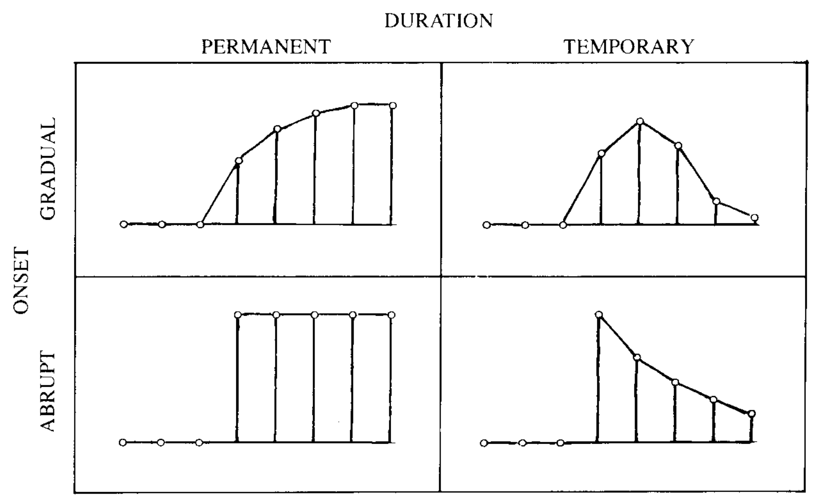

FIGURE 3.2.3 A Simple Theory of Impact

series level. The log-transformed series was then stationary in the larger sense and an appropriate ARIMA model could be fit. In estimating an impact model of such a transformed series, the ω_0 coefficient is interpreted as the pre- to postintervention change in the natural logarithm of the time series.

However precise and correct this interpretation may be, it lacks the easy interpretability of parameters estimated in the raw metric. Any social scientist who has tried to explain transformations and transformed effects to an audience of policy makers will immediately see the practical issue raised here. On one hand, in order to make a proper and correct assessment of impact, we must work in the natural log metric. Yet, on the other hand, by working in the log metric, we lose the easy interpretability of the model parameters.

Fortunately, a simple convention allows us to perform the analysis in the log metric but state our findings in terms of the raw metric. To demonstrate this convention, we must first develop the relationship of impact and ARIMA components in the log metric. Using the simplest intervention component in the log metric, we have

$$\mathrm{Ln}(Y_t) = \omega_0 I_t + \mathrm{ARIMA}.$$

The inverse procedure of the natural log operator is exponentiation, that is,

$$\mathrm{Ln}(x) = k$$
$$e^{\mathrm{Ln}(x)} = e^k = x.$$

To transform the model back into the raw metric then, we exponentiate it:

$$e^{\mathrm{Ln}(Y_t)} = e^{(\omega_0 I_t + \mathrm{ARIMA})}$$
$$Y_t = e^{(\omega_0 I_t)}\, e^{(\mathrm{ARIMA})}.$$

The term $e^{(\mathrm{ARIMA})}$ merely denotes a multiplicative shock form of the ARIMA model. For example, an ARIMA(0,1,0) process in the log metric is:

$$\mathrm{Ln}(Y_t) = Y_0 + a_1 + a_2 + \ldots + a_{t-1} + a_t$$

and exponentiating this,

$$Y_t = e^{(Y_0 + a_1 + a_2 + \ldots + a_{t-1} + a_t)}$$

but let $a_t^* = e^{(a_t)}$ and then

$$Y_t = Y_0^*(a_1^*)(a_2^*)\ldots(a_{t-1}^*)(a_t^*),$$

which is a multiplicative shock model. Similarly, an ARIMA(1,0,0) process in the natural log metric is:

$$(1 - \phi_1 B)\,\mathrm{Ln}(Y_t) = a_t.$$

As we demonstrated in the previous chapter, the ARIMA(1,0,0) process can be written as an infinite series of past shocks,

$$\mathrm{Ln}(Y_t) = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \ldots + \phi_1^{n-1} a_{t-n-1} + \phi_1^n a_{t-n} + \ldots$$

and exponentiating this process, we have

$$Y_t = a_t^*(e^{\phi_1 a_{t-1}})\ldots(e^{\phi_1^n a_{t-n}}),$$

which is a multiplicative shock model. Finally, an ARIMA(0,0,1) process in the log metric is:

$$\mathrm{Ln}(Y_t) = a_t - \Theta_1 a_{t-1}$$

$$\text{so} \quad Y_t = \frac{a_t^*}{e^{(\Theta_1 a_{t-1})}}$$

which is a multiplicative shock model. By exponentiating the ARIMA model, then, we merely change from additive to multiplicative shocks. The exponentiated model still describes the preintervention equilibrium state of the process.

The impact component of the model multiplies the equilibrium state. But prior to the intervention, when $I_t = 0$, the model is:

$$Y_t = e^{(\omega_0 I_t)} e^{(ARIMA)} = e^{(0)} e^{(ARIMA)} = e^{(ARIMA)}.$$

After the intervention, when $I_t = 1$, the model is:

$$Y_t = e^{(\omega_0 I_t)} e^{(ARIMA)} = e^{(\omega_0)} e^{(ARIMA)}.$$

It is convenient to think of pre- and postintervention equilibrium levels of the time series process. The ratio of post- to preintervention equilibrium is:

$$\frac{\text{postintervention equilibrium}}{\text{preintervention equilibrium}} = \frac{e^{(\omega_0)} e^{(ARIMA)}}{e^{(ARIMA)}} = e^{(\omega_0)}.$$

In fact, while the parameter ω_0 is not easily interpreted in the log metric, the term $e^{(\omega_0)}$ can be interpreted as *the ratio of the postintervention series level to the preintervention series level*. This ratio can be expressed as the percent change in the expected value of the process associated with the intervention:

$$\text{percent change} = (e^{(\omega_0)} - 1)\,100.$$

For example, in Section 2.12.2 we identified a noise model for the log-transformed Boston Armed Robbery series. Using the abrupt, permanent pattern of impact proposed by Deutsch and Alt (1977), we specify the tentative impact model

$$\mathrm{Ln}(Y_t) = \omega_0 I_t + \frac{(1-.43B)(1+.19B^{12})}{1-B} a_t.$$

The estimated impact parameter is:

$$\hat{\omega}_0 = -.2070 \text{ with t statistic } = -1.33,$$
so $$e^{-.2070} = .8130$$
and $$\text{percent change} = (.8130 - 1)\,100 = -18.7\%.$$

Thus, we find that introduction of a gun control law in the 112th month is associated with an 18.7% reduction in armed robberies; however, this reduction is not significantly different from zero at the .05 level. This result contrasts sharply with that of Deutsch and Alt who, using an inappropriate noise model as previously discussed, found a statistically significant decline. The reader is referred to Hay and McCleary (1979) for a detailed discussion of the effect of inappropriate noise models on impact parameter estimates in this and other related series.

Interpreting log impact as percent change can be easily applied to dynamic models of impact. To illustrate this, consider the time series plotted in Figure 3.3(a). These are monthly public drunkenness arrests for Minneapolis. In June 1971, the 66th month of this series, public drunkenness was decriminalized in Minnesota. Aaronson et al. (1978; McCleary and Musheno, 1980) claim that decriminalization affected an abrupt and profound drop in the level of arrests and, given the visual appearance of the data, their claim seems reasonable. Starting in the 67th month, the level of this series appears to drop substantially.

A less notable impact (but an impact which is nevertheless noticeable) concerns the series variance. Prior to the intervention, month-to-month fluctuations are relatively large, while postintervention, month-to-month fluctuations are relatively small. Of course, this postintervention change in variance is not a unique impact of decriminalization, but rather is due to a "floor" effect of the sort we alluded to in Section 2.4. As a result of decriminalization, the process drops to a new equilibrium level near the "floor." At this new equilibrium level, the series variance is constrained (see McCleary and Musheno, 1980). The log-transformed series is shown in Figure 3.3(b). In the natural logarithm metric, variance is more nearly constant throughout the length of the series.

As a preliminary step to impact analysis, we will build an ARIMA model for this time series. Due to the magnitude of the intervention, only the 66 preintervention observations will be used for identification.

Identification

An ACF and PACF estimated from the first 66 observations of the log-transformed series, Figure 3.3(c), indicate that the series is seasonally nonstationary. The ACF and PACF for the seasonally differenced log series, Figure 3.3(d), suggest an ARIMA $(0,0,0)\,(0,1,1)_{12}$ model for the noise component.

Estimation

To this noise component, we add an impact component to reflect a gradual, permanent effect pattern. The full model is thus

$$\mathrm{Ln}\,(Y_t) = \frac{\omega_0}{1-\delta_1 B} I_t + \frac{\Theta_0 + (1-\Theta_{12}B^{12})}{1-B^{12}} a_t.$$

(text continued on p. 181)

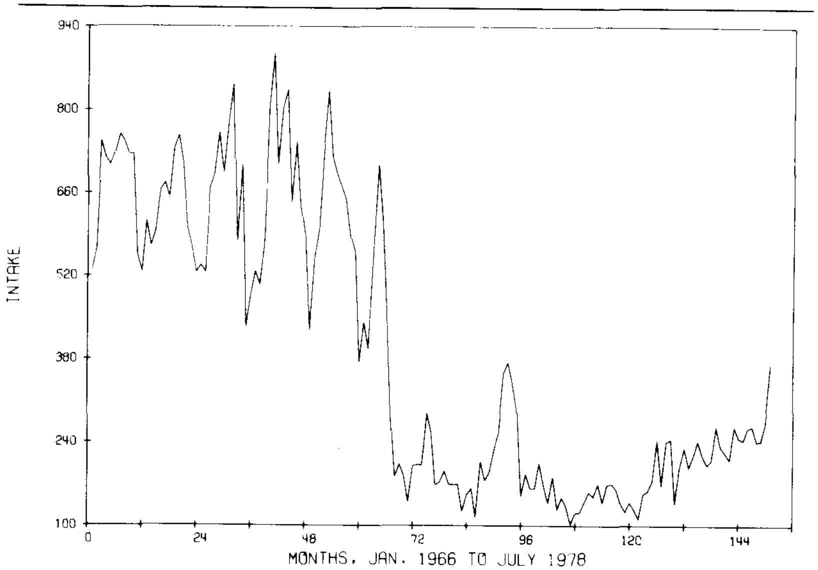

FIGURE 3.3(a) Minneapolis Public Drunkenness Intakes

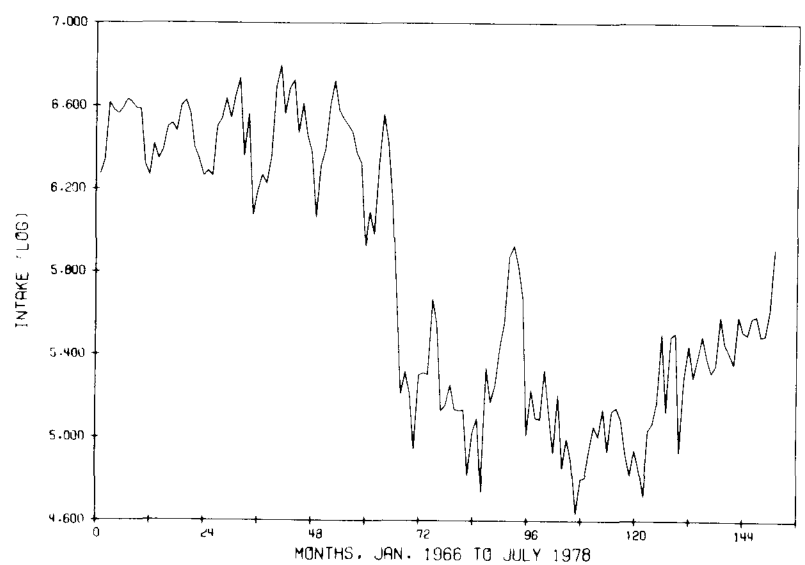

FIGURE 3.3(b) Minneapolis Public Drunkenness Intakes (Logged)


```
SERIES.. DRUNK    (NOBS=  66)  PUBLIC DRUNKENESS - PRE INTERVENTION
TRANSFORMED BY LOG.
NO. OF VALID OBSERVATIONS  =    66.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25,  66) =  167.30      SIG =     .000

 LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                   +----+----+----+----+----+----+----+----+----+----+
   1   .585  .123                     (     IXXXXX)XXXXXXXXX
   2   .327  .160                   (       IXXXXXXX*
   3   .061  .170                   (       IXX     )
   4  -.176  .170                   (   XXXXI       )
   5  -.346  .173                  *XXXXXXXXI        )
   6  -.371  .183                  *XXXXXXXXI        )
   7  -.364  .194                 (XXXXXXXXXI         )
   8  -.335  .204                 ( XXXXXXXXI         )
   9  -.068  .212                (        XXI          )
  10   .195  .213                (          IXXXXX     )
  11   .407  .215                (          IXXXXXXXXXX)
  12   .486  .227                (          IXXXXXXXXXX)X
  13   .371  .242               (           IXXXXXXXXX  )
  14   .155  .250              (            IXXXX        )
  15  -.083  .252              (          XXI            )
  16  -.240  .252              (      XXXXXXI            )
  17  -.350  .256              (   XXXXXXXXXI            )
  18  -.419  .263              (  XXXXXXXXXXI            )
  19  -.373  .273             (    XXXXXXXXXI             )
  20  -.263  .260             (      XXXXXXXI             )
  21  -.027  .284             (            XI             )
  22   .097  .284             (             IXX           )
  23   .301  .285             (             IXXXXXXXX     )
  24   .415  .289             (             IXXXXXXXXXX   )
  25   .314  .298            (              IXXXXXXXX      )
                           -2SE                           +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

 LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                   +----+----+----+----+----+----+----+----+----+----+
   1   .585  .123                     (     IXXXXX)XXXXXXXXX
   2  -.024  .123                     (    XI     )
   3  -.184  .123                     (XXXXXI     )
   4  -.209  .123                     (XXXXXI     )
   5  -.186  .123                     (XXXXXI     )
   6  -.051  .123                     (    XI     )
   7  -.110  .123                     (  XXXI     )
   8  -.158  .123                     ( XXXXI     )
   9   .219  .123                     (     IXXXXX)
  10   .216  .123                     (     IXXXXX)
  11   .188  .123                     (     IXXXXX)
  12   .071  .123                     (     IXX   )
  13  -.136  .123                     (  XXXI     )
  14  -.139  .123                     (  XXXI     )
  15  -.154  .123                     ( XXXXI     )
  16  -.065  .123                     (   XXI     )
  17   .013  .123                     (     I     )
  18  -.046  .123                     (    XI     )
  19   .037  .123                     (     IX    )
  20  -.010  .123                     (     I     )
  21   .054  .123                     (     IX    )
  22  -.213  .123                     (XXXXXI     )
  23   .008  .123                     (     I     )
  24   .166  .123                     (     IXXXX )
  25  -.023  .123                     (    XI     )
                                    -2SE         +2SE
```

