FORECASTING FROM TIME SERIES MODELS.

This is an unpublished work which was part of the project by C. Granger, A. Espasa and A. Pérez-Espartero to update the book Forecasting in Business and Economics*.

3. FORECASTING FROM TIME SERIES MODELS.

**The project was abandoned in 2009*

3.1 UNCERTAINTY IN TRENDS. INTEGRATED MODELS OF ORDER I (d, m).

In the previous chapter it has been discussed that trends refer to the acyclical long-term evolution of a time series, that trends are caused mainly by changes in technology and changes in attitudes or in the structure of the society and that trends evolve smoothly. Because of this last property we could consider as a first approximation deterministic schemes and in particular polynomial trends of low order, usually of order one:

$$x_t = a + b\ t + w_t. \tag{3.1}$$

With polynomial trends like (3.1) the w_t shocks appearing at anytime in the system affect to x_t but not to the trend. In fact, model (3.1) implies that there is none uncertain future event which could affect the trend.

The absence of uncertainty in trends seems unrealistic . In fact factors causing trends evolve smoothly but with uncertainty and this uncertainty must also be incorporated in models representing trends. In chapter 2 this has been done that by considering segmented trends. Conditional to the fact the segmentation points are known, as it was assumed in the previous chapter, these segmented trends are deterministic. But if for a particular time series, segmentations have taken place during the past it can be assumed that with probability one they will also happen in the future. About those forthcoming segmentations there is high uncertainty because it is not known when they will occur and which will be their magnitudes. Therefore, segmented trends imply that trends are not deterministic but stochastic. At certain points, t_1, t_2, …, special stochastic shocks appear and the structure of the trend changes. Conditional to the knowledge of these data points models as (2.33) can be formulated and in them the segmented trend structure for the past is known without error resulting a deterministic trend for the past but not for the future. Because of that it was commented in chapter 2 the problems of forecasting with segmented trends. In particular it was seen that these models do not adapt themselves to new breaks which could occur after the sample period and if a new break happens the model will be wrong for ever unless an analyst intervenes and reformulates the model accordingly with the new situation.

It can be concluded that T (d^s) models are more realistic than T (d) ones. But the segmented trend models which have been considered do not include a stochastic formulation of the occasional breaks and therefore of the trends and consequently one is obliged to proceed conditional to the knowledge of the time at which breaks take place. Some authors have tried stochastic formulations for breaks in trends in which the breaks are unforecastable and they conclude suggesting to work conditionally as it was done in chapter 2. Perron (1989) discusses that the conditional analysis is justified if the break in trend are exogenous to the economic phenomena under study.

The way in which a particular economic agent incorporates changes in technology, habits, etc. could be just from time to time. But aggregating data through a big number of agents the resulting time series could show changes in trend much more often, mainly in the level factor of the trend. According with that our next step will be to consider stochastic trends in which the changes in the structure of trend happen at any time t. It will be seen that this last condition is very convenient to incorporate in the model for x_t a stochastic component, trend, which is not observed. Proceeding in this way one ends up with models which could not fully anticipate changes in trend but once they happen the models adapt inmediately to the new situation.

Let us recall the T (1^s) model:

$$x_t = a + \sum_{j=1}^{r} a_j \zeta_{jt} + w_t, \qquad (3.2)$$

where the trend T_t, at time t is

$$T_t = a + \sum_{j=1}^{r} a_j \zeta_{jt}$$

and, therefore, it only changes at times $t_1, t_2, \ldots, t_r$.

The purpose now is to replace T_t in (3.2) by a stochastic trend, say μ_t. Since trends are not observed but they are included in the values for x_t one could use x_{t-1} instead. Then:

$$x_t = x_{t-1} + w_t, \qquad (3.3)$$

where, as it has been assumed in chapter 2, the w_t components are identically distributed with zero mean.

In this equation the trend value at (t-1) is incorporated in x_{t-1} and the stochastic change of the trend at time t in w_t. Therefore w_t is not longer a deviation from trend but a deviation from the evolution path given by x_{t-1}. This is so because a trend, say μ_t, is in (3.3) but is not explictly defined.

In (3.3) there is a trend -a acyclical component which perpetuates into the future- because x_{t-1} enters with coefficient one, forcing the incorporation of the past in a persistent way. In fact, and because this unit coefficient x_t incorporates x_{t-1}, and through it x_{t-2} and so on. Therefore x_t in (3.3) can be formulated as

$$x_t = x_0 + w_t + w_{t-1} + \ldots + w_2 + w_1,$$

showing that x_t equals to the value of x at some initial time, x_0, plus all the stochastic elements, w_j, which have appeared since then. Therefore, in (3.3) any w_j component persists in future x_t values, generating a trend in them. On the other hand, if x_{t-1} would enter in (3.3) with coefficient less than one in absolute value, say 0.5, one had that

$$x_t = x_0 + w_t + 0.5\, w_{t-1} + 0.5^2\, w_{t-2} + \ldots + 0.5^{t-1}\, w_1.$$

In this case very far distant w_j components have a neglegible effect on x_t, they do not persist, and no trend is generated in x_t. Finally coefficient values greater than one in absolute terms will generate an explosive behaviour for future values, which generally is not found in real data. Therefore, coefficients for x_{t-1} in (3.3) greater than one in absolute value are excluded in time series analysis.

Model (3.3) is sometimes called a reduced form model because it results from an unspecified model for the trend and a deviation from this trend. Some authors, as Harvey (1989), have been interested in formulating two-equation model for x_t, one for the trend, say μ_t, and another for x_t as a trend plus an stationary deviation, η_t, from it. In this case:

$$x_t = \mu_t + \eta_t \tag{3.4a}$$
$$\mu_t = \mu_{t-1} + \varepsilon_t. \tag{3.4b}$$

In equation (3.4b) we have that trend at time t -μ_t- is given by the trend at time (t-1) plus a random shock, ε_t, which take place at time t. These random shocks are the ones which cause the changes in trend.

