2007

A Dynamic-Trend Exponential Smoothing Model

Don Miller
Virginia Commonwealth University

Dan Williams
CUNY Bernard M Baruch College

A Dynamic-Trend Exponential Smoothing Model

Don Miller and Dan Williams

A Dynamic-Trend Exponential Smoothing Model

Abstract

Forecasters often encounter situations in which the local pattern of a time series is not expected to persist over the forecasting horizon. Since exponential smoothing models emphasize recent behavior, their forecasts may not be appropriate over longer horizons. In this paper, we develop a new model in which the local trend line projected by exponential smoothing converges asymptotically to an assumed future long-run trend line, which might be an extension of a historical long-run trend line. The rapidity of convergence is governed by a parameter. A familiar example is an economic series exhibiting persistent long-run trend with cyclic variation. This new model is also useful in applying judgmental adjustments to a statistical forecast. For example, this new model can converge an exponential smoothing forecast to a judgment-imposed future trend line that represents – say – a 10% increase over the extrapolated trend. The accuracy of this new method will be compared (later – haven't done this yet) to that of existing methods in forecasting a sample of cyclical series with long-run trends.

1. Introduction

Forecasters often encounter situations in which the local pattern of a time series is not expected to persist over the forecasting horizon. Thus, the forecast of an exponential smoothing model may not be appropriate over a longer forecasting horizon. This problem arises in a variety of contexts, including: (1) forecasting series exhibiting persistent long-run trend with cyclic variation; (2) judgmental adjustment of a statistical forecast; and (3) forecasting a series that has experienced a recent disruption of its underlying pattern. In this paper, we present a new exponential smoothing model called *dynamic-trend* that is useful in each of these situations.

Long-run trends with cyclic variation

The presence of cycles may create a short-term (local) trend that differs from the long-run trend. Holt's method of linear exponential smoothing is ill-equipped for such a situation because it tends to misinterpret cyclic variation as a local trend. It may forecast accurately for the short term but can miss badly as the local trend is extended over the forecasting horizon. Damped-trend exponential smoothing (Gardner and McKenzie, 1985) is ineffective in this situation because it damps the local trend asymptotically to zero. Snyder (2006) developed an "augmented damped-trend" model in which the local trend converges over time to a fixed long-run value. This helps, but it is insufficient because its forecast does not return to the level of the long-run trend line.

Several studies have provided ways to reconcile local and long-run forecasts. Carbone and Makridakis (1986) developed two models, one short-term and one long-run, along with a formula for reconciling them at various forecasting horizons. Armstrong and Collopy (1992) used a similar approach within rule-based forecasting. The latter paper also incorporated domain knowledge regarding "causal forces." Both achieved improvements in accuracy compared to conventional methods, but they are relatively complex.

Judgmental adjustments

Forecasting accuracy can often be improved via judgmental adjustment of a statistical forecast (Sanders, 2005). When this occurs, using a dynamic-trend model can simplify the ongoing forecasting process considerably. Judgmental adjustments are typically applied after a statistical model has produced a forecast, i.e., applied outside the model. This can lead to modeling problems down the road. Once the anticipated change is in the data, the statistical model may not be useful until it can adapt to the new pattern. The statistical model may also be compromised for tracking purposes: If the future changes (predicted judgmentally) occur as expected, the model itself will produce large errors. If change does *not* occurr as expected, the model may produce small errors.

With the dynamic-trend model, a parameter controls the rate of transition from the extrapolation forecast to the judgmentally-adjusted forecast. Thus, dynamic-trend incorporates judgmental adjustments *within* the model. It remains viable for updating and tracking as new data become available, whether the change occurs or not. (Level-adjusted exponential smoothing, Miller and Williams, 1999, accomplishes this when a *future* level shift is expected.)

To be effective, judgmental adjustments should be based on information about the future that is not contained in the existing data (Goodwin, 2005). Using a dynamic-trend model promotes this. Since judgmental adjustments are incorporated within the model, some level of interaction between the manager and the forecaster is required. The manager is likely to feel a need to explain the adjustment to the forecaster, thereby reducing the likelihood that adjustments are based something other than outside-the-data information about future events.

Pattern breaks in recent data

When there is a pattern break in the most recent data, several very different future patterns may seem plausible, ranging for example from never returning to the historical pattern to a rapid return. The dynamic-trend model supports the forecaster by providing a way to depict very different possible futures via manipulation of the parameters of a single statistical model.

In section 2, we formulate the model. In section 3, we provide examples of using the model for cyclic series with long-run trends, when judgmental adjustments are made, and when pattern breaks occur near the end of a series. We compare the model's accuracy compared to conventional models for selected series from the M3-competiton (Makridakis and Hibon, 2000). We summarize results and offer conclusions in Section 4.