FIGURE 3.3(c) ACF and PACF for the Raw (Logged) Preintervention Series

```
SERIES.. DRUNK     (NOBS=  66)  PUBLIC DRUNKENESS - PRE INTERVENTION
TRANSFORMED BY LOG.
DIFFERENCED  1 TIME(S) OF ORDER 12.
NO. OF VALID OBSERVATIONS  =    54.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 25,  54) =  31.360      SIG =      .177

  LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                    +----+----+----+----+----+----+----+----+----+----+
    1   .154  .136                    (      IXXXX  )
    2   .218  .139                    (      IXXXXX )
    3   .252  .145                    (      IXXXXXX)
    4   .259  .153                   (       IXXXXXX )
    5   .029  .161                   (       IX      )
    6   .200  .161                   (       IXXXXX  )
    7   .226  .166                   (       IXXXXXX )
    8  -.101  .171                  (     XXXI        )
    9  -.067  .173                  (      XXI        )
   10   .084  .173                  (        IXX      )
   11   .028  .174                  (        IX       )
   12  -.307  .174                  (XXXXXXXXI        )
   13  -.008  .184                  (        I        )
   14  -.005  .184                  (        I        )
   15  -.270  .184                  ( XXXXXXXI        )
   16   .040  .191                 (         IX        )
   17   .031  .191                 (         IX        )
   18  -.118  .191                 (      XXXI         )
   19  -.188  .192                 (    XXXXXI         )
   20  -.011  .196                 (         I         )
   21   .002  .196                 (         I         )
   22  -.168  .196                 (     XXXXI         )
   23   .031  .198                 (         IX        )
   24  -.047  .199                 (        XI         )
   25  -.121  .199                 (      XXXI         )
                                 -2SE               +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

  LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                    +----+----+----+----+----+----+----+----+----+----+
    1   .154  .136                    (      IXXXX  )
    2   .199  .136                    (      IXXXXX )
    3   .208  .136                    (      IXXXXX )
    4   .188  .136                    (      IXXXXX )
    5  -.102  .136                    (   XXXI      )
    6   .092  .136                    (      IXX    )
    7   .145  .136                    (      IXXXX  )
    8  -.238  .136                    (XXXXXXI      )
    9  -.171  .136                    (  XXXXI      )
   10   .041  .136                    (      IX     )
   11   .084  .136                    (      IXX    )
   12  -.303  .136                   X(XXXXXXI      )
   13  -.037  .136                    (     XI      )
   14   .149  .136                    (      IXXXX  )
   15  -.133  .136                    (   XXXI      )
   16   .198  .136                    (      IXXXXX )
   17   .016  .136                    (      I      )
   18  -.067  .136                    (    XXI      )
   19  -.017  .136                    (      I      )
   20  -.155  .136                    (  XXXXI      )
   21   .071  .136                    (      IXX    )
   22  -.059  .136                    (     XI      )
   23   .022  .136                    (      IX     )
   24  -.133  .136                    (   XXXI      )
   25  -.012  .136                    (      I      )
                                   -2SE          +2SE
```

FIGURE 3.3(d) ACF and PACF for the Seasonally Differenced (Logged) Preintervention Series


```
SERIES.. RESIDUAL (NOBS= 138)  ESTIMATED MODEL RESIDUALS
NO. OF VALID OBSERVATIONS   =  138.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 24, 138) =  131.31      SIG =      .000

  LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                    +----+----+----+----+----+----+----+----+----+----+
    1   .490  .090                      (    IXXXX)XXXXXXX
    2   .420  .100                      (    IXXXX)XXXXXX
    3   .440  .120                     (     IXXXXX)XXXXX
    4   .290  .130                    (      IXXXXXX*
    5   .230  .130                    (      IXXXXXX)
    6   .240  .130                    (      IXXXXXX)
    7   .120  .140                    (      IXXX   )
    8   .070  .140                    (      IXX    )
    9   .050  .140                    (      IX     )
   10   .010  .140                    (      I      )
   11  -.020  .140                    (     XI      )
   12   .010  .140                    (      I      )
   13  -.140  .140                    (  XXXXI      )
   14  -.130  .140                    (   XXXI      )
   15  -.060  .140                    (    XXI      )
   16  -.070  .140                    (    XXI      )
   17  -.020  .140                    (     XI      )
   18  -.030  .140                    (     XI      )
   19   .040  .140                    (      IX     )
   20  -.020  .140                    (     XI      )
   21  -.080  .140                    (    XXI      )
   22   .020  .140                    (      IX     )
   23  -.120  .140                    (   XXXI      )
   24  -.160  .140                    ( XXXXXI      )
   25  -.150  .140                    (  XXXXI      )
                                   -2SE          +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

  LAG   CORR    SE  -1  -.8  -.6  -.4  -.2   0   .2   .4   .6   .8   +1
                    +----+----+----+----+----+----+----+----+----+----+
    1   .490  .090                      (    IXXXX)XXXXXXX
    2   .240  .090                      (    IXXXX)X
    3   .230  .090                      (    IXXXX)X
    4  -.040  .090                      (   XI    )
    5  -.030  .090                      (   XI    )
    6   .050  .090                      (    IX   )
    7  -.080  .090                      (  XXI    )
    8  -.060  .090                      (  XXI    )
    9  -.030  .090                      (   XI    )
   10  -.010  .090                      (    I    )
   11  -.020  .090                      (   XI    )
   12   .050  .090                      (    IX   )
   13  -.180  .090                      *XXXXI    )
   14  -.040  .090                      (   XI    )
   15   .060  .090                      (    IXX  )
   16   .070  .090                      (    IXX  )
   17   .090  .090                      (    IXX  )
   18  -.040  .090                      (   XI    )
   19   .120  .090                      (    IXXX )
   20  -.100  .090                      ( XXXI    )
   21  -.140  .090                      (XXXXI    )
   22   .050  .090                      (    IX   )
   23  -.170  .090                      (XXXXI    )
   24  -.130  .090                      ( XXXI    )
   25  -.030  .090                      (   XI    )
                                     -2SE      +2SE
```

FIGURE 3.3(e) *Diagnosis:* ACF and PACF for the Residuals of an ARIMA (0,0,0) $(0,1,1)_{12}$ Model and Intervention Component

```
SERIES.. RESIDUAL (NOBS= 149)  ESTIMATED MODEL RESIDUALS
NO. OF VALID OBSERVATIONS   =  149.

AUTOCORRELATIONS OF LAGS 1 - 25.
Q( 24, 149) =  46.086      SIG =     .004

 LAG   CORR    SE  -1    -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1   -.340  .080                      XXXXX(XXXI    )
  2   -.050  .090                          (   XI    )
  3    .060  .090                          (    IXX  )
  4   -.040  .090                          (   XI    )
  5   -.140  .090                          (XXXXI    )
  6    .140  .090                          (    IXXXX)
  7   -.100  .090                          ( XXXI    )
  8   -.060  .100                          (  XXI    )
  9   -.010  .100                          (    I    )
 10    .010  .100                          (    I    )
 11    .100  .100                          (    IXXX )
 12   -.020  .100                          (   XI    )
 13    .000  .100                          (    I    )
 14   -.020  .100                          (   XI    )
 15   -.050  .100                          (   XI    )
 16   -.040  .100                          (   XI    )
 17    .040  .100                          (    IX ' )
 18   -.070  .100                          (  XXI    )
 19    .100  .100                          (    IXXX )
 20    .010  .100                          (    I    )
 21   -.190  .100                          *XXXXI    )
 22    .220  .100                          (    IXXXX)X
 23    .000  .100                          (    I    )
 24   -.070  .100                          (  XXI    )
 25    .110  .100                          (    IXXX )
                                         -2SE      +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 25.

 LAG   CORR    SE  -1    -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1   -.340  .080                      XXXXX(XXXI    )
  2   -.190  .080                          X(XXXI    )
  3   -.030  .080                           (   XI   )
  4   -.030  .080                           (   XI   )
  5   -.190  .080                          X(XXXI    )
  6    .020  .080                           (    IX  )
  7   -.090  .080                           ( XXI    )
  8   -.120  .080                           (XXXI    )
  9   -.140  .080                           *XXXI    )
 10   -.100  .080                           (XXXI    )
 11    .080  .080                           (    IXX )
 12    .010  .080                           (    I   )
 13    .000  .080                           (    I   )
 14   -.040  .080                           (   XI   )
 15   -.090  .080                           ( XXI    )
 16   -.100  .080                           (XXXI    )
 17   -.070  .080                           ( XXI    )
 18   -.100  .080                           (XXXI    )
 19    .050  .080                           (    IX  )
 20    .060  .080                           (    IXX )
 21   -.210  .080                          X(XXXI    )
 22    .040  .080                           (    IX  )
 23    .040  .080                           (    IX  )
 24   -.010  .080                           (    I   )
 25    .070  .080                           (    IXX )
                                          -2SE     +2SE
```

FIGURE 3.3(f) *Diagnosis:* ACF and PACF for the Residuals of an ARIMA $(0,1,0)\ (0,0,1)_{12}$ Model and Intervention Component

Parameter estimates for this tentative model are:

$$\hat{\omega}_0 = -\ .992 \text{ with t statistic} = -\ 4.17$$
$$\hat{\delta}_1 = .190 \text{ with t statistic} = .99$$
$$\hat{\Theta}_0 = -\ .007 \text{ with t statistic} = -\ .77$$
$$\hat{\Theta}_{12} = 1.039 \text{ with t statistic} = 24.73.$$

This tentative model is clearly unacceptable. The estimate of Θ_{12} lies outside the bounds of invertibility.

Diagnosis

There is little to be gained returning to the ACFs and PACFs used to identify the ARIMA (0,0,0) $(0,1,1,)_{12}$ noise component. The problem here is that a model identified from the preintervention series only will not adequately reflect the stochastic behavior of the entire series. The noise component in this example appears to be too complicated to be identified with such a short series. We can nevertheless make some educated guess as to the appropriate noise component on the basis of the information available at this point.

First, the unacceptable estimate of Θ_{12} suggests that seasonal differencing is *not* required. Second, the ACF and PACF of the model residuals, Figure 3.3(e), indicate a nonstationary process. We thus tentatively specify an ARIMA (0,1,0) $(0,0,1)_{12}$ model for the noise component.

Estimation

For the tentative model

$$\text{Ln}\,(Y_t) = \frac{\omega_0}{1 - \delta_1 B} I_t + \frac{\Theta_0 + (1 - \Theta_{12}B^{12})}{1 - B} a_t,$$

parameter estimates are:

$$\hat{\omega}_0 = -.543 \text{ with t statistic} = -2.76$$
$$\hat{\delta}_1 = .451 \text{ with t statistic} = 1.71$$
$$\hat{\Theta}_0 = .004 \text{ with t statistic} = .21$$
$$\hat{\Theta}_{12} = -.153 \text{ with t statistic} = -1.80.$$

Two parameter estimates are only marginally significant while the estimate of Θ_0 is clearly insignificant. The estimate of Θ_{12} now lies within the bounds of invertibility.

Diagnosis

The residual ACF and PACF, Figure 3.3(f) now indicate stationarity. Spikes at the first and fourth lags, however, indicate unmodeled moving average terms of those orders. A $\Theta_1 B$ and a $\Theta_4 B^4$ term must be incorporated into the noise component but there is no indication in the ACF or PACF as to how this should be done. There are three possibilities including the two-factor models

$$\mathrm{Ln}(Y_t) = \frac{\omega_0}{1-\delta_1 B} I_t + \frac{\Theta_0 + (1-\Theta_1 B - \Theta_4 B^4)(1-\Theta_{12}B^{12})}{1-B} a_t$$

and

$$\mathrm{Ln}(Y_t) = \frac{\omega_0}{1-\delta_1 B} I_t + \frac{\Theta_0 + (1-\Theta_1 B)(1-\Theta_4 B^4 - \Theta_{12}B^{12})}{1-B} a_t$$

and the three factor model

$$\mathrm{Ln}(Y_t) = \frac{\omega_0}{1-\delta_1 B} I_t + \frac{\Theta_0 + (1-\Theta_1 B)(1-\Theta_4 B^4)(1-\Theta_{12}B^{12})}{1-B} a_t.$$

To be sure, these three models are nearly identical whenever Θ_4 is small; when Θ_4 is zero, they are identical. Judging from the size of the lag-4 ACF in Figure 3.3(f), however, the estimated value of Θ_4 will not be small. The differences in these three models are in their cross-product terms. Expanding the first two-factor noise component,

$$(1-\Theta_1 B - \Theta_4 B^4)(1-\Theta_{12}B^{12}) = \\ (1-\Theta_1 B - \Theta_4 B^4 - \Theta_{12}B^{12} + \Theta_1\Theta_{12}B^{13} + \Theta_4\Theta_{12}B^{16}).$$

Expanding the second two-factor noise component

$$(1-\Theta_1 B)(1-\Theta_4 B^4 - \Theta_{12}B^{12}) = \\ (1-\Theta_1 B - \Theta_4 B^4 - \Theta_{12}B^{12} + \Theta_1\Theta_4 B^5 + \Theta_1\Theta_{12}B^{13}).$$

And expanding the three-factor noise component

$$(1-\Theta_1 B)(1-\Theta_4 B^4)(1-\Theta_{12}B^{12}) = \\ (1-\Theta_1 B - \Theta_4 B^4 - \Theta_{12}B^{12} + \Theta_1\Theta_4 B^5 + \Theta_1\Theta_{12}B^{13} \\ + \Theta_4\Theta_{12}B^{16} - \Theta_1\Theta_4\Theta_{12}B^{17}).$$

The analyst might ordinarily look to the residual ACF in Figure 3.3(f) for evidence favoring one of these three components. A spike at ACF(17), for example, would argue for the three-factor model. We see nothing in the ACF to inform this decision, however. Lacking information from this source, the analyst is advised to estimate all three models, deciding the competition with a comparison of residual statistics. Following this advice, the "best" model of the three is:

$$\mathrm{Ln}(Y_t) = \frac{\omega_0}{1-\delta_1 B} I_t + \frac{\Theta_0 + (1-\Theta_1 B - \Theta_4 B^4)(1-\Theta_{12}B^{12})}{1-B} a_t.$$

This model is the "best" because, of the three competing models, it has the lowest residual mean square (RMS) statistic. To learn this fact, of course, the analyst must estimate all three models.