Substituting μ_t in (3.4a) by its value in (3.4b) we have

$$x_t = \mu_{t-1} + \varepsilon_t + \eta_t \tag{3.5}$$

and adding and substructing η_{t-1} in (3.5) and noting that $\mu_{t-1} + \eta_{t-1}$ is x_{t-1} one can write

$$x_t = x_{t-1} + (\varepsilon_t + \eta_t - \eta_{t-1}) \tag{3.6}$$

and denoting

$$w_t = \varepsilon_t + \eta_t - \eta_{t-1} \tag{3.7}$$

we end up with model (3.3).

A possible way to arrive to model (3.3) is by the two equation model (3.4), but this is not the unique way. In fact from (3.4) we arrive to a model (3.3) in which w_t is restricted by its definition in (3.7). Consequently the reduced form resulting from model (3.4) is called a "restricted reduced form", meanwhile (3.3) is an (unrestricted) reduced form with no restrictions on w_t but the general ones of being stationary and invertible, concept which will be properly defined later. Since (3.4) includes a definition of the structure of the trend, sometimes model (3.4) is denoted as structural model. But structural models are defined by unobserved components, like the trend μ_t in (3.4a), and by given an specific definition of them (3.4b). Since the only data available is for the time series [x_t] the precise properties of the unknown trend can not be properly tested and at the same time the structural model is not neutral in deriving a model (reduced form model) for the data [x_t], because it has been seen in the previous example that the structural model (3.4) imposes the restriction (3.7) to the reduced form (3.3). Therefore mistakes in the structural formulation of the trend will end up with a wrong reduced form model for x_t. Because of that many analysts prefer to work with reduced form models, it means models for the observed data without an explicit formulation for trend which could induce restrictions in the model for [x_t] which could not be justified by data. Nevertheless for the purposes of this book one could work with structural models or with reduced form models. For the reasons mentioned we restrict ourselves to reduced form models.

The above discussion helps to understand that in (3.3) w_t is not the deviation from trend at time t or (t-1). The trend at (t-1) is included in x_{t-1} and at time t the trend changes by a random magnitud ε_t, which is included in w_t.

It is clear now that the LOL segmented trend in (3.2) is given by [a + $\sum_{j=1}^{r} a_j \zeta_{jt}$] and that (3.3) includes an stochastic trend but without a definition for it.

As it has been mentioned in previous chapters we are interested in models which incorporate all the main factors which determines the generation of the time series under consideration, but without been particulary interested in estimating separately these factors. In fact, in models like (3.3) this can only been done using restrictions which could be questionable. Therefore since x_{t-1} is stochastic and enters with unit coefficient in (3.3), one has a model for x_t with stochastic trend.

The coefficient one of x_{t-1} in (3.3) is a root in the equation ruling the dynamics of x_t, as it will be seen later, and this type of stochastic trends are also denoted as unit-root trends.

In the model (3.3) for x_t the factor level x_{t-1} changes at any time t, therefore now the trend in x_t changes with every observation as it was seen with the example given in (3.4). Certainly a time series generated by these models shows evolutivity, but an evolutivity of a specific kind. In fact, in (3.3) a particular type of stochastic trend incorporated through the term x_{t-1} has been formulated. This particular structure implies that

$$x_t - x_{t-1} = \Delta x_t = w_t \qquad (3.8)$$

and therefore x_t in (3.3) shows evolutivity but its first differences - Δx_t- not, they are stationary. This type of evolutivity is called homogeneous evolutivity.

In (3.3) x_t evolves integrating -because of the unit root- its own past and it can be said that the whole sequence of x_t ´s, denoted as $\{x_t\}$, is an integrated process, but after differencing once the resulting process is stationary. This is so because the model has one unit root but only one. In this case then it is said that $\{x_t\}$ is integrated of order one and denoted I(1). Note that in this terminology I(•) trends are always stochastic. With the I(1) models of equation (3.3) the corresponding trends have only a level factor, therefore those models can generate series with local oscillations in level but not with systematic growth.

Figure 2.4 shows the daily yen-dollar exchange rate (X_{1t}) which evolves in a way such that model (3.3) could be appropiate for these data. In figure 2.8 the daily increments (first differences) of X_{1t} are represented and it could be appreciated that for ΔX_{1t} the mean level appears as constant with zero value and that $\Delta X1_t$ can be considered as stationary when X_{1t} is not. In practice this is a rough but useful way to detect if a time series can be considered as I(1). The procedure consists in plotting x_t against time and observe if it could be said if the mean is constant or it evolves along time clearly showing different values for different subsamples. If this is the case, transform the original x_t data in Δx_t, plot it and observe if now the mean is constant. If this is the case x_t can be considered as generated by an I(1) process. Appropriate tests for the hypothesis that x_t is I (1) will be comment later. See summary 3.1A.

Model (3.3) is only valid for sereis exhibiting local oscillations in level. A model which could represent time series with systematic growth is

$$x_t = x_{t-1} + b + w_t. \qquad (3.9)$$

In fact, in (3.9) x_t includes through the unit root all the past and at the same time incorporates an additional fix factor b which in a systematic way increases the previous level. In (3.9) x_t has an evolutivity path with two components: an stochastic level component given by x_{t-1} and a deterministic incremental component given by coefficient b and consequently the model generates time series with systematic growth. The difference between (3.9) and the linear trend model

$$x_t = a + bt + w_t \qquad \text{or} \qquad (3.10)$$
$$x_t = T_{t-1} + b + w_t, \qquad (3.11)$$

Summary 3.1.A

UNCERTAINTY IN TRENDS. MODELS WITH ONE UNIT ROOT FOR DATA WITH LOCAL OSCILLATIONS IN LEVEL.

With polynomial trends there is none uncertain future event which could affect the trend. This is quite unrealistic.

With segmented trends future breaks are uncertain but usually those models do not incorporate a stochastic formulation of the breaks.

UNIT ROOT TREND MODELS: they incorporate stochastic trends by including the inmediate past of the data (x_{t-1}) with coefficient (root) one.

The unit root perpetuates into the future everything which enters in the data at any time and thus generates a trend.

In formulating a unit root model in terms of past data a trend is incorporated in the model but a direct estimation of trend is not possible without further restrictions.