2. Model Formulations

The dynamic-trend model can be based on any extrapolation model that projects a local pattern. Here we develop its application to Holt's linear exponential smoothing model. We begin with the standard formulation of Holt's model.

Holt's linear exponential smoothing

$$L_t = \alpha X_t + (1 - \alpha)(L_{t-1} + B_{t-1}), \quad (1)$$

$$= F_{t-1}(1) + \alpha e_t \quad (2)$$

$$B_t = \beta (L_t - L_{t-1}) + (1 - \beta)(B_{t-1}), \quad (3)$$

$$= B_{t-1} + \alpha\beta e_t \quad (4)$$

$$F_t(m) = L_t + mB_t \quad (5)$$

L_t is the local level of the series; B_t Is the local trend; $F_t(m)$ is the forecast at origin t for m periods ahead; and $e_t = X_t - F_{t-1}(1)$, the one-step-ahead forecasting error at period t. $F_t(m)$ is a projection of the local trend line, determined at t, over the next m periods. The smoothing parameters for level and trend, α and β, are usually restricted to the range (0, 1). Equations (2) and (4) are simpler, error-correction forms of (1) and (3). The Holt model becomes simple exponential smoothing if both β and the initial trend B_1 are set to 0.

The dynamic-trend model

The dynamic-trend forecast starts with a short-term forecast such as that of Holt's model and transitions asymptotically to a long-run (future) trend line.

Let $L^*_t = A^* + B^* t$ represent the long-run trend line at period t.

The dynamic-trend model blends the Holt model and the basic trend:

$$L_t = \alpha X_t + (1-\alpha)[L_{t-1} + (1-\phi_1)(L^*_{t-1} - L_{t-1})] \quad + \quad [B_{t-1} + (1-\phi_2)(B^* - B_{t-1})], \tag{6}$$

$$= F_{t-1}(1) + \alpha e_t \tag{7}$$

$$B_t = \beta\{L_t - [L_{t-1} + (1-\phi_1)(L^*_{t-1} - L_{t-1})]\} + (1-\beta)[B_{t-1} + (1-\phi_2)(B^* - B_{t-1})], \tag{8}$$

$$= \phi_2 B_{t-1} + (1-\phi_2) B^* + \alpha\beta e_t \tag{9}$$

$$F_t(1) = [\phi_1 L_t + (1-\phi_1) L^*_t] \quad + \quad [\phi_2 B_{t-1} + (1-\phi_2) B^*] \tag{10}$$

$$F_t(m) = [\phi_1^m L_t + (1-\phi_1^m) L^*_t] + [\sum [\phi_2^i B_{t-1} + (1-\phi_2^i) B^*] \tag{11}$$

$$= F_t(m-1) + [\phi_1^{m-1}(1-\phi_1)(L^*_t - L_t)] + [\phi_2^m B_t + (1-\phi_2^m) B^*] \tag{12}$$

L_t and B_t are the level and trend at period t. L_t^* is the level of the long-run trend line at t, and B^* is the slope of the long-run trend line. $F_t(m)$ is the forecast, determined at t, for the next m periods. The parameters α and β smooth the level and trend, as in the Holt model. The parameter ϕ_1 governs the rate at which the Holt level transitions to the level of the basic trend line, and ϕ_2 governs the rate at which the Holt trend transitions to the slope of the basic trend line. Both ϕ_1 and ϕ_2 are restricted to the range (0, 1). The nearer ϕ_1 and ϕ_2 are to 0, the more rapid the transition.

In (6), $(1 - \phi_1)(L^*_{t-1} - L_{t-1})$ is the amount by which the level is predicted to transition at period t toward the level of basic trend line. Similarly, in (6), $(1 - \phi_2)(B^* - B_{t-1})$ is the amount by which the trend is predicted to transition at period t toward the slope of basic trend line. Equations (7) and (9) are simpler, error-correction forms of (6) and (8). Equation (12) is an equivalent form of (11) that may be more convenient for computation within a spreadsheet. In (12), $[\ \phi_1^{m-1}\ (1 - \phi_1)\ (L^*_t - L_t)\ ]$ is the amount by which the level is predicted to transition toward the long-run trend line from period t+m-1 to t+m, and $[\ \phi_2^m\ B_t\ +\ (1 - \phi_2^m)\ B^*\ ]$ is the predicted trend in period t + m.