Estimation

Parameter estimates for this tentative model are:

$$\begin{aligned}
\hat\omega_0 &= -.6116 \text{ with t statistic } = -4.16\\
\hat\delta_1 &= .5186 \text{ with t statistic } = 4.30\\
\hat\Theta_0 &= .0040 \text{ with t statistic } = 1.09\\
\hat\Theta_1 &= .5052 \text{ with t statistic } = 8.14\\
\hat\Theta_4 &= .5727 \text{ with t statistic } = 8.94\\
\hat\Theta_{12} &= -.2384 \text{ with t statistic } = -2.74.
\end{aligned}$$

With the exception of the Θ_0 estimate, all parameter estimates are statistically significant and otherwise acceptable. A diagnostic check of the model residuals indicates that they are not different than white noise.

Whenever a statistically insignificant parameter is dropped from a tentative model, the remaining parameters must be reestimated. The estimate of Θ_0 is not statistically significant and must be dropped from the model for this series. The new estimates of the impact component parameters are:

$$\hat{\omega}_0 = -.6070 \quad \text{with t statistic} = -4.05$$
$$\hat{\delta}_1 = .41287 \text{ with t statistic} = 2.83,$$

which is a substantial change. This should warn the reader to always estimate all parameters of a model simultaneously. In all cases, when a parameter is dropped from a model, estimates of the remaining parameters will change if only slightly.

We can now interpret this result. In the first month following decriminalization, the level of the log-transformed time series process changed by the quantity $\hat{\omega}_0$. This amounted to a decrease and, in successive months, the process level continued to drop. Log levels for the first six months following decriminalization are expected to be:

$$\begin{aligned}
&\text{first month:} && \hat{\omega}_0 && = -\ .607\\
&\text{second month:} && \hat{\omega}_0\,(1 + \hat{\delta}_1) && = -\ .858\\
&\text{third month:} && \hat{\omega}_0\,(1 + \hat{\delta}_1 + \hat{\delta}_1^2) && = -\ .961\\
&\text{fourth month:} && \hat{\omega}_0\,(1 + \hat{\delta}_1 + \hat{\delta}_1^2 + \hat{\delta}_1^3) && = -1.004\\
&\text{fifth month:} && \hat{\omega}_0\,(1 + \hat{\delta}_1 + \hat{\delta}_1^2 + \hat{\delta}_1^3 + \hat{\delta}_1^4) && = -1.021\\
&\text{sixth month:} && \hat{\omega}_0\,(1 + \hat{\delta}_1 + \hat{\delta}_1^2 + \hat{\delta}_1^3 + \hat{\delta}_1^4 + \hat{\delta}_1^5) && = -1.029
\end{aligned}$$

and so forth. The postintervention level of the log process continues to drop but by smaller and smaller increments. The asymptotic change in log level is:

$$\text{asymptotic change} = \frac{\hat{\omega}_0}{1 - \hat{\delta}_1} = -1.034.$$

And thus, by the end of the sixth postintervention month, the log process has achieved 99.52% of its asymptotic change in level. While the process level continues to drop after the sixth month, this change is negligible.

To translate this finding into the raw metric, the analyst needs only to exponentiate the asymptotic change in level:

$$\frac{\text{postintervention equilibrium}}{\text{preintervention equilibrium}} = e^{-1.034} = .35558.$$

This result in turn is translated into a percent change estimate of impact:

$$\text{percent change} = .35558 - 1.0 = -64.44\%.$$

As the preintervention mean for this series is approximately 651 arrests per month, the percent change represents a reduction of approximately 420 arrests per month. This interpretation is consistent with the visual evidence in the plotted time series, Figure 3.3(a).

We urge the reader to replicate this analysis. The manner in which a full model was built is typical of the iterative procedure except, of course, that the preintervention series was not long enough to permit a confident identification of the noise component. Through a conservative series of model-building steps, we were nevertheless able to arrive at an adequate but parsimonious representation of the process underlying this time series.

3.4 Higher Order and Compound Intervention Components

In Section 3.2.3, we developed a modeling strategy based on two characteristics of an impact: onset and duration. An impact can be *either* abrupt *or* gradual in onset and *either* permanent *or* temporary in duration. Using the zero- and first-order transfer functions associated with three distinct patterns of impact, almost any problem in any substantive area can be analyzed. There are two advantages to working from this perspective. First, the results of the analysis are easily interpreted and, second, because the transfer functions are related at the bounds of system stability, alternative impact hypotheses can ordinarily be ruled out through analysis. But should an impact assessment require a more complicated effect model, higher order and compound intervention components can be used. Unfortunately, interpretability of the impact assessment may suffer in many cases.

A higher order intervention component is defined as one with powers of B higher than zero or one. For example,

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B - \delta_2 B^2} I_t$$

is a second-order transfer function. The bounds of system stability for higher order transfer functions are identical with the bounds of invertibility for higher order autoregressive operators (see Section 2.5). For this second-order transfer function then, the bounds are:

$$\begin{aligned}
-1 < \delta_2 &< +1\\
\delta_1 + \delta_2 &< +1\\
\delta_2 - \delta_1 &< +1.
\end{aligned}$$

The behavior of this or any higher order transfer function can be determined by expanding the inverse operator as an infinite series. A simpler recursive method is more practical, however. Working with the Y^*_t time series in this case

$$Y^*_t = \frac{\omega_0}{1 - \delta_1 B - \delta_2 B^2} I_t$$

$$(1 - \delta_1 B - \delta_2 B^2)\, Y^*_t = \omega_0 I_t$$

$$Y^*_t = \delta_1 Y^*_{t-1} + \delta_2 Y^*_{t-2} + \omega_0 I_t.$$

Prior to onset of the event, when $I_t = 0$, $Y^*_t = 0$. At $i + 1$, however, the step function changes from zero to one and the value of Y^*_{i+1} is expected to be:

$$\begin{aligned} Y_{i+1} &= \delta_1 Y^*_i + \delta_2 Y^*_{i-1} + \omega_0 I_{i+1} \\ &= \delta_1 (0) + \delta_2 (0) + \omega_0 (1) = \omega_0. \end{aligned}$$

In the next moment,

$$\begin{aligned} Y^*_{i+2} &= \delta_1 Y^*_i + \delta_2 Y^*_i + \omega_0 I_{i+2} \\ &= \delta_1 (\omega_0) + \delta_2 (0) + \omega_0 (1) = \delta_1 \omega_0 + \omega_0. \end{aligned}$$

And in the next moment,

$$\begin{aligned} Y^*_{i+3} &= \delta_1 Y^*_{i+2} + \delta_2 Y^*_{i+1} + \omega_0 I_{i+3} \\ &= \delta_1 (\delta_1 \omega_0 + \omega_0) + \delta_2 (\omega_0) + \omega_0 (1), \\ &= \delta_1^2 \omega_0 + \delta_1 \omega_0 + \delta_2 \omega_0 + \omega_0, \end{aligned}$$

and so forth. The reader may use this recursive method to determine the behavior of any higher order transfer function. In this case, *two* rate parameters, δ_1 and δ_2, determine the rate at which the process achieves a new equilibrium level. The interpretability of this model is somewhat lessened by including an extra rate parameter, however.

As a general strategy, the analyst can *fit* a higher order transfer function to the time series: δ- and ω-parameters are added to the model until a desired degree of fit is achieved. In model *fitting*, however, the analyst must be conscious of what has been lost. First, the impact assessment is no longer a *confirmatory* analysis based on a null hypothesis. Second, the estimate of effect is likely to be biased. Model fitting takes advantage of properties of the specific realization and it may be incorrect to attribute those properties to the generating process itself. We will demonstrate a model-fitting procedure in the next section.

While the use of higher order transfer functions often reduces model interpretability, the use of *compound* intervention components may increase model interpretability. A *compound* intervention component is defined for our purposes as a sum of two low-order transfer functions. Suppose, for example, that a time series has been impacted by two distinct interventions at two different times. A model reflecting this compound effect would be:

$$Y_t = f(I_t) + f(I_{t+n}) + N_t.$$

The sense of this compound intervention component is that the generating process is impacted once and then is impacted again n observations later.

The distinct impacts of a compound intervention component need not occur at different times, of course. There are many situations in which, on the basis of theory, a single intervention may have two distinct impacts, and in these situations, a compound intervention component can be used as a model of the effect. One of the most useful compound components in our experience is the one composed of the zero-order transfer function

$$f(I_t) = \omega_0 I_t$$

and the first-order transfer function applied to a pulse

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B} (1 - B)\, I_t.$$

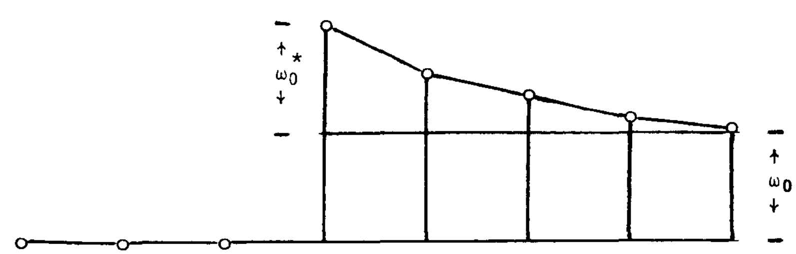

FIGURE 3.4 Pattern of Impact Expected of a Compound Intervention Component

Combining these two components, a compound component is formed as

$$f(I_t) = \omega_0 I_t + \frac{\omega_0^*}{1 - \delta_1 B}.$$

An asterisk (*) is used only to denote the fact that these two zero-order ω-parameters are not the same. Figure 3.4 shows the pattern of impact determined by this compound model. In a literal sense, this impact is the sum of the impacts expected of the two component elements.

We have found this compound intervention model useful in many situations in which an intervention may produce an abrupt but short-lived demonstration effect as well as a larger permanent effect. By accounting for the temporary demonstration effect with a decaying spike, this model permits a more precise estimate of the permanent effect.

To illustrate the use of this compound model, we return to the Perceptual Speed time series. In Section 3.2.1, we used a first-order transfer function to estimate a *gradual, permanent* reduction in perceptual speed scores due to the onset of a chlorpromazine regimen. In many psychological phenomena such as learning, habit formation, and so forth, gradual impacts are expected theoretically as a result of a simple growth process. A "learning curve," for example, may result from a gradual accumulation of correct responses or, alternatively, from a gradual decay of incorrect responses.

There is another mechanism which can result in a gradual, permanent impact, however. First, a treatment itself can result in an abrupt, constant impact of the sort

Before o o o o o o

o o o o o o After.

Associated with the treatment, however, might be a *placebo* (or *trauma; novelty; reactivity;* and so forth) effect of the sort

o

o
o
Before o o o o o o o o o After.

This temporary effect is due only to the novelty (from the perspective of the single schizophrenic patient in this case) of the change in state from "no

treatment" to "treatment." After a few days, the effect wears off, leaving only the true physiological effect. When these two distinct impacts are added together, a *gradual, permanent* impact is realized. For all practical purposes, of course, the analyst cannot tell whether the observed impact is due to a single process (attenuation of a natural resistance to chlorpromazine in this case) or due to two distinct processes (physiological impact plus novelty).

Apologizing for this simplification of theory, we can reanalyze the Perceptual Speed time series with an assumption that there are two distinct impacts of the chlorpromazine regimen. Using the same noise component selected in Section 3.2.1, the impact assessment model corresponding to a double impact theory is:

$$Y_t = [\omega_0 + \frac{1 - B}{1 - \delta_1 B}\omega_0^*]\, I_{61} + \frac{1 - \Theta_1 B}{1 - B} a_t.$$

Parameter estimates for this model are:

$$\begin{aligned}
\hat{\Theta}_1 &= .77 \text{ with t statistic} = 12.97\\
\hat{\omega}_0 &= -28.8 \text{ with t statistic} = -3.50\\
\hat{\delta}_1 &= .51 \text{ with t statistic} = .56\\
\hat{\omega}_0^* &= 18.05 \text{ with t statistic} = 1.72.
\end{aligned}$$

The estimates of δ_1 and ω_0^* are not statistically significant so, strictly speaking, these parameters should be dropped from the model. There are a number of reasons which have nothing to do with the null hypothesis which might explain the lack of statistical significance, however. We will address these reasons shortly.

Assuming that statistical significance is not an issue here, and accepting the estimated parameters of this model at face value, the impact on the n^{th} day of the chlorpromazine regimen is given by

$$\begin{aligned}
\text{impact on day } n &= \hat{\omega}_0 + \hat{\delta}_1^{n-1}\hat{\omega}_0^*\\
&= -28.8 + (.51)^{n-1}(18.05),
\end{aligned}$$

which is merely the sum of the two simple impacts. In Section 3.2.1, on the other hand, impact on the n^{th} day of the chlorpromazine regimen was given by

$$\text{impact on day } n = \sum_{k=0}^{n} \hat{\delta}_1^k \hat{\omega}_0$$
$$= \sum_{k=0}^{n} (.53)^k(-28.49).$$

If the results of these two models are compared for the first seven days of the regimen,

	Compound Model	Gradual Model
First day	−10.75 units	−13.30 units
Second day	−19.64 units	−20.08 units
Third day	−24.15 units	−23.87 units
Fourth day	−26.44 units	−25.89 units
Fifth day	−27.60 units	−26.96 units
Sixth day	−28.19 units	−27.53 units
Seventh day	−28.49 units	−27.83 units,

it is clear that these two models give similar estimates of the impact. Estimates of asymptotic effect (−28.8 units for the compound model *versus* −28.49 units for the gradual model) are nearly identical. *There is nevertheless a substantive difference between the impact estimates of these two models.* Whereas the gradual model predicts the effect based on a single mechanism, the compound model predicts the effect based on two distinct mechanisms. Unique treatment and novelty effects are untangled and estimated.