The model

$$x_t = x_{t-1} + w_t \qquad (3.3)$$

is a unit-root model in which the trend has just a level factor which changes at any time t.

The unit root in the above model disappears taking first differences of data, which consequently are stationary. Because of that (3.3) is called an integrated model of order one, I (1).

where

$$T_{t-1} = a + b\,(t-1),$$

is that in the latter both components in the trend are deterministic and in (3.9) the level x_{t-1} to which a deterministic increment b is added is stochastic.

Taking first differences in (3.9) we obtained

$$\Delta x_t = b + w_t \qquad (3.12)$$

and in (3.12) it can be observed that the mean of the increments Δx_t is given by the parameter b and if, as it is usually the case, b is positive x_t will show systematic growth. Certainly a negative b will induce systematic declaining.

With model (3.9) one also has that x_t shows evolutivity but Δx_t in (3.12) not. In other words Δx_t is stationary and therefore x_t in (3.9) is I (1). This I (1) notation is incomplete because variables with very different trends, x_t in (3.3) with local oscillations in level and x_t in (3.9) with systematic growth (with deterministic increment factor), are denoted the same: I (1). The important difference between those models comes from the fact that the increments, Δx_t, in the first case have a zero mean as it can be seen from equation (3.8) and in the second one have a non-zero mean b, equation (3.12). Therefore following Espasa and Peña (1995) we can use the notation I (1, m) where m equals zero if the mean of Δx_t is zero and equals one if the mean of Δx_t is non zero. Thus in (3.3) one has an I (1,0) model and in (3.9) and I (1, 1). Note that this new terminology makes explicit the number h^* of factors in the trend. The value h^* is given by

$$h^* = 1 + m.$$

Figure 2.12 plots the data for the US real GDP denoted as time series $X9_t$. Comparing these data with time series $X1_t$ corresponding to yen-dollar exchange rate (figure 2.4) it is clear that $X9_t$ shows systematic growth and $X1_t$ just local oscillations in level. Given the different nature of the economic phenomena represented by these two time series, these differences in trend is what can be expected on theoretical grounds. The first differences of GDP, $\Delta X9_t$, are plotted in figure 3.1 and it can be observed in this plot that the mean is non-zero and the GDP series can be taken as I (1,1). On the other hand for the yen-dollar exchange rate, $\Delta X1_t$ in figure 2.8 clearly shows a zero mean and $X1_t$ can be considered as I(1,0). In both cases, $X9_t$ and $X1_t$, differencing one gets a transformed time series which is stationary and consequently has a constant mean, but in the case of $\Delta X9_t$ this mean is non-zero inducing growth in $X9_t$ and in the case of $\Delta X1_t$ the mean is zero and no systematic growth appears in $X1_t$.

In (3.9) the model is denoted as I (1, 1) and being $h^* = 2$ the model includes systematic growth in trend. Since the model is integrated, the trend is stochastic and this property can be seen in the level factor x_{t-1}, but since m = 1 the increment factor of the trend is deterministic.

It is also interesting to have models generating systematic growth with the two stochastic components, level and incremental factor. A model for that purpose is the following:

$$x_t = x_{t-1} + (x_{t-1} - x_{t-2}) + w_t. \qquad (3.13)$$

Figure 3.1

US GDP, first differences

This model incorporates a stochastic level factor x_{t-1} and a stochastic incremental factor $(x_{t-1} - x_{t-2})$ and this is done through a unit root in both cases. The result is a two unit root model. Taking first differences in (3.13) we have

$$\Delta x_t = (x_t - x_{t-1}) = (x_{t-1} - x_{t-2}) + w_t$$

and the model for Δx_t is not stationary, still has a stochastic unit root factor. In fact the above equation shows that Δx_t is I (1, 0). Taking first differences again we get

$$\Delta^2 x_t = [(x_t - x_{t-1}) - (x_{t-1} - x_{t-2})] = w_t \qquad (3.14)$$

and in (3.14) it can be seen that the second differences of x_t, $\Delta^2 x_t$, are stationary. Each time that first differences are applied a unit root is eliminated. In (3.13) one has a model with two unit roots and now one needs to apply twice first differences in order to obtain a transformation of the data, $\Delta^2 x_t$, which is stationary. Because of that it can be said that model (3.13) is I (2) and since the stationary transformation $\Delta^2 x_t$ in (3.14) has a zero mean it can be said more precisely that model (3.9) is I (2,0). Since (3.13) is an integrated model – contemporaneus values depend on past values with a unit root- its trend is stochastic. In this case the trend has two factors, h* = 2, and none of them is deterministic (m=0), therefore this trend is fully stochastic. This is quite clear in (3.13) where the level is included in x_{t-1} and the incremental factor in $(x_{t-1} - x_{t-2})$ and both are stochastic with unit roots. Since after differencing, in this case twice, one gets a transformation of data which is stationary it can be said that the evolutivity in (3.9) is also homogeneous. The level and incremental factors of the trend at time (t-1) are in x_{t-1} and $(x_{t-1} - x_{t-2})$ but since these magnitudes are affected by w_{t-1} and w_{t-2} both elements include something else than trend. Therefore (3.9) is a model with systematic growth trend but as in the I (1) case one can not derived directly from it a full specification for the trend, and as mentioned before in general we do not care about it.

In chapter 2 the Core CPI for US, which excludes food and energy prices from total consumer price index, was presented. An annual time series of this variable, denoted $Z7_t$, is plotted in figure 3.2, panel A. Since in these data the proportionality property is present, it is convenient to work with the logarithmic transformation, as it was done in figure 3.2. This figure shows that in Core CPI there is systematic growth. Panel B in the same figure shows the plot of

$$\log Z7_t - \log Z7_{t-1} = \pi_t$$

which as discussed in the previous chapter is a good approximation of Core inflation. In this plot it can be observed that the Core inflation does not oscillates around a fix mean and therefore it cannot be considered as stationary. In fact Core inflation shows local oscillations in level, a type of behaviour similar to the one we discussed for the yen-dollar exchange rate. It can be concluded that Core inflation is I (1). Figure 3.3 gives the plot of

$$\Delta^2 \log Z7_t = \Delta \pi_t = [Z7_t - Z7_{t-1}] - [Z7_{t-1} - Z7_{t-2}],$$

which is the plot of the second differences of log Core CPI or the first differences of Core inflation. In this figure it can be observed that $\Delta \pi_t$ oscillates around a constant mean and therefore $\Delta \pi_t$ is stationary or I (0) and then π_t is I (1) or more precisely I (1,0) because $\Delta \pi_t$ has a zero mean. Finally $\log Z7_t$ is I (2,0).