Expression (11) can be expressed in closed form as follows:

$$F_t(m) = [\ \phi_1^m\ L_t\ +\ (1 - \phi_1^m)\ L^*_t\ ]\ +\ mB\ -\ [\ \phi_2\ (1 - \phi_2^m)\ /\ (1 - \phi_2)\ ]\ (T_t - B^*)\ ], \quad \text{if } \phi_2 < 1 \qquad (13)$$

$$= \phi_1^m\ L_t\ +\ (1 - \phi_1^m)\ L^*_t\ +\ mB_t \quad \text{if } \phi_2 = 1 \qquad (14)$$

This model contains many conventional models as special cases, including simple and linear exponential smoothing, damped-trend exponential smoothing (Gardner and McKenzie, 1985), Snyder's augmented damped-trend model (Snyder, 2006), and the Theta model (Assimakopoulos and Nikolopoulos, 2000; Hyndman and Billah, 2003). Each of these models results from specific settings of ϕ_1 and ϕ_2, as follows:

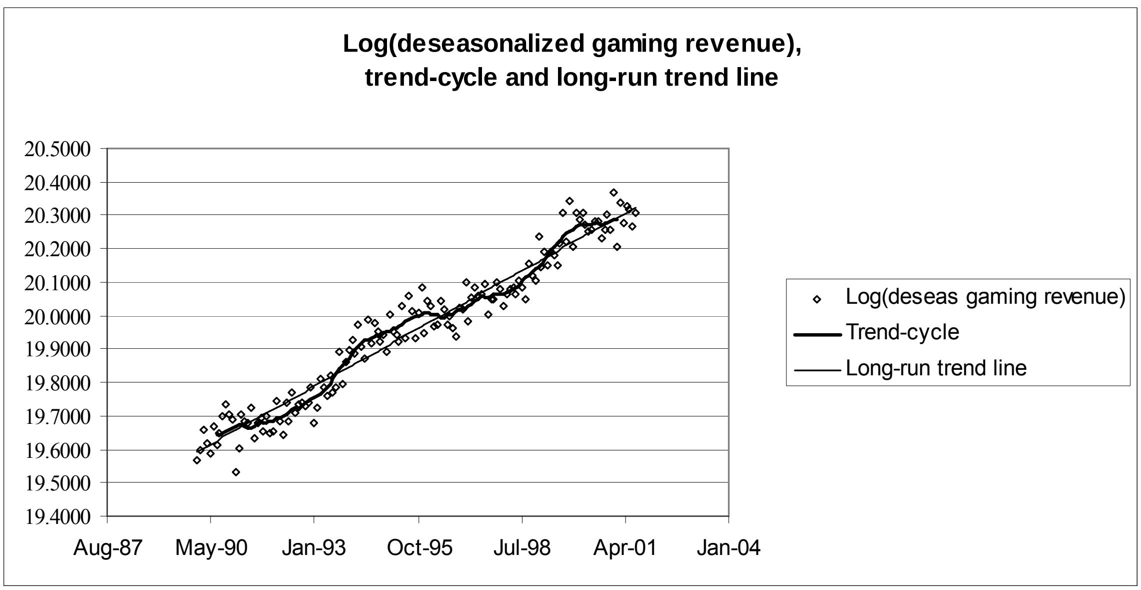
1. Holt's linear exponential smoothing: Set $\phi_1 = 1$ and $\phi_2 = 1$.
2. Damped-trend: Set $\phi_1 = 1$ and the basic trend $B^* = 0$. The damping parameter is ϕ_2.
3. Snyder's augmented damped-trend: Set $\phi_1 = 1$.
4. Theta model: Set $\phi_1 = 1$ and $\phi_2 = 1$ (producing Holt's model). Set the initial trend B_1 to ½ B^* (where B^* = the slope of the fitted trend line through the original series.) Set $\beta = 0$. (Thus, the trend remains constant at ½ B^*.)

3. Examples and Evaluation

Long-run trend with cyclic variation

Consider the use of the dynamic-trend model for forecasting monthly gaming revenues for Clark County, Nevada (Las Vegas)[1]. Gaming revenues grew steadily on a percentage basis from January 1990 to August 2001 (when growth was disrupted by the events of 9-11), a period of almost 11 years. Figure 1 is a plot of the logarithm of gaming revenues (seasonally adjusted), which exhibit a linearly increasing trend with cyclic variation, along with the the long-run trend line determined by least-squares fit to the data and the estimated trend-cycle.

Figure 1



We fit Holt, damped-trend, augmented damped-trend, and dynamic-trend models to the entire series through August 2001 (n = 140). (We stopped here because the events of 9/11 disrupted the series.) Parameter values and initial conditions were determined in the following way: For Holt, we optimized α and β by choosing the values that minimized

[1] We express our appreciation to Dr. Keith Schwer, former Director of The Center for Business and Economic Research at the University of Nevada, Las Vegas for providing these data.

within-sample 1-month-ahead root-mean-square error (RMSE). For damped-trend, we used the same values for α and β, then chose the optimal value of ϕ. For augmented damped-trend, the long-run trend was determine by least-squares fit to the in-sample data. As with damped-trend, we used the Holt values for α and β, then optimized ϕ. For the dynamic-trend model, we used the same long-run trend line and the same values for α and β, then chose the optimal values of ϕ_1 and ϕ_2. We then repeated the exercise in the same way, except that we optimized the parameters by minimizing 12-month-ahead RMSE.