Because the transfer function parameter estimates for the compound model were statistically insignificant, one might argue that the compound impact theory is unjustified. Multicollinearity is always a problem with compound models, however; we will discuss this problem at some length in Chapter 6. Another problem affecting this particular time series is the outlier in the 31st observation. We will discuss the problem of outliers generally in impact assessments at the end of this chapter but, in all cases, transfer function parameter estimates will be sensitive to outliers. Finally, depending

upon theoretical perspective, statistical significance may be a minor concern. Given a theory where the compound impact model is "true," the transfer function parameter estimates will be interpreted without qualification. In the next section, we will demonstrate the procedures of model *fitting* (as opposed to model building). A comparison of these two sections will illuminate the role of theory in impact assessment.

3.5 U.S. Suicides

The examples preceding this have been straightforward impact assessment analyses, aimed at estimating the effects of discrete interventions (events) on time series. We will now address a related use of impact assessment models. Figure 3.5(a) shows a time series of annual U.S. suicide rates for the 1920–1969 period. The series begins at a level of approximately seventeen suicides (per 100,000 total population), jumps up to approximately 29, and then returns gradually to the starting level. A striking feature of this plot is the set of observations for the 1930s. These are outliers but, in this case, recording errors are not suspected as in the Hyde Park Purse

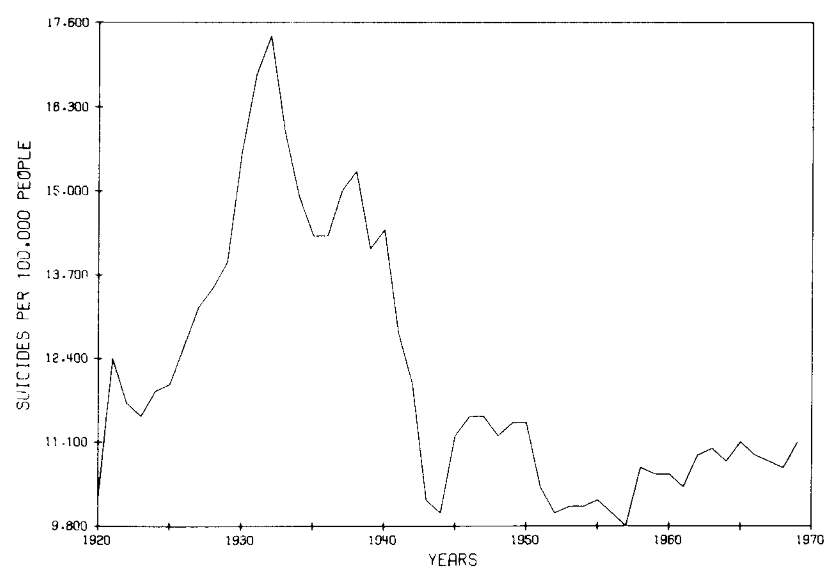

FIGURE 3.5(a) U.S. Suicide Rate, 1920–1969

```
SERIES.. SUICIDES (NOBS=  50)  U.S. SUCIDE RATE, 1920-1969
NO. OF VALID OBSERVATIONS   =     50.

AUTOCORRELATIONS OF LAGS 1 - 10.
Q( 10,  50) =  164.72        SIG =     0.000

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1    .917  .141                      (      IXXXXXX)XXXXXXXXXXXXXXXX
  2    .822  .232                 (           IXXXXXXXXXXX)XXXXXXXXX
  3    .715  .284               (             IXXXXXXXXXXXXXX)XXXX
  4    .625  .318             (               IXXXXXXXXXXXXXXXX*
  5    .551  .342            (                IXXXXXXXXXXXXXX  )
  6    .473  .359           (                 IXXXXXXXXXXXXX    )
  7    .403  .371          (                  IXXXXXXXXXXX       )
  8    .334  .380          (                  IXXXXXXXXX         )
  9    .241  .386          (                  IXXXXXXX           )
 10    .127  .389          (                  IXXX               )
                          -2SE                                 +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 10.

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1    .917  .141                      (      IXXXXXX)XXXXXXXXXXXXXXXX
  2   -.115  .141                      (  XXXI      )
  3   -.126  .141                      (  XXXI      )
  4    .049  .141                      (      IX    )
  5    .050  .141                      (      IX    )
  6   -.111  .141                      (  XXXI      )
  7    .001  .141                      (      I     )
  8   -.027  .141                      (     XI     )
  9   -.222  .141                      (XXXXXXI     )
 10   -.198  .141                      ( XXXXXI     )
                                     -2SE        +2SE
```

FIGURE 3.5(b) ACF and PACF for the Raw Series

Snatchings time series. These extreme values are instead attributed to the obvious exogenous effect of the Great Depression.

Referring to these observations as "outliers" should not imply a substantive and/or theoretical argument about the sociological relevance of economic depression. Some would argue that the Depression was an unfortunate "accident" caused by the improbable intersection of several economic and political events. Others would argue that the Depression was merely an extreme instance of the periodic crises which characterize capitalist economic systems. Although time series analysis could certainly be used profitably to illuminate these questions, that is not our purpose here. We are concerned only with the statistical issues associated with these Depression outliers.

During the 1920–1969 period, the time series appears to be relatively flat, fluctuating about a mean of (approximately) 17 suicides. The exception


```
SERIES.. SUICIDES (NOBS=  50)  U.S. SUCIDE RATE, 1920-1969
DIFFERENCED  1 TIME(S) OF ORDER  1.
NO. OF VALID OBSERVATIONS   =     49.

AUTOCORRELATIONS OF LAGS 1 - 10.
Q( 10,  49) =  7.4952        SIG =      .678

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1    .216  .143                      (      IXXXXX )
  2    .043  .149                      (      IX     )
  3   -.121  .150                      (   XXXI      )
  4   -.068  .152                     (     XXI       )
  5    .032  .153                     (       IX      )
  6   -.012  .153                     (       I       )
  7    .035  .153                     (       IX      )
  8    .173  .153                     (       IXXXX   )
  9    .180  .157                     (       IXXXX   )
 10    .130  .161                     (       IXXX    )
                                    -2SE           +2SE

PARTIAL AUTOCORRELATIONS OF LAGS 1 - 10.

 LAG   CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                    +----+----+----+----+----+----+----+----+----+----+
  1    .216  .143                      (      IXXXXX )
  2   -.004  .143                      (      I      )
  3   -.136  .143                      (   XXXI      )
  4   -.036  .143                      (     XI      )
  5    .072  .143                      (      IXX    )
  6   -.049  .143                      (     XI      )
  7    .028  .143                      (      IX     )
  8    .186  .143                      (      IXXXXX )
  9    .115  .143                      (      IXXX   )
 10    .058  .143                      (      IX     )
                                     -2SE         +2SE
```

FIGURE 3.5(c) ACF and PACF for the Regularly Differenced Series

to this, of course, is the period of time during and following the Depression onset. While the process as a whole may be stationary (or trendless at least), these observations change the appearance of the series. They are outliers and, in a statistical sense only, they violate the assumption of stationarity. The generating process itself may well be a stationary process but this realization of the process is not. From this series alone, the analyst could not infer a stationary property of the generating process.

The consequences of these outliers for identification and estimation are similar to those we discussed in Section 2.12.4 for the Hyde Park time series. Figure 3.5(b) shows the ACF and PACF estimated from the raw suicide time series. While these statistics suggest a nonstationary process, that appearance may be due only to the Depression outliers. Figure 3.5(c) shows the ACF and PACF estimated from the differenced suicide time

```
SERIES.. SUICIDES (NOBS=  50)  U.S. SUCIDE RATE, 1920-1969
SELECTED OBSERVATIONS    24 -     50.
NO. OF VALID OBSERVATIONS  =      27.

AUTOCORRELATIONS OF LAGS 1 - 10.
Q( 10,  27) =  31.691      SIG =      .000

  LAG    CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                     +----+----+----+----+----+----+----+----+----+----+
    1    .653  .192                 (         IXXXXXXXXX)XXXXXX
    2    .319  .262              (            IXXXXXXXX    )
    3    .165  .276             (             IXXXX         )
    4    .051  .280             (             IX            )
    5   -.195  .280             (        XXXXXI             )
    6   -.437  .285             (  XXXXXXXXXXXI             )
    7   -.417  .309            (    XXXXXXXXXXI              )
    8   -.305  .329           (       XXXXXXXXI               )
    9   -.239  .339          (          XXXXXXI                )
   10   -.251  .345          (          XXXXXXI                )
                           -2SE                               +2SE


PARTIAL AUTOCORRELATIONS OF LAGS 1 - 10.

  LAG    CORR    SE  -1   -.8  -.6  -.4  -.2   0    .2   .4   .6   .8  +1
                     +----+----+----+----+----+----+----+----+----+----+
    1    .653  .192                 (         IXXXXXXXXX)XXXXXX
    2   -.186  .192                 (    XXXXXI         )
    3    .070  .192                 (         IXX       )
    4   -.086  .192                 (       XXI         )
    5   -.329  .192                 ( XXXXXXXXI         )
    6   -.259  .192                 (   XXXXXXI         )
    7    .076  .192                 (         IXX       )
    8   -.020  .192                 (        XI         )
    9   -.015  .192                 (         I         )
   10   -.157  .192                 (     XXXXI         )
                                  -2SE                +2SE
```

FIGURE 3.5(d) ACF and PACF for the Raw Series (1943–1969 Only)

series. These indicate white noise but this appearance too may be due only to the Depression outliers.

It is instructive to compare the effects of multiple outliers in this series with the effect of a single outlier in the Hyde Park series. Here the outliers give an overall trend to the series which is reflected in the ACF shown in Figure 3.5(b). Once this trendlike component is removed through differencing, however, the ACF, shown in Figure 3.5(c), indicates a white noise process. The differenced suicide series and the undifferenced Hyde Park series in fact have almost identical ACFs and this should not be surprising. *In both cases, outliers inflate the estimate of process variance and, as a result, understate the values of low-order ACFs.*

Since the Depression observations of this time series are not due to recording error (and hence cannot be "corrected" or adjusted), a noise component must be built around the outliers.

Identification

Figure 3.5(d) shows an ACF and PACF estimated from the 1943–1969 segment of the series. This series is too short to permit a confident identification. Although the ACF and PACF indicate an ARIMA (1,0,0) process, the large standard errors give us little confidence in estimates of the high-order ACF (k) and PACF (k). The reader who replicates our analysis will discover that deleting one or two observations from this segment changes the appearance of these statistics markedly. Acknowledging the weakness of this identification, we tentatively select an ARIMA (1,0,0) model for the noise component.

Estimation

Parameter estimates for the noise component model are:

$$\hat{\phi}_1 = .67 \text{ with t statistic} = 4.47.$$

Although this estimate is based only on the last 27 observations of the series, the estimate is statistically significant and otherwise acceptable. Diagnostic checks of the residuals indicate that the residuals of this model are not significantly different than white noise.

Accepting the ARIMA (1,0,0) model for a noise component, the exogenous shock of the Great Depression will be represented by an impact assessment model of the form

$$Y_t = f(I_t) + \frac{a_t}{1 - \phi_1 B}.$$

In this case, social theories suggest an abrupt, temporary pattern of impact associated with the transfer function

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B}(1 - B)\, I_t$$

$$\text{where} \quad I_t = 0 \text{ prior to } 1930$$
$$= 1 \text{ from } 1930 \text{ on.}$$

visual inspection of the plotted time series raises questions as to the adequacy of this model, however. *Note that the series does not reach its highest point in 1930, but continues to climb higher in 1931 and 1932 before starting to decay.* Also, decay from the 1932 zenith is interrupted by an increase in 1937. (By many accounts, 1937 was the worst year of the

Depression.) To incorporate these impacts into the model, we specify the compound model

$$f(I_t) = \frac{\omega_0 - \omega_1 B - \omega_2 B^2 - \omega_7 B^7}{1 - \delta_1 B}(1 - B)\, I_t.$$

As we have not encountered a higher order transfer function up to this point, this model requires some explanation.

The exogenous shock of the Great Depression is represented by a pulse function, $(1 - B)I_{1930}$, visited upon the economic-social system in 1930. The scalar weight, ω_0, translates this pulse into an increase in the suicide rate which begins immediately to decay. The *rate* of decay is determined by the parameter δ_1. The scalar weights ω_1, ω_2, and ω_7 translate the pulse into delayed spikes, one each occurring in 1931, 1932, and 1937. Each of these delayed spikes also begins immediately to decay at a rate determined by the parmeter δ_1. The seventh-order transfer function is thus identical to *four* decaying spikes. The transfer function in fact can be rewritten as

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B}(1 - B)\, I_t - \frac{\omega_1}{1 - \delta_1 B}(1 - B)\, I_{t-1}$$
$$- \frac{\omega_2}{1 - \delta_1 B}(1 - B)\, I_{t-2} - \frac{\omega_7}{1 - \delta_1 B}(1 - B)\, I_{t-7},$$

which makes this point clear. To draw the expected pattern of impact determined by the transfer function, the reader need only draw four decaying spikes (beginning in 1930, 1931, 1932, and 1937) one on top of the other.

Estimation

Parameter estimates for the full intervention model are:

$$\begin{aligned}
\hat{\phi}_1 &= .87 \text{ with t statistic } = 8.21 \\
\hat{\delta}_1 &= .80 \text{ with t statistic } = 8.00 \\
\hat{\omega}_0 &= 2.08 \text{ with t statistic } = 3.12 \\
\hat{\omega}_1 &= -1.94 \text{ with t statistic } = -3.21 \\
\hat{\omega}_2 &= -1.60 \text{ with t statistic } = -2.62 \\
\hat{\omega}_7 &= -1.26 \text{ with t statistic } = -2.19.
\end{aligned}$$

All coefficients are statistically significant and otherwise acceptable. Our

estimate of ϕ_1 is dangerously close to the bounds of invertibility, however. When autoregressive parameters are so large, we prefer to respecify the noise component. The noise component is thus respecified as ARIMA (0,1,0) and parameters are reestimated as

$$\begin{aligned}
\hat{\omega}_0 &= 1.855 \text{ with t statistic } = 3.26 \\
\hat{\omega}_1 &= -1.851 \text{ with t statistic } = -3.29 \\
\hat{\omega}_2 &= -1.614 \text{ with t statistic } = -2.69 \\
\hat{\omega}_7 &= -1.117 \text{ with t statistic } = -2.04 \\
\hat{\delta}_1 &= .733 \text{ with t statistic } = 5.43.
\end{aligned}$$

All parameters are statistically significant at a .05 level and are otherwise acceptable.