The properties of models I (1,0) I (1,1) and I (2,0) are summarized in table 3.1.

Figure 3.2

Log US CORE CPI

US CORE INFLATION (π_t = log CPI_t – log CPI_{t-1})

Figure 3.2.

STATIONARY TRANSFORMATION OF US CORE CPI.

Table 3.1. Integrated trends.

MODEL	PROPERTIES
(3.3) I (1.0) $x_t = x_{t-1} + w_t$	(a) It has a trend with a level factor. (b) The trend is not defined, but its value at (t-1) is included in x_{t-1}. At time t its value change due to an stochastic shock which in (3.3) is included in w_t. Therefore the trend is fully stochastic. (c) Model (3.3) has a trend because it is incorporated through the unit coeffcient (root) of x_{t-1}. Then (3.3) represents an integrated process of order one: I (1). (d) Applying first differences - $\Delta x_t = w_t$ - the unit root is eliminated and the resulting transformed data is stationary with zero mean. In more precise terms it can be said that (3.3) represents an integrated process of order I (1,0). The sum of the arguments in I (•, •) is one pointing out that the trend in (3.3) only has one factor and can only generate time series with local oscillations in level. (e) Since x_t is I (1, 0), Δx_t is I (0,0) or stationary with zero mean. (f) Example: the daily yen-dollar exchange rate can be considered as I (1,0). See figures 2.4 and 2.8.
(3.9) I (1,1) $x_t = x_{t-1} + b + w_t$	(a) It has a two-factor trend: level and increment. (b) The trend is not fully defined. The increment factor is deterministic and given by parameter b, but the level factor is not defined. Its value at time (t-1) is included in x_{t-1} and at time t its value changes due to an stochastic shock, which in (3.9) is included in w_t. Therefore the trend is stochastic but with deterministic increment factor. (c) Model (3.9) has a trend because it is incorporated in the model with a unit root. Then (3.9) represents an integrated process of order 1: I (1). (d) Applying first differences -$\Delta x_t = b + w_t$- the unit root is eliminated and the resulting transformed data is stationary with non-zero mean. In more precise terms it can be said that (3.9) represents an integrated process of order I (1,1) The sum of the arguments of I (•, •) is two, pointing out that the trend in (3.9) has two factors. The second argument is non-zero and it indicates that the second factor, the increment, is deterministic. Therefore (3.9) can generate time series with systematic growth in which the growth has a constant mean. (e) Since x_t is I (1, 1), Δx_t is I (0, 1) or stationary with non-zero mean. (f) Example: the quarterly US GDP in real terms can be considered as I (1,1). See figures 2.12 and 3.1.

Table 3.1, Continuation

MODEL	PROPERTIES
(3.13) I (2,0) $x_t = x_{t-1} + (x_{t-1} - x_{t-2}) + w_t$	(a) It has a trend with two factors: level and increment. (b) The trend factors are not defined but their values at (t-1) are included in x_{t-1} and $(x_{t-1} - x_{t-2})$. At time t the values of these factors change due to two stochastic shocks, which in (3-13) are included in w_t. Therefore both factors are stochastic. (c) These factors really generate a trend because they are incorporated in the model with a unit root each one, resulting in a two-unit root trend. Model (3.13) represents an integrated process of order 2: I (2). (d) Applying first differences one unit root is eliminated. Therefore twice first differences are required to transform x_t as stationary $\Delta^2 x_t = w_t$. Since the mean of $\Delta^2 x_t$ is zero,it can be said in more precise terms that x_t is I (2,0). The sum of the arguments in I(•, •) is two, pointing out that the trend in (3.13) has two factors and since the second argument is zero both factors are stochastic. Therefore (3.13) can generate time series with systematic growth in which the growth is also non-stationary. (e) Since x_t is I (2,0), Δx_t is I (1,0). (f) Example Core CPI is I (2,0) and Core Inflation I (1,0). See 3.2 and 3.3.

In the discussion of model (3.3) it became clear that the trend in this model change with every observation due to an stochastic shock, say ε_t, which is included in w_t. Then through the unit root this shock is maintained in the future. In model (3.9) the deterministic increment factor of the trend certainly does not change with time, but the level factor changes at each time t in a similar way than in model (3.3) by the incorporation of a stochastic shock and in this case also by the incorporation of a deterministic increment. In model (3.13) the trend has two stochastic factors and the corresponding structural model, which is not going to be described here, now has three equations. One defining x_t as the sum of a trend plus a stationary factor and the other two defining the two stochastic components of trend. Both components change each time due to two different stochastic shocks, which in the reduced equation (3.13) are included in w_t.

In all these three models, (3.3), (3.9) and (3.13), the stochastic shocks entering in the trend at time t are unpredictable. Therefore if at time t^* the trend suffers a big stochastic shock this shock certainly is not incorporated in x_{t*-1} and with information till (t^* - 1) the above models will make a bad forecast for x_{t*}. The situation will be similar to the forecast with a segmented trend model with information till time (t^* -1) if at time t^* a new break in the trend takes place. However, with the integrated model the big shock occured at time t^* will be incorporated into the trend and into the model through the unit root and it will be included in all forecasts done with information at least till time t^*. It can be seen then that before a big shock the integrated models will produce a bad forecast but once the shock has taken place the models adapt themselves and produce forecasts according to the new situation. This is a big difference with segmented trend models which after a new break they remain wrong for ever. For this reason even when an economic phenomenon could be generated by a segmented trend model one could try to approximate it by using a corresponding integrated model. This model will induce small changes in trend at each time t which, in general, could not be much distorting since they are due to minor random shocks, which by nature will tend to have positive sign half of the times and negative sign another half. On the other hand when a greater shock arrives, the integrated model will incorporate it inmediately for future forecasts. See summary 3.1B.