Table 2 provides model details and the results of fitting. Dynamic-trend had the smallest RMSE and Holt the largest, but the differences are not great. For 12-month-ahead comparisons, dynamic-trend is clearly the best fit (RMSE = .0520), followed by augmented-damped trend (RMSE = .0606), then Holt (RMSE = .0630).

Table 2

Quantity	Holt	Augmented damped trend	Dynamic-trend
Initial level (L_1)	19.594	19.594	19.594
Initial trend (B_1)	.00526	.00526	.00526
α	.217	.217	.217
β	.010	.010	.010
ϕ_1	*	*	.911
ϕ_2 (this is "ϕ" for "augmented")	*	.754	.832
Long-run trend line, intercept A*	*	*	19.586
Long-run trend line, slope B*	*	.00526	.00526
In-sample RMSE (1-ahead)	.0475	.0473	.0465
In-sample RMSE (12-ahead)	.0630	.0606	.0520

* = not applicable to this method

A good fit does not necessarily lead to good forecasting. We developed five out-of-sample 12-month- ahead forecasts using Holt, damped-trend, augmented damped-trend, and dynamic-trend models. These forecasts were produced every 18 months starting in December 1993. Each time the forecasts were updated, we re-optimized the parameters and re-estimated the long-run trend line. The initial values for level and trend were developed from the parameters of the long-run trend line ($L_1 = A^* + B^*$, $B_1 = B^*$).

We found that relative model performances depended on (1) the accuracy of the estimated long-run trend, and (2) the recent data pattern leading up to the forecasting period. Figures 2, 3, and 4 illustrate how these factors affected model accuracy for three of the five forecasts. Damped-trend is not included in the figures as it fares poorly – which is not surprising since it isn't designed for series with sustained trend.

- December 1993 (Figure 2): The dynamic-trend forecast is the least accurate, and the Holt forecast is most accurate. The reason for dynamic-trend's poor performance is that the estimated long-run trend based on the first four years' data is inaccurate. Subsequent data revels that December 1993 marks the beginning of a return to the long-run trend from the bottom of a down cycle.
- June 1995 (Figure 3): The local level and trend are close to the level and trend of the long-run trend line. Thus, the three forecasts are similar and there is little difference in model performance.
- December 1996 (Figure 4): The Holt model is strongly affected by a temporary downturn in the data, and its forecast is wildly inaccurate. The dynamic-trend forecast is the most accurate.