To intrepret this finding, we begin with a zero level in the year preceding the Depression. Increases in successive years are expected to be:

$$\begin{aligned}
&1930: \hat{\omega}_0 = +1.855 \\
&1931: \hat{\delta}_1\hat{\omega}_0 - \hat{\omega}_1 = +3.210 \\
&1932: \hat{\delta}_1^2\hat{\omega}_0 - \hat{\delta}_1\hat{\omega}_1 - \hat{\omega}_2 = +3.967 \\
&1933: \hat{\delta}_1^3\hat{\omega}_0 - \hat{\delta}_1^2\hat{\omega}_1 - \hat{\delta}_1\hat{\omega}_2 = +2.908 \\
&1934: \hat{\delta}_1^4\hat{\omega}_0 - \hat{\delta}_1^3\hat{\omega}_1 - \hat{\delta}_1^2\hat{\omega}_2 = +2.131 \\
&1935: \hat{\delta}_1^5\hat{\omega}_0 - \hat{\delta}_1^4\hat{\omega}_1 - \hat{\delta}_1^3\hat{\omega}_2 = +1.562 \\
&1936: \hat{\delta}_1^6\hat{\omega}_0 - \hat{\delta}_1^5\hat{\omega}_1 - \hat{\delta}_1^4\hat{\omega}_2 = +1.145 \\
&1937: \hat{\delta}_1^7\hat{\omega}_0 - \hat{\delta}_1^6\hat{\omega}_1 - \hat{\delta}_1^5\hat{\omega}_2 - \hat{\omega}_7 = +1.956 \\
&1938: \hat{\delta}_1^8\hat{\omega}_0 - \hat{\delta}_1^7\hat{\omega}_1 - \hat{\delta}_1^6\hat{\omega}_2 - \hat{\delta}_1\hat{\omega}_7 = +1.434
\end{aligned}$$

and so forth. The pattern of impact from this transfer function is *isomorphic* to the effect seen in the plotted time series, Figure 3.5(a). Diagnostic checks of the model residuals indicate that they are not different than white noise.

This model must now be interpreted. The use of a seventh-order transfer function as a model of the Great Depression had only one justification: *it fit the data well*. There is no theoretical basis for assuming that depressions generally will have an impact of this sort on suicide rates. This use of a compound intervention component may be contrasted with the use of a compound intervention component in the analysis of the Perceptual Speed time series in Section 3.4. There a compound intervention component was justified by a theory which predicted that there would be *two* distinct impacts from a single intervention. The results of that analysis could be interpreted within the context of that theory. In this present example, however, no confident statement can be made about the impact of the Depression on suicides except this obvious one: There *appears* to be a substantial effect.

In a broader sense, this impact assessment model has no interpretation whatsoever. The model built here nevertheless may be seen as a "clean" picture of the suicide rate generating process and may be used for a number of purposes not related to impact assessment.

Vigderhous (1978) and Mark (1979), for example, have suggested causal models for this time series, using a time series of annual U.S. unemployment rates to predict suicide rates. A causal model requires a "clean" picture of both unemployment and suicide time series. In the simplest case, a causal model might be diagramed as

$$\text{unemployment} \dashrightarrow \text{suicide}.$$

Yet if the Depression outliers are not removed from both time series, a more complicated causal model,

$$\begin{array}{ccc} \text{unemployment} & \dashrightarrow & \text{suicide}, \\ \uparrow & & \uparrow \\ \text{- - - - -} & \text{depression} & \text{- - - - -} \end{array}$$

might lead to spurious causal inferences. There is no doubt that both unemployment rates and suicide rates were at their highest levels during the Great Depression (and at their lowest levels during World War II). This does not mean that "unemployment causes suicides," however.

In Chapters 4 and 5, we develop the use of ARIMA models and methods for univariate forecasting and multivariate causal analysis. In both of these applications, the analyst must assume that the generating process of a time series in invariant; that the process will continue to generate realizations that are identical with one another within the limits of sampling variance. To be sure, wars and depressions are part of any social science generating process. When only one short realization of the process is available, however, and when that realization includes a war or depression, a number of practical problems arise.

Here it may be instructive to compare social and industrial processes. In the field of industrial control, generating processes are always invariant. Manufacturing processes are relatively constant over time. Error in the process is usually due only to slight variations in the quality of process inputs

(electricity, water, oil, and so on), and as the quality of these inputs is tightly controlled, errors are slight. Input variance is so relatively small, in fact, that it is always well described as white noise.

In the social sciences, on the other hand, generating processes are subject to a wide range of exogenous (input) forces which are not tightly controlled. So long as no single exogenous force exerts a primary influence on the process, the analyst is justified in treating the sum of many exogenous forces as white noise. When a dormant force suddenly asserts itself, however, the effect of that variable cannot be treated ideally as "just another random shock," part of the white noise process. This is a *practical* dictum, of course, for in an infinitely long realization of the suicide rate process, the Great Depression would indeed be "just another random shock."

Wars and depressions commonly exert primary influences on most social indicators. Interaction among nations is ordinarily a stochastic process whose impact on social indicators is adequately described as white noise; except, as a practical matter, in the extreme case of a declared war. Economic fluctuations are similarly stochastic, exerting no strong deterministic influence on social indicators except, as a practical matter, in the extreme case. When only a finite realization of a social process is available, wars and depressions are best thought of as cataclysmic *events* rather than as points on a continuum.

This point is brought home by the analysis of the U.S. Suicide rate time series. While there is no problem *fitting* an intervention component to the series, it would be incorrect to say that this component is a *model* of the Great Depression impact. It is not. Nevertheless, the results of this analysis give a "clean" picture of the suicide rate generating process which may then be used in forecasting or causal modeling.

3.6 A Final Note on the Outlier Problem

In Section 2.12.4, we discussed the problem of outliers in identification. As illustrated by the Hyde Park Purse Snatchings time series, outliers inflate the estimate of process variance and thus understate low-order autocorrelation. This problem is analogous to the problem of estimating a noise component from an impacted series. If the impact is large, as in the Directory Assistance time series, for example, the estimated ACF is practically meaningless. A noise component must be estimated from the preintervention time series only. Identification notwithstanding, outliers cause another specific problem for impact assessment analyses: Transfer function parameter estimates may be unduly influenced by outliers.

The Hyde Park Purse Snatchings time series had an outlier (due to a recording error which was later corrected). In the 42nd observation of the series, a community whistle alert program was implemented (see Reed, 1978). We will estimate the impact of that program on purse snatchings both with and without the incorrect deviant observation.

First, with the outlier: In Section 2.12.4, we identified an ARIMA (0,0,0) model for this series; the raw series was not different than white noise. Reed proposed an abrupt, constant impact for the program, so the impact assessment model is set tentatively as

$$y_t = \omega_0 I_{42} + a_t,$$

$$\text{where } I_{42} = 0 \text{ for the first 41 observations}$$
$$= 1 \text{ for the next 30 observations.}$$

The parameter estimate for this model is:

$$\hat{\omega}_0 = -4.17 \text{ with t statistic} = -1.78.$$

The sign of the parameter estimate is negative, as one would expect if the program had any impact at all, and it is statistically significant at the .10 level. Diagnostic checks of the model residuals indicate that they are not different than white noise. However, given our discussion of ACF distortion associated with outliers, this is to be expected.

Second, with the corrected observation: Considering the effect of a large outlier on the estimation of impact parameters, one might suspect that, because the outlier is in the postintervention segment, it would *inflate* the estimate of postintervention level and thus *deflate* the estimate of impact. By this line of reasoning, one would expect an analysis of the corrected time series to show a larger drop in purse snatchings coincident with the intervention.

But in practice this line of reasoning is flawed. Outliers bias estimates of both the noise and impact component parameters, resulting in a *joint* effect. Replacing the incorrect deviant observation (66 purse snatchings) with its correct value (12) might nevertheless be expected to result in a larger estimate of impact. *Using the statistically inadequate ARIMA (0,0,0) noise component, this is indeed true.* For the model

$$y_t = \omega_0 I_{42} + a_t,$$

the parameter estimate is:

$$\hat{\omega}_0 = -5.97 \text{ with t statistic} = -3.50.$$

A diagnostic check of these model residuals would reveal that the model is not statistically adequate, however. Using the more appropriate ARIMA (2,0,0) for a noise component, parameter estimates are:

$$\hat{\phi}_1 = .28 \text{ with t statistic} = 2.37$$
$$\hat{\phi}_2 = .36 \text{ with t statistic} = 3.09$$
$$\hat{\omega}_0 = -3.07 \text{ with t statistic} = -.87.$$

Both autoregressive parameters are statistically significant and otherwise acceptable. The transfer function parameter is not statistically significant, however. Diagnostic checks of the model residuals indicate that they are not different than white noise, so the model is statistically adequate.

An interpretation of these results is that *there is no evidence to support the hypothesis that this program had an impact on purse snatchings*.

A more important result of this analysis concerns the effect of an outlier on the estimate of impact. Using the same statistically inadequate noise component, the estimate of effect changed by over 40% when the outlier was removed from the time series. The degree of distortion attributable to outliers will depend upon their size and number, the size of the effect, and upon the length of the time series. There will always be some distortion, however.

3.7 Conclusion

We have required a mental leap of the reader from Chapter 2 to Chapter 3. We developed an atheoretical, mechanical model-building strategy in Chapter 2 which leads to an adequate, parsimonious ARIMA model. In Chapter 3, however, we developed a subtler strategy which is not amenable to description as a rigid series of steps or as a flow chart, see Figure 2.11(a). Unlike univariate ARIMA modeling, the *use* of ARIMA models and methods in impact assessment requires a thoughtfully flexible strategy which may change from situation to situation.

The impact itself causes real practical problems for the identification of a noise component. In analyzing the Sutter County Workforce series, for example, the impact was so slight that noise component identification was a simple task. In analyzing the Directory Assistance time series, on the other hand, the impact was so profound that the noise model had to be identified from the 146 preintervention observations. The impact was equally profound in the analysis of the Minneapolis Public Drunkenness series, but in that analysis, only 66 preintervention observations were available. In effect, the noise component had to be identified from the residuals. The procedure in each case was determined by idiosyncracies of the time series under analysis. The decisions involved in a *general* modeling procedure are too many to be described in a simple flow chart.

The considerations involved in selecting an appropriate intervention component are even more complicated. The impact assessment model must be *interpretable,* so the "best" model is always the statistically adequate model whose substantive implications make the most sense. Whereas two competing ARIMA noise components can be compared absolutely by purely statistical criteria (their RMSs, for example), two impact assessment models cannot always be compared along purely statistical dimensions.

In cases in which the analyst can theoretically justify the simple *form* of the expected impact, an intervention component can be selected a priori. The results of the impact assessment analysis then constitute a testing of the theoretically generated null hypothesis. When theory does not point to a single expected impact, however, it is often possible to narrow the possibilities to a few alternatives. A logical system that we have found useful assumes that an impact will be *either* abrupt *or* gradual, *either* permanent *or* temporary. Given these possibilities, the analyst can select a zero- or first-order transfer function for the intervention component. Because these transfer functions are related at their extremes, alternative models can almost always be ruled out in the analysis.

When the expected pattern of impact cannot be limited to a few alternatives, the repertoire of intervention components can be expanded to include *compound* components. The compound intervention component is the sum of two low-order transfer functions and the expected impact of the compound component is simply the sum of the expected impacts of its element components. The compound intervention component is thus easily interpreted. At the extreme, the analyst may *fit* an intervention component to a time series by simply adding δ- and ω-parameters to the intervention component until the fit is complete. An impact assessment analysis based on a *fitted* model is uninterpretable, however. As demonstrated in our analysis of the U.S. Suicide Rate time series, fitted models may be useful but they are not generally interpretable.

These theoretical considerations make ARIMA impact assessment modeling a "confirmatory" analysis. In the analysis of the Sutter County Workforce time series, for example, only one pattern of impact was considered because, on theoretical grounds, only one pattern of impact was plausible. The findings of that analysis were interpretable only in the context of the theory. Similarly, the Perceptual Speed time series was analyzed from two different perspectives. Neither analysis was more or less correct than the other outside of the theoretical context. While the results of these analyses

could be used to illuminate the theoretical context, the results will require a theoretical interpretation.

For Further Reading

The *design* of time series quasi-experiments is thoroughly discussed by Cook and Campbell (1979: Chapter 5) or Glass et al. (1975: Chapters 1–4). The Glass et al. work also has an extensive bibliography of published research. The *analysis* of time series quasi-experiments is developed by McCain and McCleary (1979), McDowall and McCleary (1980), Hibbs (1977), Glass et al. (1975: Chapters 5–7), or Box and Tiao (1975). Glass et al. do not develop seasonal ARIMA models or dynamic intervention components. This work is now outdated and is not generally recommended. The Box-Tiao article must be regarded as the source work for this field. Unfortunately, it may not be accessible to the mathematically unsophisticated reader. The Hibbs article has proved to be extremely influential and is highly recommended.

NOTE TO CHAPTER 3

1. A *gradual, temporary* pattern of impact can be determined by mapping any unimodal function (Normal, Poisson, and so forth) to a pulse. The function will ordinarily be determined by theory in a substantive area. As theory varies tremendously, we have not covered these methods. For an instructive example, however, the reader should see Izenman and Zabel (1980).

4 Univariate ARIMA Forecasts

In this chapter, the shortest one of the volume, we describe univariate ARIMA forecasting methods. While ARIMA methods give the "best" short-range forecasts for a wide variety of time series, there are other univariate forecasting methods which, for some data and in some situations, give "better" forecasts. Much of the forecaster's work involves preparing the "best" forecast in a particular situation. Preparing the forecast itself is not a difficult task and requires little experience. Recognizing the idiosyncracies of each situation, however, and accounting heuristically for these idiosyncracies in the forecast, requires some experience. The reader who is interested primarily in forecasting will not benefit greatly from our treatment of this area. We direct those readers to other sources, particularly to Makridakis and Wheelwright (1978) and to Pindyck and Rubinfeld (1976), where univariate forecasting methods are developed in a richer context.