From the above discussion it becomes clear that the main difference between segmented trend models and unit-root models or models with integrated trends lie in how often the trend changes. It has been argued before that changes in the level of the trend could occured much more often than changes in the growth factor, therefore for time series with systematic growth models with one unit root can be seen as very useful. Then it could be a more open question if one is going to use an I (1, 1) model which takes the increments of the data as stationary with a constant mean or if one is going to employ an I (2, 0), where those increments are not stationary. Clearly a third alternative appears as interesting. A model having a unit root but with a segmented mean in the first differences. This model can be written as

$$x_t = x_{t-1} + b + \sum_{j=1}^{r} b_j \zeta_{jt} + w_t \qquad (3.15)$$

and taking first differences we have

$$\Delta x_t = (x_t - x_{t-1}) = b + \sum_{j=1}^{r} b_j \zeta_{jt} + w_t. \qquad (3.16)$$

Summary 3.1.B

MODELS WITH UNIT ROOT TRENDS.

$$x_t = x_{t-1} + w_t \quad \text{and} \quad (3.3)$$

$$x_t = x_{t-1} + b + w_t \quad (3.9)$$

are both I (1), but (3.3) only integrates the past and (3.9) also incorporates a constant term b each time.

TREND in (3.3) only shows local oscillations in level and trend in (3.9) systematic growth with a constant (deterministic) mean value b.

Because the mean of Δx_t in (3.3) is zero but in (3.9) is different from zero it can be said that (3.3) is integrated I (1,0) and (3.9) I (1,1).

With the I (1,m) notation, m = 0, 1, the value

$$h^* = 1 + m$$

indicates the number of factors in the definition of the trend.

A unit-root model generating systematic growth with both stochastic level and incremental factor is

$$x_t = x_{t-1} + (x_{t-1} - x_{t-2}) + w_t \ . \quad (3.13)$$

This model has two unit roots, one integrating the level and another integrating the increments.

Taking first differences in (3.13)

$$\Delta x_t = \Delta x_{t-1} + w_t$$

the resulting model still has a unit root. To obtain an stationary transformation one needs to apply twice first differences.

$$\Delta^2 x_t = w_{t.} \quad (3.14)$$

In (3.15) $\Delta^2 x_t$ has a zero mean.

Model (3.13) is denoted as integrated of order I (2, 0).

Model (3.15) can be denoted as I (1, 1^s) because has one unit root and the differenced series has a non-zero mean which is segmented. The segmentation as in the previous chapter is incorporated in the notation by using the superscript s.

Model I (1, 1^s) makes much sense for many economic time series showing systematic growth for which one could expect that the mean of growth changes occasionally but not every time as it happens with model I (2, 0). But still the I (2, 0) has the advantage that when a big shock to the growth factor will take place the model will adapt automatically to the new situation and model I (1, 1^s) not.

We can conclude this discussion by saying that for series with systematic growth model I (1, 1^s) could be very useful in describing the main features of the data, but for everyday forecasting purposes model I (2, 0) has the advantage that it adapts itself to the new situation after big shocks.

The I (1, m) or I (2, m) notation can be generalized as I (d, m) as it is done in Espasa and Peña (1995) o better as I (d, m^s). In this last notation "d" refers to the number of unit roots in the model or what is the same, the number of times which one needs to apply first differences before obtaining a transformation of the data which is stationary or contains a deterministic trend. Therefore a varible x_t generated by a model I (d, m^s) is called integrated of order d. The coefficient m takes the value zero if the d-times differenced data, $\Delta^d x_t$, has a zero mean, the value one if it has a constant non-zero mean and the value m if the difference data contains a deterministic trend with m components. Thus if x_t is such that

$$\Delta x_t = a + bt + w_t \qquad (3.17)$$

then d = 1, m = 2 and therefore x_t is I (1, 2).

If the possible constant mean or deterministic trend in $\Delta^d x_t$ is segmented then this feature is indicated by using a superscript s in m.

With this notation the T (2) model in the previous chapter is denoted now as I (0, 2).

In the I (d, m^s) notation the number h^* of components in the trend is given by

$$h^* = d + m.$$

If m equals zero all trend components are stochastic, if d equal zero all trend components are deterministic and if both d and m are non-zero the trend has stochastic and deterministic components and the component of highest order is deterministic.With this terminology $h^* > 0$ denotes the presence of trend except in the case I (0, 1), where data oscillates around a constant mean and therefore the corresponding variable is stationary. Thus I (d, m) is a valid notation for trend for I (d, m \ d≠0 or d=0 and m > 1). It must be observed that in the case where the model contains some segmented factor then the I (d, m^s) terminology is always a valid notation for trend.

As it has been discussed previously in different ocassions the trends in economic time series can be considered to have up to two components, therefore the usual models of type I (d, m^s) will be such that

$$0 \leq h^* = d + m \leq 2.$$

Therefore model (3.17) where $h^* = 3$ are very rarely appropiate for economic time series. In fact the models which turn to be most useful for economic time series are, I (1, 0) for series with local oscillations in level and I (1, 1), I (1, 1^s) or I (2, 0) for series with systematic growth. See summary 3.1.C

Summary 3.1.C.

I (d, m^s) NOTATION

The I (d, m^s) terminology is used to denote the type of trend in an economic time series x_t. Thus

$$h^* = d + m$$

indicates the number of factors in the trend and:

d: indicates the number of trend factors captured through past values of x_t with a unit root. Those factors are then stochastic.
Each time that we apply first differences to x_t we eliminate one of the mentioned unit-root factors.

m can take the following values:

- zero if the mean of $\Delta^d x_t$ is zero
- one if the mean of $\Delta^d x_t$ is a non-zero constant
- m* if $\Delta^d x_t$ includes a time polynomial with m* factors.

Therefore m indicates the number of deterministic factors in the trend of x_t.

The subscript s indicates that the non-zero mean or polynomial trend in $\Delta^d x_t$ is segmented.