Figure 2

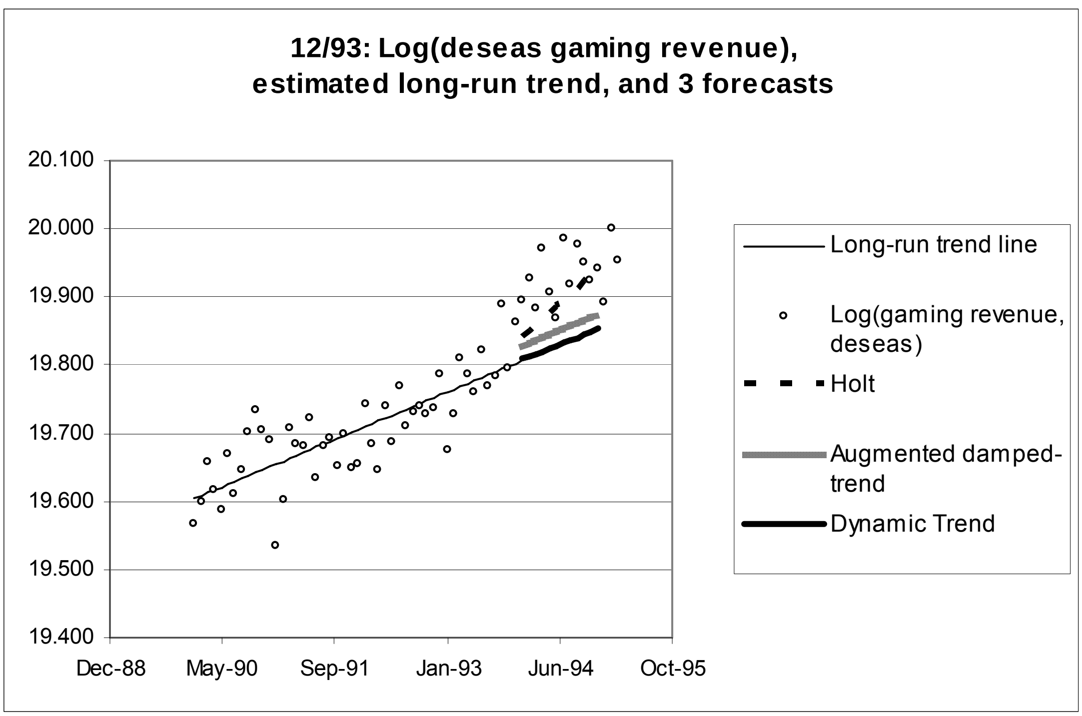


Figure 3

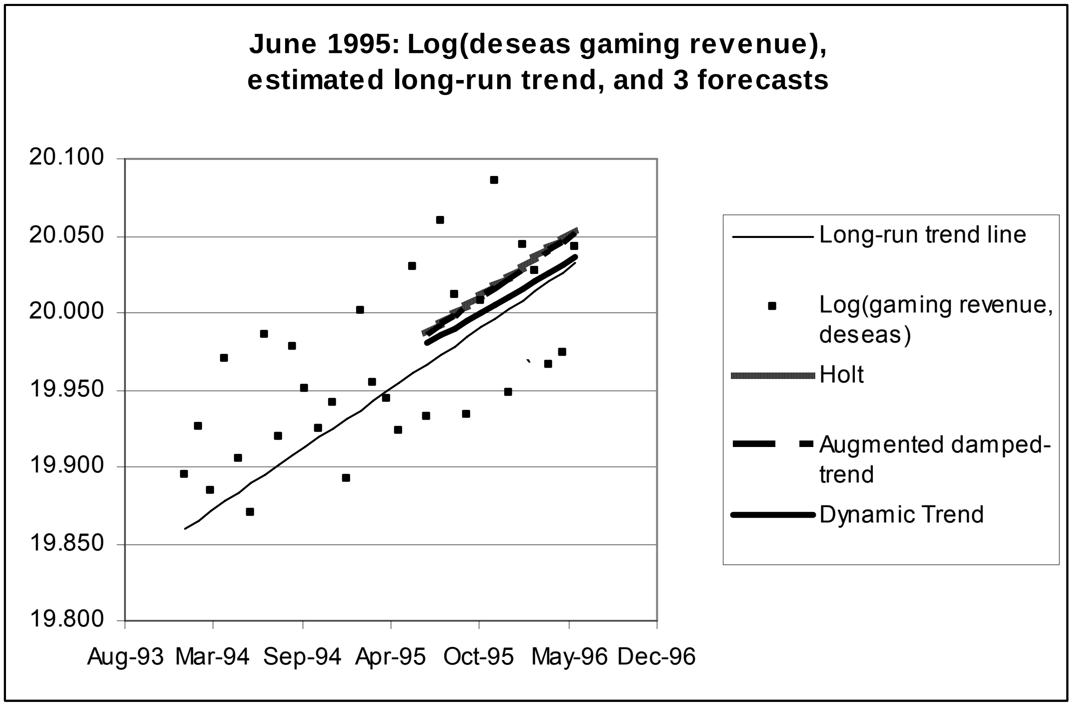


Figure 4

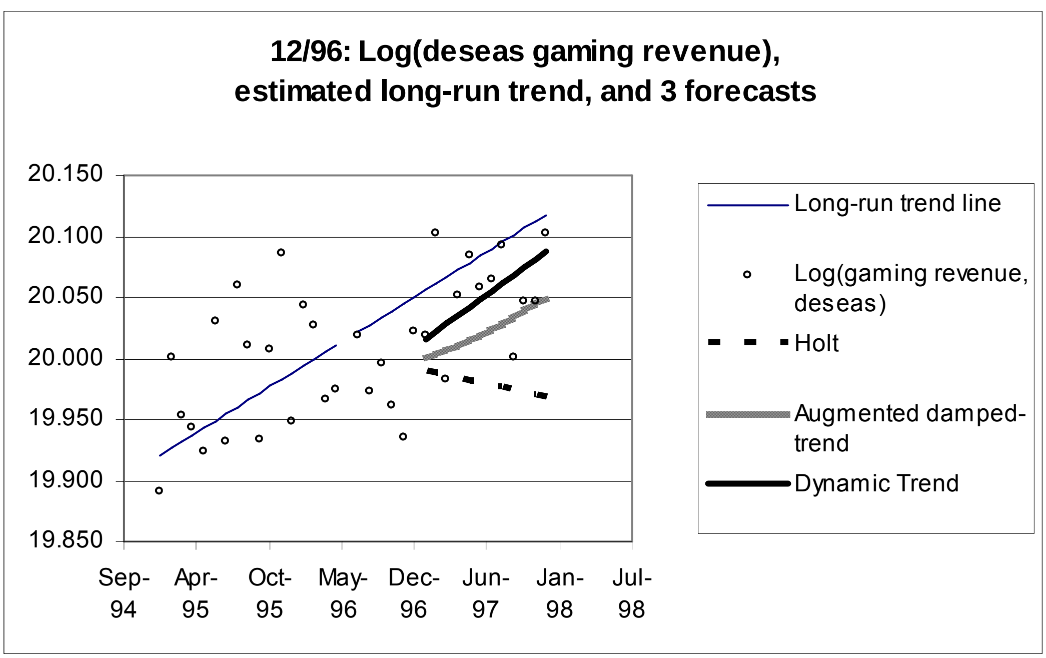


Table 3 provides the out-of-sample mean absolute percentage error (MAPE) for 12 month forecasts for all five forecasts. Dynamic-trend was the most accurate in all cases except the 12/93 forecast, for which the data were insufficient to estimate the long-run trend accurately.

	MAPE (Out-of sample, over 12 months)			**Estimated long-run trend line**	
	Holt	Augmented damped-trend	Dynamic-trend	intercept	slope
12-93**	**0.212**	0.386	0.496**	19.59977	0.00422
Jun-95	0.224	0.219	**0.214**	19.56824	0.00595
Dec-96	0.38	0.202	**0.158**	19.57718	0.00563
Jun-98	0.192	0.183	**0.145**	19.59031	0.00522
Dec-99	0.165	0.146	**0.08**	19.58855	0.00526

** Data (3 years) were insufficient to estimate the long-run trend accurately

Empirical evaluation of model performance

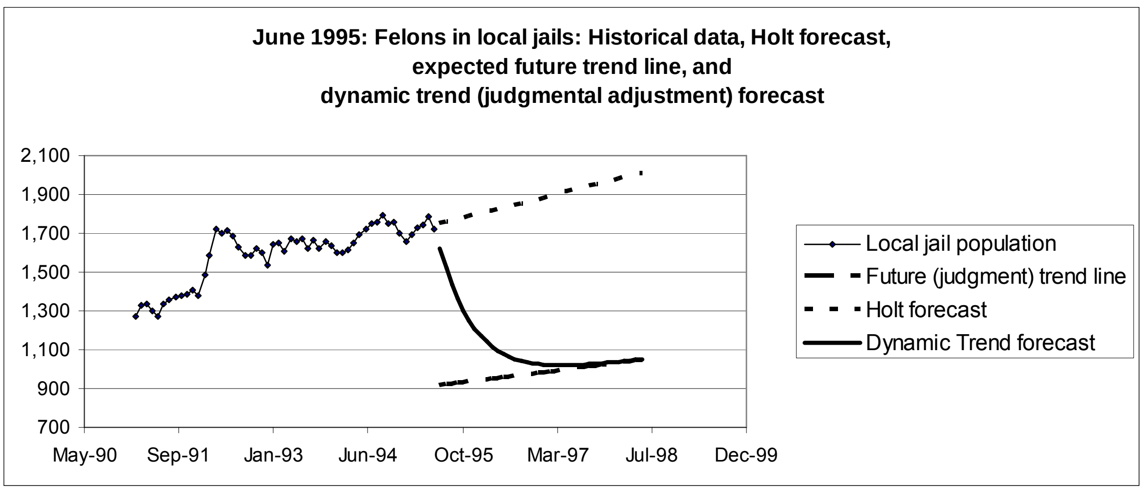
(Here we will describe the results of forecasting withheld data for lots of series with long-run trend and cyclic variation,probably taken from the M- and/or M3-competitions.)

Judgmental adjustment of statistical forecasts

In 1995, the Virginia General Assembly approved the construction of a number of new prisons in conjunction with mandated sentences and elimination of parole. Because of insufficient space in prisons, felons were also being housed in local jails, creating overcrowded conditions there as well. This number had been increasing steadily for a number of years. When new prisons were completed, the number of felons housed in local jails was expected to decline from about 1,740 to 1,000 over two years, at which point it would start increasing at one-half its current trend. Figure 4 plots the monthly local prison population through June 1995, along with the Holt forecast, the future trend line that was expected once the decline was complete, and the dynamic-trend forecast that represents the judgmental adjustment of the Holt forecast.

Since the judgmental adjustment is based on information not contained in the data, the model used to produce the pre-judgment, statistical forecast does not have to include a dynamic-trend component. For this example, the pre-adjustment forecast developed at June 1995 was produced by fitting a Holt model [$\alpha = .7$, $\beta = .01$ in expressions (1) – (4)] to the historical data through June 1995. The expected future trend line was the line with level = 1,000 at June 1997 and slope = 3.73 per month (½ the Holt trend value at June 1995 of 7.45 per month). The dynamic-trend forecast used the parameter values (ϕ_1 = .892 and ϕ_2 = .5) in expressions (8) and (9) which produced a transition from the Holt forecast to the expected future trend line (approximately) in June 1997.