Our decision to de-emphasize univariate ARIMA forecasting methods in this volume is based on two points. First, almost every book written on the topic of applied time series analysis is concerned exclusively with forecasting. We would have little original thought to add to this body of work. Second, univariate forecasts are usually reliable only in the short range (two or three periods into the future, that is), so univariate ARIMA forecasting is not itself likely to become a widely used method of social research. In the fields of business and management, short-range forecasts can be extremely useful. Managers use month-to-month forecasts to optimize control of inventories, to allocate and schedule salesmen, and so forth. Social scientists, of course, do not ordinarily have such well-defined problems. Unlike the firm, the economy or the social system is seldom representable by a few crucial indicators. More important, economies and social systems are so relatively cumbersome that it would be practically impossible for a policy maker to react to monthly changes in a social indicator.

In contrast to univariate ARIMA forecasts, *multivariate* ARIMA forecasts can be extremely useful in social research. A multivariate forecasting model will ideally account for the joint variation of several social indicators and, based on this structure, will give reliable *long*-range forecasts of a time series. We will develop multivariate ARIMA forecasting methods in the next chapter but that presentation will assume a knowledge of the univariate material developed in this chapter. Beyond this, by learning the algebra of univariate ARIMA forecasting models, the reader will gain a final crucial insight into the nature of the general ARIMA model. Finally, univariate ARIMA forecasts are useful as metadiagnostic tools in many situations. Other things equal, the relative utility and validity of two competing ARIMA models can be compared by contrasting the forecasting abilities of the two models. We will discuss this technique in the concluding section of this chapter.

4.1 Point and Interval Forecasts

All univariate forecasting methods (including ARIMA methods) are based on the same logic. First, the expected value of the time series process is calculated and, second, the expected value is extrapolated into the future. The underlying assumption of this logic is that the process is invariant and this may not always be a wise assumption. It is nevertheless an assumption which the forecaster must be willing to make.

If the current time series observation is Y_t, then we are interested in predicting the values of $Y_{t+1}, Y_{t+2}, \ldots, Y_{t+n}$. We will denote our ARIMA forecast of Y_{t+n} by $Y_t(n)$. We call $Y_t(n)$ *the origin-t forecast of Y with a lead time of n observations*.

As a first step in generating an estimate of $Y_t(n)$, we calculate the expected value of the Y_t process. Our calculations will be simplified considerably if we work in terms of the deviate process, y_t. Noting that the Y_t and y_t processes are related by

$$Y_t = y_t + \Theta_0,$$

we can translate our calculations back into the Y_t metric simply by adding a constant to our result.

Now there are actually *two* expected values of a time series process which can be used for univariate forecasts: the *un*conditional and the *conditional* process expectations. To illustrate the differences between these two expectations, consider the ARIMA(1,0,0) process

$$(1 - \phi_1 B)y_t = a_t.$$

As demonstrated in Chapter 2, this process can be expressed identically as an exponentially weighted sum of past shocks:

$$\begin{aligned} Y_t &= (1 - \phi_1 B)^{-1} a_t \\ &= (1 + \phi_1 B + \phi_1^2 B^2 + \ldots + \phi_1^n B^n + \ldots)\, a_t \\ &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \ldots + \phi_1^n a_{t-n} + \ldots . \end{aligned}$$

Taking the expected value of this expression,

$$Ey_t = Ea_t + \phi_1 Ea_{t-1} + \phi_1^2 Ea_{t-2} + \ldots + \phi_1^n Ea_{t-n} + \ldots .$$

Because the expected value of any random shock is zero, the expected value of the infinite series is zero and

$$\begin{aligned} Ey_t &= 0 + \phi_1 (0) + \phi_1^2 (0) + \ldots + \phi_1^n (0) + \ldots \\ &= 0 \end{aligned}$$

and thus

$$EY_t = Ey_t + \Theta_0 = \Theta_0.$$

This is the unconditional expectation of an ARIMA (1,0,0) process. Extrapo lating this term into the future,

$$\begin{aligned} y_t(1) &= Ey_{t+1} = 0 \\ y_t(2) &= Ey_{t+2} = 0 \\ &\vdots \\ y_t(n) &= Ey_{t+n} = 0. \end{aligned}$$

When the unconditional expectation of the process is used as a univariate forecast, the process mean is the forecast regardless of lead time.

The problem with forecasts based on the unconditional expectation of a process is that much valuable information is ignored. While each past random shock of the process was *expected* to be zero, for example, almost all of these shocks will not be exactly zero. (A time series process in which each shock is zero will have zero variance.) In fact, we know the precise values of some of these shocks (indirectly) and this information could be useful if it were incorporated into the forecast model.

The *conditional* process expectation uses this information. The conditional expectation of y_{t+1} is:

$$E(y_{t+1} \mid y_t, y_{t-1}, \ldots, y_2, y_1).$$

The conditional expectation of y_{t+1} is conditional upon the t preceding observations of the time series process. Expressing the ARIMA (1,0,0) process again as the exponentially weighted sum of past random shocks, the conditional expectation of y_{t+1} is:

$$Ey_{t+1} = Ea_{t+1} + \phi_1 a_t + \phi_1^2 a_{t-1} + \ldots + \phi_1^n a_{t-n-1} + \ldots$$

Now, to be sure, only t random shocks of this process are known. The value of a_{t+1}, the next shock of the process, and the values of $a_0, a_{-1}, \ldots, a_{-\infty}$, which predate the start of the observed time series, are unknown. While distantly past shocks have not been observed, however, and thus are unknown, the *sum* of these shocks is known. Specifically,

$$y_t = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \ldots + \phi_1^n a_{t-n} + \ldots$$

and this known quantity can be substituted into the expression for the conditional expectation of y_{t+1}. Thus,

$$Ey_{t+1} = Ea_{t+1} + \phi_1 y_t$$

and as the expected value of a_{t+1} is zero,

$$Ey_{t+1} = \phi_1 y_t.$$

Conditional expectation forecasts of the ARIMA (1,0,0) process are, then,

$$\begin{aligned} y_t(1) &= E(a_{t+1} + \phi_1 y_t) \\ &= \phi_1 y_t \\ &\vdots \\ y_t(n) &= E(a_{t+n} + \phi_1 a_{t+n-1} + \ldots + \phi_1^{n-1} a_{t+1} + \phi_1^n y_t) \\ &= \phi_1^n y_t. \end{aligned}$$

It should be intuitively plausible that the "best" forecast of a time series process is the *conditional* expectation of the process. What we mean by

"best" in this context is that the conditional expectation forecast has the lowest possible *mean-square forecast error* (MSFE) of any expectation-based forecast.[1]

The notion of MSFE is that point estimates (as opposed to interval estimates which we will discuss shortly) of $y_t(1), y_t(2), \ldots, y_t(n)$ will be compared with their observed values. When sufficient time has elapsed so that values of $y_{t+1}, y_{t+2}, \ldots, y_{t+n}$ are known, MSFE is computed from the formula

$$MSFE = 1/n \sqrt{\sum_{i=1}^{n} [y_{t+i} - y_t(i)]^2}.$$

The MSFE can be used to compare two forecast models of the same time series. A more important use, however, is in the estimation of *interval* forecasts.

It will ordinarily be of some use to set confidence intervals around each point estimate of $y_t(n)$. While the analyst is naturally interested in the expected value of y_{t+n}, this value is meaningless without some idea of how far away the real value of y_{t+n} is likely to be from this expected value. The point estimate forecast of the process may be generated directly from the difference equation form of the ARIMA model (though this may be computationally inefficient in some cases). To generate interval estimates, however, the model must be solved for Y_t. Solved, the ARIMA model expresses Y_t as a weighted sum of past shocks which, by convention, is expressed as

$$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots + \psi_k a_{t-k} + \ldots.$$

The ψ-weights in this solved ARIMA (p,d,q) $(P,D,Q)_S$ model are determined by the model structure (the values of p, d, q, P, D, Q, and S, that is) and by the values of the ϕ and Θ parameters. An ARIMA (1,0,0) $(1,0,0)_4$ model, for example,

$$(1 - \phi_1 B)(1 - \phi_4 B^4) y_t = a_t,$$

is solved as

$$\begin{aligned} y_t &= (1 - \phi_1 B)^{-1}(1 - \phi_4 B^4)^{-1} a_t \\ &= (1 + \phi_1 B + \phi_1^2 B^2 + \ldots)(1 + \phi_4 B^4 + \phi_4^2 B^8 + \ldots) a_t \\ &= a_t + \phi_1 a_{t-1} + \ldots + (\phi_1^4 + \phi_4) a_{t-4} \\ &\qquad + (\phi_1^5 + \phi_1\phi_4) a_{t-5} + \ldots. \end{aligned}$$

so the ψ-weights of an ARIMA (1,0,0) $(1,0,0)_4$ model are:

$$\begin{aligned}\psi_1 &= \phi_1\\ \psi_2 &= \phi_1^2\\ \psi_3 &= \phi_1^3\\ \psi_4 &= \phi_1^4 + \phi_4\\ \psi_5 &= \phi_1^5 + \phi_1\phi_4\end{aligned}$$

and so on. Any ARIMA (p,d,q) $(P,D,Q)_S$ model can be rewritten in its ψ-weight form. To do this in the general case, the model is first solved for y_t and then the common powers of B are collected as in this example. Rewriting an ARIMA model in its ψ-weight form involves tedious arithmetic which we will avoid from this point on. Of course, in preparing forecasts, the analyst will always use a computer forecasting program to estimate the ψ-weights required for interval estimates of $y_t(n)$.

So long as the ϕ and Θ parameters of the ARIMA model lie within the bounds of stationarity-invertibility, the infinite sequence of ψ-weights converges to zero. The value of the k^{th} weight, ψ_k, is thus approximately zero and this is a fortunate (though inevitable) consequence. As there are only t values of a time series available for computation, the ψ-weight form would not always be useful in preparing interval estimate forecasts. The sequence of ψ-weights is infinite but a_0 and all preceding shocks are unknown. Because the sequence of ψ-weights converges to zero, however, we may usually take advantage of the fact that

$$\psi_k a_{t-k} \simeq Ea_{t-k} = 0.$$

So the expected values (zero) of distantly remote shocks may be substituted into the ψ-weight form without appreciably affecting the precision of interval estimate forecasts.[2]

To derive interval forecasts, we note that y_{t+1} can be expressed as the series

$$y_{t+1} = a_{t+1} + \psi_1 a_t + \psi_2 a_{t-1} + \ldots + \psi_k a_{t-k+1} + \ldots$$

As the value of the future random shock, a_{t+1}, is unknown, we must substitute its expected value (zero) into this expression. Then assuming that all ψ-weights and past shocks are known (or, alternatively, assuming that $\psi_k a_{t-k+1}$ is approximately zero), the conditional expectation of y_{t+1} is:

$$\begin{aligned}Ey_{t+1} &= Ea_{t+1} + \psi_1 a_t + \ldots + \psi_k a_{t-k+1} + \ldots\\ &= \psi_1 a_t + \ldots + \psi_k a_{t-k+1} + \ldots.\end{aligned}$$

Using this conditional expectation as a forecast of y_{t+1}, the *error in forecasting* is:

$$\begin{aligned}e_{t+1} &= y_{t+1} - Ey_{t+1}\\ &= y_{t+1} - y_t(1) = a_{t+1}.\end{aligned}$$

The forecast error in other words is equal to y_{t+1} minus its forecasted value. This error will always be equal to the random shock, a_{t+1}, and the forecast *variance* is thus

$$VAR(1) = Ee_{t+1}^2 = \sigma_a^2,$$

which is the variance of the white noise process and, also, a function of the expected MSFE. An interval forecast of y_{t+1} is thus

$$-1.96\sqrt{VAR(1)} < y_t(1) < +1.96\sqrt{VAR(1)}.$$

We expect y_{t+1} to lie in this interval 95% of the time.

If we now wish to forecast the next value of the process, we begin with the ψ-weight expression for y_{t+2},

$$y_{t+2} = a_{t+2} + \psi_1 a_{t+1} + \psi_2 a_t + \ldots + \psi_k a_{t-k+2} + \ldots.$$

The first two shocks are unknown. Substituting their expected value, we obtain the conditional expectation

$$Ey_{t+2} = \psi_2 a_t + \ldots + \psi_k a_{t-k+2} + \ldots.$$

Using this conditional expectation as a forecast, the error is:

$$\begin{aligned}e_{t+2} &= y_{t+2} - y_t(2)\\ &= a_{t+2} + \psi_1 a_{t+1}.\end{aligned}$$

And variance of this forecast is

$$\begin{aligned}VAR(2) &= Ee_{t+2}^2 = E[(a_{t+2} + \psi_1 a_{t+1})^2]\\ &= E(a_{t+2}^2 + 2a_{t+2}\psi_1 a_{t+1} + \psi_1^2 a_{t+1}^2)\\ &= Ea_{t+2}^2 + 2\psi_1 Ea_{t+2}a_{t+1} + \psi_1^2 Ea_{t+1}^2\\ &= \sigma_a^2 + \psi_1^2\sigma_a^2 = (1 + \psi_1^2)\,\sigma_a^2.\end{aligned}$$

Interval estimates of $y_t(2)$ are thus

$$-1.96\sqrt{VAR(2)} < y_t(2) < +1.96\sqrt{VAR(2)}$$

with the same 95% interpretation. We note finally that VAR(2) will always be larger than VAR(1) except when $\psi_1^2 = 0$.

To forecast the next value of the process, we begin with the expression for y_{t+3},

$$y_{t+3} = a_{t+3} + \psi_1 a_{t+2} + \psi_2 a_{t+1} + \psi_3 a_t + \dots.$$

The first three shocks of this expression are unknown, so substituting their expected value, the conditional expectation of y_{t+3} is:

$$Ey_{t+3} = \psi_3 a_t + \dots + \psi_k a_{t-k+3} + \dots.$$

Using this conditional expectation as our forcast, the error is:

$$\begin{aligned} e_{t+3} &= y_{t+3} - y_t(3) \\ &= a_{t+3} + \psi_1 a_{t+2} + \psi_2 a_{t+1}. \end{aligned}$$

Forecast variance is thus

$$\begin{aligned} VAR(3) &= Ee_{t+3}^2 = E[(a_{t+3} + \psi_1 a_{t+2} + \psi_2 a_{t+1})^2] \\ &= (1 + \psi_1^2 + \psi_2^2)\,\sigma_a^2. \end{aligned}$$

Continuing this procedure, we can demonstrate that the forecast variance for $y_t(n)$ is:

$$\begin{aligned} VAR(n) &= Ee_{t+n}^2 = E[(a_{t+n} + \psi_1 a_{t+n-1} + \dots + \psi_{n-1} a_{t+1})^2] \\ &= (1 + \psi_1^2 + \dots + \psi_{n-1}^2)\,\sigma_a^2. \end{aligned}$$

As lead time increases, forecast variance increases and the width of interval estimates of $y_t(n)$ increase. This is an intuitively satisfying result. The farther out into the future we predict, the farther out on a limb we climb.