3.2. STOCHASTIC SEASONALITY. MODELS I (d, m^s) (SS).

In chapter 2 we mentioned some factors causing seasonality. Those factors do not change rapidly along time and as a first approximation in section 2.9 seasonality was considered as fixed and therefore as deterministic. But future evolution of the factors behind seasonality is uncertain and seasonality should , in general, be considered as stochastic.

Simillary as it was done with model (2.29) with segmented mean -I $(0,1^s)$- one could considered that the seasonal factors of models in section (2.9) change rarely from time to time. Conditional to the knowledge of the time points at which these changes have taken place one could formulate a model with changing seasonal factors such that inside the sample the seasonal factors could be taken as deterministic but future factors are clearly stochastic. A seasonal model of this type has similar inconvenients to those pointed out for segmented trend models. Consequently, and proceeding in a similar way than in previous section, we can considered stochastic seasonal models in which the seasonal factors change, usually very smoothly, at any time t. In the formulation of stochastic seasonallity it makes not much sense to consider that the trend could be deterministic and we will introduce stochastic seasonality only upon the models discussed in section 3.1.

An I (1,0) model with deterministic seasonallity can be written as

$$x_t = x_{t-1} + \sum_{j=1}^{s} a^*_j S_{jt} + w_t, \qquad (3.18)$$

where S_{jt}, j = 1, …, s, are the seasonal dummy variables defined in section (2.9) and the seasonal factors follows the restriction

$$\sum_{j=1}^{s} a^*_j = 0. \qquad (3.19)$$

In (3.18) at each time t only one a^*_j parameter appears because all the others are multiplied by seasonal dummies with zero value.

From (3.18) we can write

$$\Delta x_t = \sum_{j=1}^{s} a^*_j S_{jt} + w_t \qquad (3.20)$$

and in (3.20) again only one a^*_j enters at each time t and according with (3.19) for this a^*_j, that without lost of generality could considered to be a^*_s, we have that

$$a^*_s = -\sum_{j=1}^{s-1} a^*_j.$$

As it happens with trend, seasonal factors are not observed but we observe x_t which includes them. Now one must take into consideration that under the hypothesis that x_t follows an I (1,0) model it has trend and this trend must be removed in order to take observed data as proxy for seasonal factors. In equation (3.20) the transformed data, Δx_t, has no trend but has seasonality. The aim now is to change the deterministic seasonal factor, say a^*_s, appearing in (3.20) by a stochastic factor, say σ_{st}, which in mean follows a restriction like (3.19). This implies that σ_{st} equals mimus $\sum_{j=1}^{s-1}\sigma_{j(t-s+j)}$ plus some stochastic perturbation.

With all these considerations we could write

$$x_t = x_{t-1} - \sum_{j=1}^{s-1} \Delta x_{t-j} + w_t. \qquad (3.21)$$

In (3.21) x_{t-1} captures the stochastic I (1,0) trend which disappear taking first differences and then the factor $[\ -\sum_{j=1}^{s-1}\Delta x_{t-j}\ ]$ incorporates seasonality. As in the previous section, (3.21) is a reduced from model which results from a structural model with three equations. One defining x_t as the sum of a trend, seasonal factors and a residual component, and the other two defining the trend and seasonal factors. These last two equations are affected by contemporaneous stochastic shocks, given a stochastic nature to trend and seasonal factors. These shocks in the reduced form model (3.21) are included in w_t. It must be noted that the factor $[\ -\sum_{j=1}^{s-1}\Delta x_{t-j}\ ]$ incorporating seasonality enters in the model with coefficient (root) one and (3.21) is also a model with seasonal unit roots. Since the seasonal factor has (s-1) elements there are also (s-1) seasonal unit roots.

Equation (3.21) can be formulated as

$$x_t = x_{t-1} - [(x_{t-1} - x_{t-2}) + (x_{t-2} - x_{t-3}) + \dots + (x_{t-s+1} - x_{t-s})] + w_t \qquad (3.22)$$

and cancelling terms we end up with

$$x_t = x_{t-s} + w_t, \qquad (3.23)$$

which is a simpler way to write (3.21). Apparently in (3.23) there is only one unit root but since it operates on the lag s it is seen from (3.22) that really there are s unit roots one positive refering to trend and (s-1) complex and negative refering to the (s-1) independent seasonal factors.

Both models (3.18) and (3.23) are I (1,0) but the first one with deterministic seasonality (DS) and the second with stochastic seasonality (SS). Therefore it is convinient to denote (3.18) as I (1,0) (DS) and (3.23) as I (1,0) (SS).

Denoting by Δ_s the seasonal difference operator which is such taht

$$\Delta_s x_t = (x_t - x_{t-s}),$$

it can be seen from equations (3.22) and (3.23) that

$$\Delta_s x_t = \sum_{j=0}^{s-1} \Delta x_{t-j}$$

which implies that seasonal differences, $\Delta_s x_t$, equals to the sum of the corresponding s consecutive first differences.

Therefore seasonal differencing implies two very different operations. First the data is transformed in first differences, eliminating LOL trends, and then at each point in time s consecutive first differences are summed, eliminating seasonality. Indeed, seasonality means cyclical evolution of one-year period and then is clear that transforming the observations –previously transformed by first differences to eliminate the trend- by sums over the cyclical period which ends at time t one removes the cyclical oscilation operating at time t.

From (3.23) it can be seen that by taking seasonal differences in (3.23) one obtains

$$x_t - x_{t-s} = \Delta_s x_t = w_t, \qquad (3.24)$$

which is stationary. This means that in (3.24) the two evolutivity factors -trend and seasonallity- present in (3.21) or (3.23) have been eliminated. But it is important to realice that with the first differencing application embedded in seasonal differencing we have removed the trend and by the summing application the seasonallity.

If in (3.21) one takes first differences only

$$\Delta x_t = -\sum_{j=1}^{s-1} \Delta x_{t-j} + w_t \qquad (3.25)$$

and it can be seen that Δx_t still has seasonal evolutivity. This is so because Δx_t in (3.25) has a unit-root dependence on its own past and this past with Δx_t sums up to complete the seasonal cycle ending at t. In other words Δx_t in (3.25) is not stationary, in order to get a stationary transformation one needs to sum the dependent variable in (3.25), through out the whole yearly cycle ending at t. Then, from (3.25)

$$\sum_{j=0}^{s-1} \Delta x_{t-j} = w_t ,$$

which is stationary.