Figure 4



How did this forecast work out? Quite well, as shown in Figure 5. The forecast errors over the 24-month horizon are relatively small, and there is no need to reconsider plans. The greatest benefit of incorporating judgment within the model is forecast management, as the model remains viable for tracking and updating. Figure 6 shows an easily produced updated forecast using data through February 1997.

Figure 5

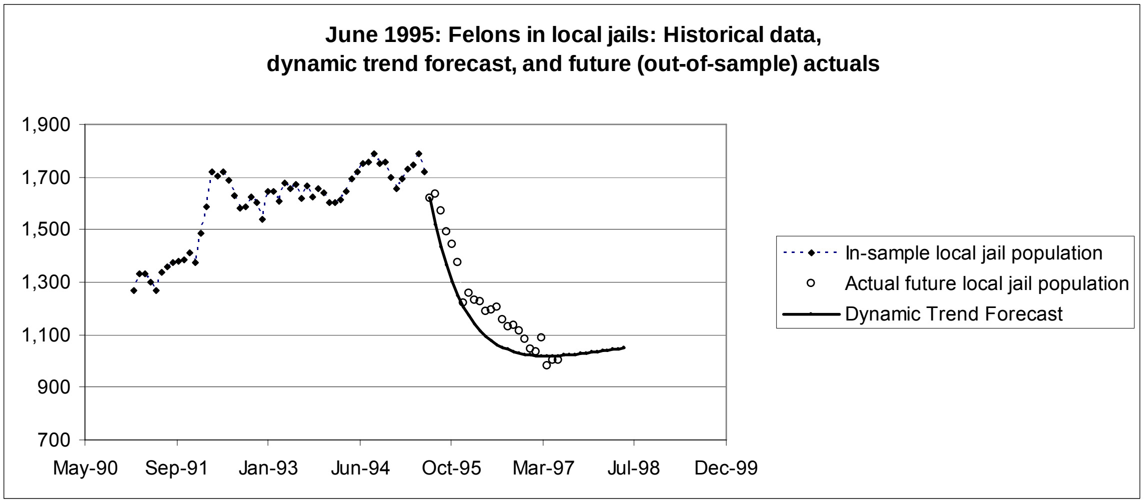
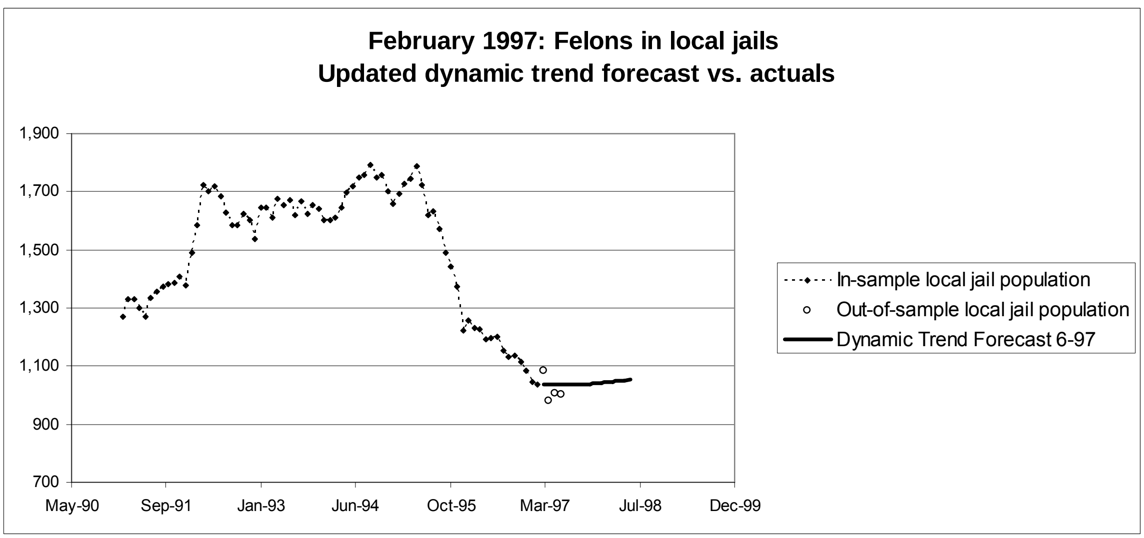