In the next section, we will examine forecast variances of several ARIMA processes. First, however, we must note that *our discussion of VAR(n) has assumed that the true values of the* ψ-weights are known. *As these weights are determined by the ARIMA structure and by the values of* ϕ *and* Θ *parameters, this amounts to an assumption that we know the true structure of the process.* This assumption is never satisfied. In practice, ambiguity in identification and errors in estimation always leave some doubt as to the true ψ-weights. When we are dealing with only an approximation of the true ARIMA structure (albeit a close approximation), the true value of VAR(n) may be larger than the expected values we have given here.

This understimation of VAR(n) is especially a problem when forecasting relatively short time series. In general, identification and estimation of an ARIMA model becomes easier and more definite as the length of the time series increases. For a sufficiently long series, due to the approximate equivalence of the various ARIMA models, any model selected through the iterative identification/estimation/diagnosis strategy outlined in Section 2.11 will give ψ-weights quite close to their true values. When a time series is relatively short, however, we will have less confidence in the model selected and, as a consequence, we will expect our estimates of VAR(n) to be understated.

Denoting an estimate of the k^{th} ψ-weight by ψ_k^*, the estimation error due only to a poor ARIMA model is:

$$u_k = \psi_k - \psi_k^*.$$

The true error associated with $y_t(n)$ is:

$$e_{t+n} = a_{t+n} + (\psi_1^* + u_1)a_{t+n-1} + \dots + (\psi_{n-1}^* + u_{n-1})a_{t+1}$$

The true forecast variance is:

$$VAR(n) = [1 + (\psi_1^* + u_1)^2 + \dots + (\psi_{n-1}^* + u_{n-1})^2]\,\sigma_a^2.$$

The estimated forecast variance will understate this true VAR(n) by the term

$$(u_1^2 + 2u_1\psi_1^* + \dots + u_{n-1}^2 + 2u_{n-1}\psi_{n-1}^*)\sigma_a^2.$$

Underestimation of VAR(n) due to a poor model is thus a function both of the size of each u_k error and the size of each true ψ_k. This points out a salient difference between ARIMA modeling generally and ARIMA modeling for forecasts. When a model is to be used strictly for forecasting, the analyst may conclude that the "best" model is one in which the standard errors of ϕ and Θ parameters are smallest. Other things equal, such models will have the smallest u_k terms.

4.2 ARIMA Forecast Profiles

We have demonstrated that the origin-t forecast with lead time of n observations for an ARIMA (p,d,q) $(P,D,Q)_S$ process is given by the conditional expectation of y_{t+n}

$$y_t(n) = \psi_n a_t + \psi_{n+1} a_{t-1} + \ldots + \psi_{n+k} a_{t-k} + \ldots.$$

The forecast variance of $y_t(n)$ is:

$$\mathrm{VAR}(n) = (1 + \psi_1^2 + \psi_2^2 + \ldots + \psi_{n-1}^2)\sigma_a^2.$$

The characteristic behavior of an ARIMA forecast, the "forecast profile," is thus determined solely by the ψ-weights of the process. Since the ψ-weights are determined by ARIMA structures, each class of ARIMA models has a characteristic profile.

White Noise

An ARIMA (0,0,0) process written as

$$y_t = a_t$$

has uniformly zero ψ-weights:

$$\psi_1 = \psi_2 = \ldots = \psi_k = 0.$$

Point forecasts of an ARIMA(0,0,0) process are thus

$$\begin{aligned} y_t(1) &= Ea_{t+1} = 0 \\ y_t(2) &= Ea_{t+2} = 0 \\ &\vdots \\ y_t(n) &= Ea_{t+n} = 0. \end{aligned}$$

Because all ψ-weights are zero, variance about these point estimate forecasts is constant for all lead times:

$$\begin{aligned} \mathrm{VAR}(1) &= \sigma_a^2 \\ \mathrm{VAR}(2) &= \sigma_a^2 \\ &\vdots \\ \mathrm{VAR}(n) &= \sigma_a^2. \end{aligned}$$

For white noise processes, the conditional and unconditional expectations are identical. The best forecast is thus the process mean and the history of the process yields no information which can be used to improve upon this prediction.

Integrated Processes

An ARIMA (0,1,0) process, or random walk, written as

$$Y_t = Y_{t-1} + a_t$$

has unit ψ-weights

$$\psi_1 = \psi_2 = \ldots = \psi_k = 1.$$

Point forecasts of an ARIMA (0,1,0) process are thus

$$\begin{aligned} Y_t(1) &= E(Y_t + a_{t+1}) = Y_t \\ Y_t(2) &= E(Y_t + a_{t+1} + a_{t+2}) = Y_t \\ &\vdots \\ Y_t(n) &= E(Y_t + a_{t+1} + \ldots + a_{t+n}) = Y_t. \end{aligned}$$

The best forecast of a random walk is the last observation. If the random walk is differenced, the forecast profile of the z_t process is a white noise profile. In terms of the Y_t process, however, forecast variance increases at a linear rate with respect to lead time:

$$\begin{aligned} \mathrm{VAR}(1) &= Ea_t^2 = \sigma_a^2 \\ \mathrm{VAR}(2) &= E[(a_{t+2} + a_{t+1})^2] = 2\sigma_a^2 \\ &\vdots \\ \mathrm{VAR}(n) &= E[(a_{t+n} + a_{t+n-1} + \ldots + a_{t+1})^2] = n\sigma_a^2. \end{aligned}$$

For each observation increase in lead time, VAR(n) increases by one unit of white noise variance.

In Figure 4.2(a), we show forecasts of "Series B," a time series of IBM stock prices introduced in Section 2.1. This series follows a random walk with a forecast profile typical of all integrated processes. After two or three steps into the future, the confidence intervals (set at 95% in this figure) become so large as to render the interval forecast meaningless. A nonstationary process in fact is defined as one with no finite variance. The limit of VAR(n) for an integrated process as lead time increases to infinity is infinity.

Autoregression

An ARIMA (1,0,0) process written as

$$y_t = \phi_1 y_{t-1} + a_t$$

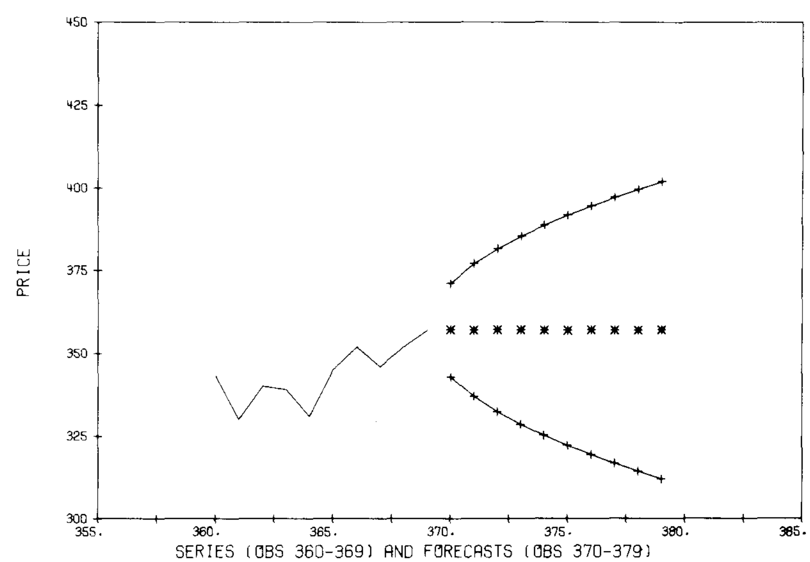

FIGURE 4.2(a) *Forecast Profile:* Series B

has exponentially decaying ψ-weights:

$$\psi_k = \phi_1^k.$$

Point forecasts of the ARIMA (1,0,0) process are thus

$$\begin{aligned} y_t(1) &= E(a_{t+1} + \phi_1 a_t + \phi_1^2 a_{t-1} + \ldots + \phi_1^k a_{t-k+1} + \ldots) \\ &= \phi_1 y_t \\ y_t(2) &= E(a_{t+2} + \phi_1 a_{t+1} + \phi_1^2 a_t + \ldots + \phi_1^k a_{t-k+2} + \ldots) \\ &= \phi_1^2 y_t \\ &\vdots \\ y_t(n) &= E(a_{t+n} + \phi_1 a_{t+n-1} + \ldots + \phi_1^n a_t + \ldots) \\ &= \phi_1^n y_t. \end{aligned}$$

Forecast variance about these point estimates is a function of the exponentially decaying ψ-weights:

$$\begin{aligned} VAR(1) &= \sigma_a^2 \\ VAR(2) &= (1 + \phi_1^2)\,\sigma_a^2 \\ &\vdots \\ VAR(n) &= (1 + \phi_1^2 + \ldots + \phi_1^{2n-2})\,\sigma_a^2. \end{aligned}$$

Confidence intervals about successive forecasts increase at a rate determined by the value of ϕ_1. When ϕ_1 is small, the increase in VAR(n) for an increase in lead time is small. When ϕ_1 is large, the increase in VAR(n) for an increase in lead time is large. In any event, it is clear that successive lead time increments produce smaller and smaller increments in forecast variance. Noting that the expression for VAR(n) is a geometric progression, forecast variance approaches a limit of

$$\lim_{n\to\infty} VAR(n) = \frac{\sigma_a^2}{1-\phi_1^2}$$

which is the variance of the y_t autoregressive process. In fact, the limit of VAR(n) as n approaches infinity will always be the variance of the ARIMA (p,d,q) $(P,D,Q)_S$ process. For the ARIMA (0,0,0) white noise process, the

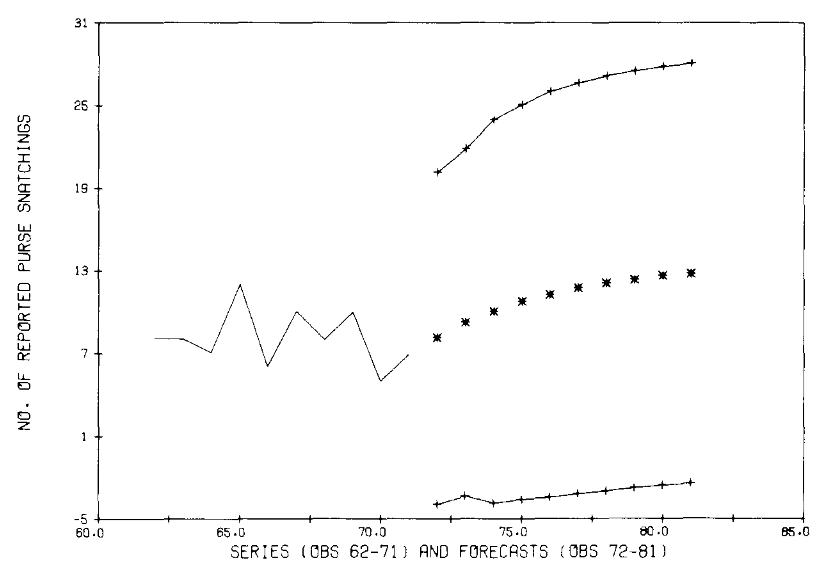

FIGURE 4.2(b) *Forecast Profile:* Hyde Park Purse Snatchings

limit of VAR(n) is σ_a^2 and for the ARIMA (0,1,0) process, VAR(n) increases without bound.

In Figure 4.2(b), we show forecasts of the Hyde Park Purse Snatchings time series. The noise component of this model is ARIMA (2,0,0) with small values of ϕ_1 and ϕ_2. The forecast profile is typical of autoregressive profiles. As lead time increases, forecasts regress to the process mean; confidence intervals about each point forecast increase with increases in lead time.

Moving Averages

An ARIMA (0,0,1) process written as

$$y_t = a_t - \Theta_1 a_{t-1}$$

has only one nonzero ψ-weight:

$$\psi_1 = -\Theta_1$$
$$\psi_2 = \psi_3 = \ldots = \psi_k = 0\,.$$

Point forecasts for the ARIMA (0,0,1) process are thus

$$y_t(1) = E(a_{t+1} - \Theta_1 a_t) = -\Theta_1 a_t$$
$$y_t(2) = E(a_{t+2} - \Theta_1 a_{t+1}) = 0$$
$$\vdots$$
$$y_t(n) = E(a_{t+n} - \Theta_1 a_{t+n-1}) = 0\,.$$

Forecast variance is determined by the single nonzero ψ-weight:

$$\mathrm{VAR}\ (1) = Ea_{t+1}^2 = \sigma_a^2$$
$$\mathrm{VAR}\ (2) = E\left[(a_{t+2} - \Theta_1 a_{t+1})^2)\right] = (1 + \Theta_1^2)\,\sigma_a^2$$
$$\vdots$$
$$\mathrm{VAR}\ (n) = E\left[(a_{t+n} - \Theta_1 a_{t+n-1})^2\right] = (1 + \Theta_1^2)\,\sigma_a^2.$$

After the second step into the future, forecast variance is constant. The limit of VAR(n) is thus

$$\lim_{n\to\infty} \mathrm{VAR}\ (n) = (1 + \Theta_1^2)\,\sigma_a^2,$$

which is the variance of the ARIMA (0,0,1) process.