The above equation can be reformulated as

$$\sum_{j=0}^{s-1} x_{t-j} = \sum_{h=1}^{s} x_{t-h} + w_t \qquad (3.26)$$

and using

$$z_t = \sum_{j=0}^{s-1} x_{t-j}$$

we can write

$$z_t = z_{t-1} + w_t,$$

which clearly shows that the dependent variable z_t in (3.26) has a LOL trend.

At this stage it is interesting to introduce some aditional notation. Let us call B to the lag operator which is such that when applied to x_t lags it one period. This means that

$$Bx_t = x_{t-1}.$$

Applying now B to x_{t-1} one obtains that

$$B(B\,x_t) = B^2\,x_t = x_{t-2}.$$

In general

$$B^j\,x_t = x_{t-j}.$$

Using the lag operator

$$\Delta = (1 - B)$$

and

$$\Delta_s = (1 - B^s) = (1 - B)\,(1 + B + \ldots + B^{s-1}),$$

which shows that the seasonal difference operator is the product of a first difference operator and a sum operator over the s consecutive observations ending at time t. This last operator can be called seasonal sum and we will denoted by U_{s-1} (B) such that

$$U_{s-1}(B) = (1 + B + B^2 + \ldots + B^{s-1}). \qquad (3.27)$$

In model (3.21) there is no systematic growth but just local oscillations in level because, as it can be seen from (3.23), the mean of the increments, now seasonal increments, is zero. The fact that in (3.23) the level factor is not the previous value of x_t as in model (3.3) but the value in the same season of the previous year indicates that the model generates local oscillations in level with stochastic seasonallity.

It is important to note that model (3.18) gives a direct estimation of seasonal factors, because they are fixed and have been included in the model. Model (3.23) includes a stochastic seasonal scheme but from it one can not get without imposing further restrictions an estimation of the seasonal factors. This question is identical to the one discussed in the previous sections when studing stochastic trends and as it was said there we do not care about it.

The main points of the previous discussion are in summary 3.2.A.

Summary 3.2.A.

STOCHASTIC SEASONALITY: MODELS I (1, 0) (SS)

(a) The future evolution of the factors causing seasonality is uncertain and seasonality should, in general, be considered as stochastic.

In this case it makes sense to consider seasonality only in models with stochastic trends.

(b) Since the seasonal factors are restricted on their means to sum zero, stochastic seasonality can be introduced in model I (1, 0) by including with unit coefficient (root) the term $\left[\; -\sum_{j=1}^{s-1} \Delta x_{t-j} \quad \right]$.

The resulting model is (3.21) and cancelling terms one gets a simpler formulation:

$$x_t = x_{t-s} + w_t \,. \qquad (3.23)$$

To make explicit that this model incorporates stochastic seasonality (SS) it will be denoted as I (1,0) (SS).

A model I (1, 0) with deterministic seasonality (DS) like (3.18) will be denoted as I (1,0) (DS).

(c) Models I (1, 0) (DS) incorporate an explicit formulation of seasonality and the seasonal factors can be estimated directly. Models I(1,0) (SS) incorporate seasonality through past values of Δx_t over the current seasonal cycle with unit roots. In this case an estimation of the seasonal factors is not possible without aditional and questionable restrictions.

(d) If a time series is generated by a I (1, 0) (SS) model, their seasonal differences are stationary.

Seasonal differencing includes two different operations: (1) a transformation of data in first differences and (2) the sum of s consecutive first differences. The first operation eliminates trend and the second seasonality.

PUT HERE AN EXAMPLE ON UNEMPLOYMENT DATA, AS A SERIES I (1,0) (SS).

For systematic growth one could consider the I(1,1) model. This model with deterministic seasonality takes the form:

$$x_t = x_{t-1} + b + \sum_{j=1}^{s} b^*_j S_{jt} + w_t, \qquad (3.28)$$

where

$$\sum_{j=1}^{s} b^*_j = 0 \text{ and}$$

therefore for any b^*_j , for instante, b^*_s, one has that

$$b^*_s = -\sum_{j=1}^{s-1} b^*_j .$$

Model (3.28) can be formulated as

$$\Delta x_t - b = \sum_{j=1}^{s} b^*_j S_{jt} + w_t, \qquad (3.29)$$

where the dependent variable (Δx_t-b) is free of trend. In (3.29) at each time t only one b^*_j enters in the equation.

One way to formulate (3.28) or (3.29) with stochastic seasonality is substituting the corresponding b^*_s coefficient in those equations by

$$-\sum_{j=1}^{s-1} (\Delta x_t - b).$$

Then

$$x_t = x_{t-1} + b - \sum_{j=1}^{s-1} (\Delta x_{t-j} - b) + w_t. \qquad (3.30)$$

In (3.30) x_{t-1} captures the stochastic level factor of the trend, b the determinsitic incremental factor and $\{-\sum_{j=1}^{s} (\Delta x_j - b)\}$ the seasonal factor.

From (3.30) one gets after a similar manipulation than before

$$x_t = x_{t-s} + s \bullet b + w_t, \qquad (3.31)$$

which is a simpler formulation, in which it is clear that the model is I (1, 1) with stochastic seasonality, because the stochastic level factor in (3.31) is a seasonal level factor.

The I (2, 0) model with deterministic seasonality can be written as

$$x_t = x_{t-1} + (x_{t-1} - x_{t-2}) + \sum_{j=1}^{s} b^*_j S_{jt} + w_t, \qquad (3.32)$$

or

$$\Delta x_t - \Delta x_{t-1} = \sum_{j=1}^{s} b^*_j S_{jt} + w_t ,$$

where the dependent variable $[\Delta x_t - \Delta x_{t-1}]$ is free of trend.