Figure 6



Disruptions of the pattern in the recent data

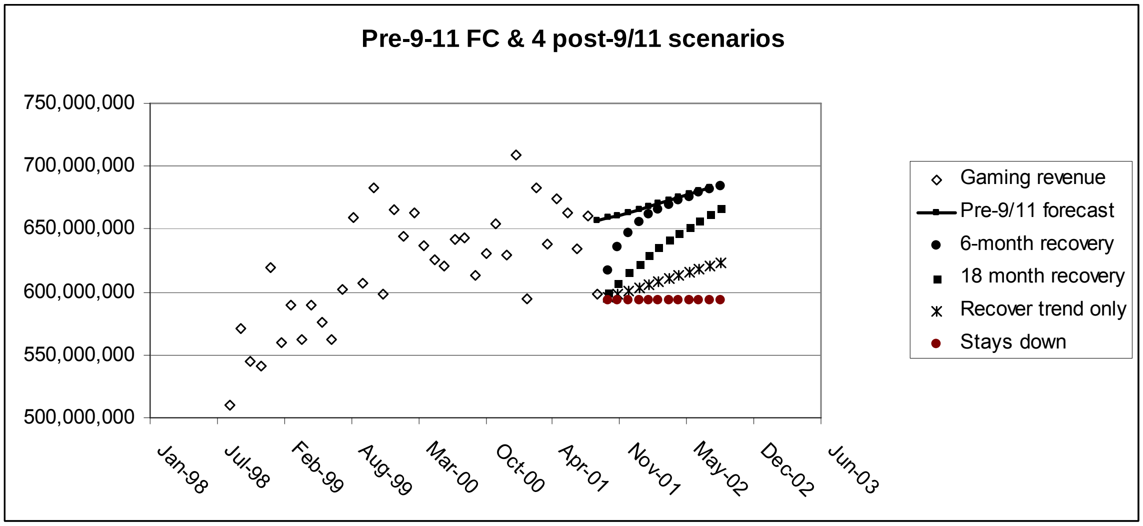
Now we return to the Clark County, Nevada (Las Vegas) gaming series. The events of 9/11 caused gaming revenues to decline precipitously. In the immediate aftermath of 9/11, most construction was suspended, services were cut, shows went dark, free entertainment was curtailed, and thousands of employees were furloughed or laid-off. Casino and hotel management had to decide how long to put plans on hold, or whether these cuts should be permanent, even greater cuts were required, or wholly new strategies were needed for a post 9/11 environment (Las Vegas Review-Herald, September 11, 2004). In the wake of 9/11, the best decision alternative depended on the nature and speed of recovery.

Figure 3 shows actual gaming revenues up to and including the September 2001 drop along with the pre-9/11 forecast developed with data through August 2001 using the dynamic-trend model ($\alpha = .235$; $\beta = .021$; $\phi_1 = .92$ and $\phi_2 = .78$). It also shows four different "recovery" paths, each of which might require different decisions. These alternative futures start from the lower, post-9/11 level, so we introduced a one-period decline of $72,000,000 (estimated judgmentally) in September 2001 via level-adjusted exponential smoothing (Williams and Miller, 1999) within the dynamic-trend model. Then we updated the forecast four times, with each new forecast representing a different recovery path. The four forecasts used the same values of α and β as the pre-9/11 model and were differentiated by the values used for ϕ_1 and ϕ_2.

In one path, revenues return to the pre-9/11 forecast within about 6 months. This path is the forecast that results from setting $\phi_1 = .6$ and $\phi_2 = .6$ in the dynamic-trend model and defining the future long-run trend line to be the pre-9/11 forecast. A second path has revenue returning to the pre-9/11 forecast more slowly, over about 18 months. This path was achieved by setting $\phi_1 = .9$ and $\phi_2 = .9$. In a third possible scenario, revenues immediately resume the trend of the pre-9/11 forecast but never recover to its level. Here, we set $\phi_1 = 1.0$ and $\phi_2 = 0$. The fourth scenario has revenue staying at the September 2001 level for the foreseeable future. For this path, we set $\phi_1 = 1.0$ and $\phi_2 = 0$, and we set the future long-run trend value to zero for all future periods. Simply by manipulating the parameters, the model can be adjusted to produce the scenarios that are meaningful to planners.

In the immediate aftermath of a disruption, each of the scenarios may seem plausible. As new, post-disruption data become available, plans can be firmed up as one or two scenarios begin to emerge. If additional scenarios of interest emerge, they can be added by further manipulation of the parameters. Custer and Miller (2007) provide a procedure for such analysis.

Figure 3



Summary and Conclusions

(later)

References

Sanders, Nada. Introduction to "When and How Should Statistical Forecasts Be Judgmentally Adjusted?," Foresight, 1, 2005

Goodwin, 2005

Gardner and McKenzie, 1985

Snyder (2006)

Carbone and Makridakis (1986)

Armstrong and Collopy (1992)

Miller and Williams, 1999

Makridakis and Hibon, 2000

Gardner and McKenzie, 1985

Snyder, 2006

Assimakopoulos and Nikolopoulos, 2000

Hyndman and Billah, 2003

Custer and Miller (2007)