In Figure 4.2(c), we show forecasts of the Swedish Harvest Index time series which we analyzed in Section 2.12.3. Here we see the distinctive

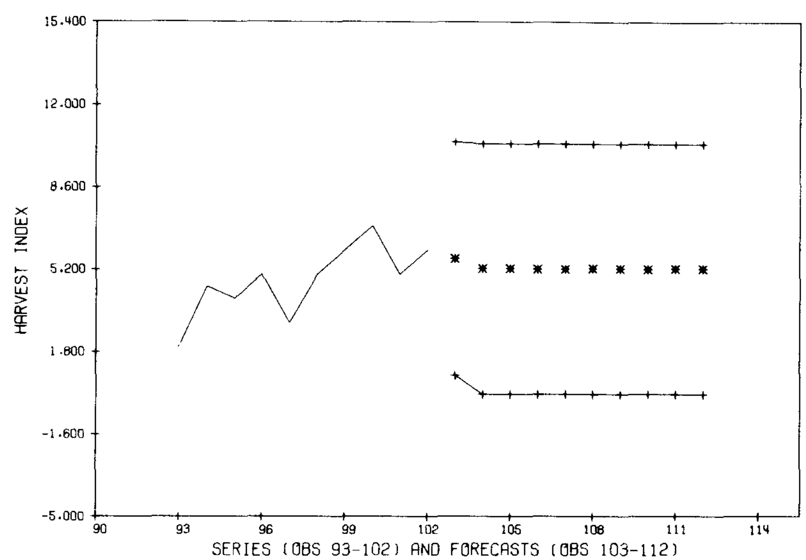

FIGURE 4.2(c) *Forecast Profile:* Swedish Harvest Index

forecast profile of a moving average process. With lead times greater than one observation, $y_t(n)$ is the process mean; VAR(n) remains constant.

As the reader may suspect by now, univariate forecasts of simple ARIMA models (as well as all univariate forecasts) tend to be statistically trivial. The best point forecast is often the process mean; and for substantial lead times, interval forecasts often approach infinity. Forecasts of complicated ARIMA seasonal models are somewhat more useful. The forecasts track the seasonal pattern of the process quite well for at least one seasonal period and thus may be used to assess "turning points" in the series. In Figure 4.2(d), we show forecasts of the Sutter County workforce series which we modeled in Section 2.12.1 as an ARIMA (0,1,1) $(0,1,1)_{12}$ process. Both moving average parameters are relatively large. The forecasts appear to track the pattern of seasonal variation quite well, although because the model is nonstationary, confidence intervals about the point forecasts grow large rapidly

4.3 Conclusion: The Uses of Forecasting

The reader who now understands how interval estimate forecasts are generated has no doubt gained a deeper insight into the nature of ARIMA

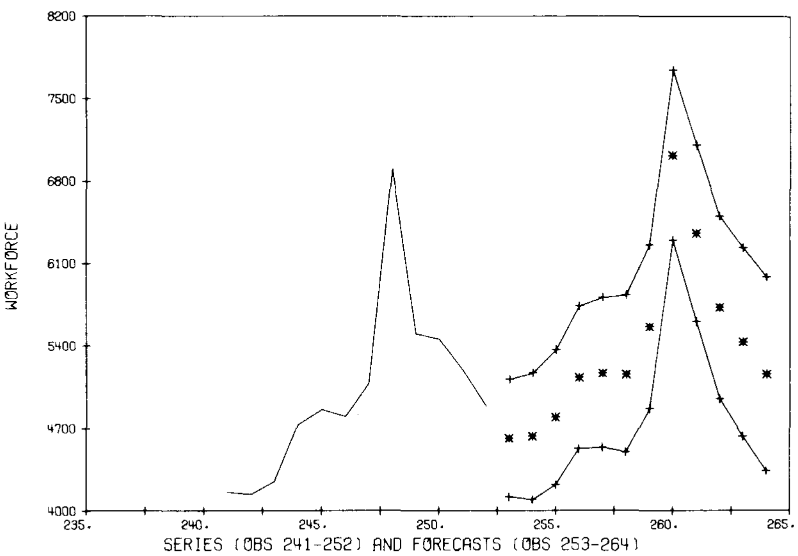

FIGURE 4.2(d) *Forecast Profile:* Sutter County Workforce

models. Beyond this not insubstantial value, however, the reader who intends to apply the principles of time series analysis to social science problems will likely see no use for univariate forecasts. In fact, however, univariate forecasts can be extremely useful as a tool of model diagnosis. We will conclude this chapter with a description of this use and with a comment on another use of forecasting which, in our opinion, is improper.

Forecasting as a Diagnostic Tool

It will often happen that two roughly identical ARIMA models produce radically different forecasts of the same time series. When this is true, the analyst will be justified in selecting the model with the "better" forecasting ability.

An example of this is seen in our analysis of the U.S. Suicide Rate time series (Section 3.5). The analysis first lead to an ARIMA (1,0,0) model for the N_t component:

$$y_t = f(I_t) + \frac{a_t}{1 - .87B}.$$

Due to the relatively large estimated value of ϕ_1, however, we respecified the N_t component as ARIMA (0,1,0):

$$Y_t = f(I_t) + \frac{a_t}{1 - B}.$$

In practice, it is almost always better to work with an ARIMA (0,1,0) model rather than an ARIMA (1,0,0) model with a large autoregressive parameter. (We will discuss the problems of estimating these parameters in Chapter 6.) There is nevertheless one situation in which the ARIMA (1,0,0) model would be preferred to the ARIMA (0,1,0) model: when the *true* noise process is autoregressive. The analyst can never know the true stochastic process underlying a time series. Yet there are many sensible procedures which the analyst can use to rule out competing models and one of these is forecasting.

While there are few major differences between ARIMA (1,0,0) and ARIMA (0,1,0) models in many contexts, this is not at all true in the context of forecasting. Using 1969 as the origin, forecasts of the two models are:

	ARIMA (1,0,0)	ARIMA (0,1,0)	ACTUAL RATE[3]
1970	11.08	11.10	11.6
1971	11.06	11.10	not available
1972	11.05	11.10	not available
1973	11.03	11.10	12.0
1974	11.02	11.10	12.1
1975	11.01	11.10	12.7
1976	11.00	11.10	12.5

The forecasts of these models are remarkably close to each other but this is expected: The two models are nearly identical. Yet forecasts of the ARIMA (1,0,0) model regress gradually to the estimated process mean (approximately 10.9 suicides per 100,000 total population) while forecasts of the ARIMA (0,1,0) model remain constant for all lead times. The actual rate does not regress to the population mean, as would be expected of an autoregressive process, but rather continues to move upward. A random walk in fact is characterized by wide swings away from the process mean. Overall, the ARIMA (0,1,0) model has a lower MSFE statistic than the ARIMA (1,0,0) model; and on this basis alone, we would select ARIMA (0,1,0) model as the "best" one for this time series.

Nevertheless, the differences between the forecasts of these two models is small and other analysts might choose the ARIMA (1,0,0) model in spite of its lesser forecasting power. In the end, model selection will hinge on a great many statistical and substantive concerns. Forecasting power is only one of these concerns and it need not be more important than any other.

Forecasting as an Impact Assessment Tool

If our experiences are typical, students of time series analysis are more fascinated with univariate forecasting than with any other application of ARIMA modeling. There seems to be a fundamental (almost spiritual) human interest in "predicting the future" which is aggravated by a course in time series analysis. The fact remains that univariate forecasts are essentially trivial and often disappointing. One can predict the future only for short lead times and in limited contexts.

Students of time series analysis often "discover" a means of assessing impacts with univariate forecasts. These methods are actually quite old and widely used in industrial engineering and quality control applications. While these methods are valid and useful in manufacturing contexts, however, they are not generally suited to social science problems and data.

As an illustration of the use of these methods in quality control engineering, consider a machine that manufactures ball bearings. Each ball bearing will vary slightly in diameter and this variance from ball bearing to ball bearing is a stochastic (time series) process. An ARIMA model of the process is used to set forecast confidence intervals around the process realization. When the manufacturing process is "in control," 95% of the ball bearings produced by the machine will lie within these confidence intervals. If a run of three or four ball bearings are observed to lie outside the confidence intervals, the quality control engineer infers that the process is "out of control." The machine is turned off and repairs are made.

The strong forecast-based inference here is possible because of certain given characteristics of the manufacturing process. One, the quality control engineer knows a priori what a "bad" ball bearing is. A "bad" ball bearing is one that is a few thousandths of an inch too small or too large; one that cannot be sold to customers. Two, process realizations are relatively long. The quality control engineer may have one thousand or more observations available for building an ARIMA model of the process. Three, process inputs are tightly controlled and known. White noise inputs to the manufacturing process arise from relatively small variations in the quality of raw materials (water, oil, electricity, steel, and so on). The quality control engineer knows not only the sources of white noise inputs but also their relative magnitudes.

These characteristics of the manufacturing process are not ordinarily seen in social processes. There is no definition of a "high" or "bad" unemployment rate, for example; social science time series realizations are ordinarily short; process inputs are unknown and erratic and there are quite often seasonal inputs which are unheard of in manufacturing processes. Forecasts of social processes are thus less certain and more prone to error than forecasts of industrial processes.

In Chapter 3, we developed impact assessment models from a foundation of scientific validity. Certain "threats to validity" are controlled through design while others are controlled through analysis. This distinction is not always clear and this is particularly true when alternative patterns of impact are compared and ruled out. When an intervention component has been misspecified, Type I or Type II decision errors are a likely result. To control this threat to validity, we have recommended a conservative strategy of model building. The strategy leads generally to a more confident statement of impact but it does so at a real cost: A relatively long postintervention time series segment is required.

Impact assessments based on forecasting do not require this great cost. Deutsch (1978) has recently proposed a variation of the time series quasi-experiment which allows for an impact assessment within a few weeks or months of the intervention. At the simplest level, Deutsch proposes to build an ARIMA forecasting model from the preintervention time series. Postintervention forecasts of the model are then compared with the actual observations. If the postintervention observations fall outside the forecast confidence intervals, Deutsch concludes that the social system has gone out of control or that the intervention has had an impact on the time series.

If "early detection methods" should prove generally reliable, Deutsch's work will represent an important advance in social science methodology. In the first published use of these methods, however, Deutsch and Alt (1977) found a statistically significant drop in gun-related crime after introduction of a strict gun-control law. In our reanalysis of those data (Sections 2.12.2 and 3.6; see also, Hay and McCleary, 1979), we found no evidence of the effect claimed by Deutsch and Alt. We attribute this difference in findings to, among other things, weaknesses of the early detection methods used by Deutsch and Alt.

A major deficiency of forecast-based impact assessments is that confidence intervals about each point estimate are subject to error. As noted in Section 4.1, the ψ-weight models used to set confidence limits about each point estimate require that the analyst know the *true* ARIMA structure of the time series process. Yet in practice, this is never the case.

Compared to industrial process time series, social science time series are relatively short. A weak seasonal component, for example, may go undetected or may have statistically insignificant parameter estimates for a time series of only 100 observations. When a few more observations are added to the time series, however, the seasonal component may suddenly assume statistical significance. Nonstationary processes present an analogous problem. With a weak trend, estimates of Θ_0 are likely to be statistically insignificant unless the time series is relatively long.

When a run of postintervention observations fall outside the forecast confidence intervals, there are always two equally plausible explanations: the process may have been impacted by a social intervention and/or the forecast confidence intervals may have been underestimated. With a relatively short time series, the latter explanation is always more plausible than the former.

But a greater problem with forecast-based impact assessments is that threats to validity cannot be controlled. A confident statement of impact requires not only a statement as to whether an impact occurred or not (which forecast-based assessment may or may not adequately give) but also a statement as to the nature of the impact. In the first place, abrupt temporary impacts will almost always "fool" a forecast-based impact assessment.[4] As an exercise, the reader may wish to try a forecast-based assessment for the Sutter County Workforce time series. In Section 3.2.2, we demonstrated that the Sutter County flood had only a substantively and statistically insignificant impact on this time series. Yet if the preintervention series is used to forecast the postintervention observations, the analyst will arrive at a radically different conclusion. The "reactive intervention" threat to internal validity will also "fool" a forecast-based impact assessment. For most ARIMA models, the last observation of the series has the greatest weight in determining forecasts, and if the last preintervention observation is an extreme value, forecast-based impact assessments will indicate a statistically and substantively significant effect.

To guard against these threats, impact assessment requires a relatively long postintervention time series segment. There is a fundamental difference between *detecting* and *modeling* an impact. Even in the more "applied" social sciences (evaluation research and policy analysis, for example), impact assessment must be concerned with the dynamic structure of social change. This concern can be addressed only from a foundation of scientific validity and from the conservative impact model-building strategy we outlined in Chapter 3.

For Further Reading

Nelson (1973: Chapter 6–8) develops univariate ARIMA forecasting at an introductory level. Granger and Newbold (1977: Chapters 4–5) develop

the same material at a slightly higher level. More comprehensive treatments of forecasting which consider non-ARIMA methods as well are given by Pindyck and Rubinfeld (1976) and Makridakis and Wheelwright (1978). Makridakis-Wheelwright is written for graduate students in business while Pindyck-Rubinfeld is written for graduate students in economics. While both works are outstanding, there is a clear difference in the levels of sophistication assumed of the reader. Granger and Newbold (1977: Chapter 8) compare the performance of a variety of forecasting methods. This work is absolutely essential for any reader who plans to do forecasting. Finally, Vigderhous (1978) or Land and Felson (1976) are excellent examples of forecasting in a social science context.

NOTES TO CHAPTER 4

1. The proof of this claim is obvious when one considers that the *conditional* expectation uses *all* of the available information about the process. See Pindyck and Rubinfeld (1976: 498–499) for a formal proof.

2. The assumption is that ψ_k is zero (or some infinitesimally small number) is satisfied whenever the time series is long, say 50 observations or more, or whenever the low-order ψ-weights are so small that the sequence converges to zero within a few weights. When this assumption is *not* satisfied, the minimum MSFE forecasts are generated by backcasting the series to obtain estimates of $y_0, y_{-1}, \ldots, y_{-\infty}$. For relatively long time series, of course, the conditional expectation of y_{t+n} is the same whether the expected values or the backcasted values of distant random shocks are used. See Box and Jenkins (1976: 199–200) for a detailed description of backcasting.

3. The values of this time series are taken from the U.S. Department of Commerce publication *Historical Statistics of the United States: Colonial Times to 1970*. The values after 1970 are taken from the *1978 Statistical Abstract of the United States*. The rates for 1971 and 1972 are not given in that volume.

4. Hay and McCleary (1979: 309–310) show that one of the time series analyzed by Deutsch and Alt has an abrupt, temporary impact effect. Using the early detection method, however, Deutsch and Alt conclude that the effect is a permanent reduction in gun-related crime.