In this case a model with stochastic seasonality takes the form

$$x_t = x_{t-1} + (x_{t-1} - x_{t-2}) - \sum_{j=1}^{s-1} (\Delta x_{t-j} - \Delta x_{t-j-1}) + w_{t.} \qquad (3.33)$$

In (3.33) x_{t-1} captures the stochastic level factor of the trend, $(x_{t-1} - x_{t-2})$ the stochastic trend incremental factor and

$-\sum_{j=1}^{s} (\Delta x_{t-j} - \Delta x_{t-j-1})$ the stochastic seasonality.

After some manipulation (3.33) can be written as

$$x_t = x_{t-1} + (x_{t-s} - x_{t-s-1}) + w_t \text{ or} \qquad (3.34)$$

as

$$x_t = x_{t-s} + (x_{t-1} - x_{t-s-1}) + w_t . \qquad (3.35)$$

In both cases we have level and incremental factors which are stochastic and one of them is seasonal, indicating that the model generates systematic growth with stochastic seasonality.

In this case neither first differences, Δx_t, nor seasonal differences, $\Delta_s x_t$, are stationary. The stationarity is obtained after applying one regular and one seasonal difference, $\Delta\Delta_s$.

It is convinient to consider different ways of representing $\Delta \Delta_s$. Thus

$$\begin{aligned} \Delta \Delta_s = (1 – B) (1 – B^s) &= 1 – B – (B^s – B^{s+1}) && (3.36a)\\ &= 1 – B^s – (B – B^{s+1}). && (3.36b) \end{aligned}$$

Also

$$\begin{aligned} \Delta \Delta_s = \Delta^2 U_{s-1} (B) &= \Delta^2 (1 + B + B^2 + \ldots + B^{s-1}) && (3.36c)\\ &= \Delta^2 + \Delta^2 (B + B^2 + \ldots + B^{s-1}) &&\\ &= [1 – L – (1-L)] + \Delta^2 (B + B^2 + \ldots + B^{s-1}). && (3.36d) \end{aligned}$$

The model for x_t takes the form

$$\Delta \Delta_s x_t = w_t.$$

Using (3.36d) for $\Delta\Delta_s$ the above model is represented according to equation (3.33). Using (3.36a) the representation (3.34) is obtained and making use of (3.36b) the equation (3.35). Finally, the right hand side term in (3.36c) makes clear the way of transforming in this case x_t in a stationary variable: applying twice first differences and then summing over the seasonal cycle.

In chapter 2 the US monthly time series corresponding to the index of industrial production (IIP) was called $X8_t$ and its plot is in figure 2.11. The plot quite clearly indicates that data shows systematic growth and also this type of behaviour is the one to be expected for a time series reflecting industrial production. Therefore for this series it can be assumed that the trend has two factors. Figure 2.11 also shows seasonal evolutivity.

In figures 3.4A and 3.4B it can be seen the first and seasonal differences of log $X8_t$, respectively. In both cases the mean level does not seem constant indicating that in order to obtain an stationary transformation one needs to apply a second differencing operation. The plot of Δ log IIP_t also shows seasonal oscillations therefore in this case the second difference operation should be seasonal.

In both cases one arrives to $\Delta\ \Delta_s$ log IIP_t which is plotted inf igure 3.5 where it can be seen that this transformation can be taken as stationary.

Figure 3.4A

Figure 3.4.B.

Figure 3.5

Alternatively looking at figure 3.4A one could considered that a model with constant mean and deterministic seasonality could serve as a first approximation. In this case

$$\Delta \log IIP_t = b + \sum_{j=1}^{s-1} b^*_j S^*_{jt} + w_t , \qquad (3.37)$$

implying that log IIP_t is given by model (3.28). The coefficients in (3.37) can be estimated, as shown in chapter 2, using the restriction that the sum of the b^*_j coefficients is zero. Then the coefficients $\hat{b}_j^*$ give an estimation of the seasonal factors of Δ log IIP_t under the assumption that they are fixed. Those coefficients are given in figure 3.6.

THE CORRESPONDING COEFFICIENTS FOR log IIP_t CAN BE DERIVED FROM PARAMETERS $\hat{b}_j^*$ AS FOLLOWS.

From the above results one could ask which model (3.28) or (3.34) could be better for log IIP_t. It has been discussed already the interest of the stochastic formulations for trend and seasonal factors and also their limitations. On the whole and mainly for forecasting proposes the stochastic formulations could be in many cases preferable, but when the length of the time series is short, less than five years say, the alternative of deterministic seasonality is advisable. In the next chapter it will be commented how one can find evidence that the deterministic approach could be more unsatisfactory than the stochastic one or how it is possible to detect if the deterministic formulation is a valid approximation for a given time series. This will be done looking at the properties of the corresponding w_t residuals in each case.

The most interesting models for economic time series allowing for deterministic or stochastic trends or seasonalities are collected in table 3.2. and the main points of the last part of this section are in summary 3.2.B.

Summary 3.2.B.

STOCHASTIC SEASONALITY: MODELS I (1, 1) (SS) AND I (2,0) (SS).

(a) For models I (1,1) stochastic seasonality (SS) can be introduced by including with unit coefficient (root) the term [$-\sum_{j=1}^{s-1} (\Delta x_{t-j} - b)$].

The resulting model is (3.30) and canceling terms one gets a simpler formulation

$$x_t = x_{t-s} + s \bullet b + w_t. \qquad (3.31).$$

This model can be denoted I (1,1) (SS).

(b) For models I (2,0) stochastic seasonality can be introduced by including the term [$-\sum_{j=1}^{s-1} (\Delta x_{t-j} - \Delta x_{t-j-1})$].

The resulting model is (3.33) and simpler formulations of it are

$$x_t = x_{t-1} + (x_{t-s} - x_{t-s-1}) + w_t \text{ or} \qquad (3.34)$$

$$x_t = x_{t-s} + (x_{t-1} - x_{t-s-1}) + w_t\,. \qquad (3.35)$$

This model can be denoted as I (2,0) (SS).

In this case neither first differences, Δx_t, or seasonal differences, $\Delta_s\, x_t$, are stationary. Stationarity is obtained applying one regular and one seasonal difference. More precisely applying twice first differences and then summing over the seasonal cycle.

An example of data for which model I (2,0) (SS) could be appropiate is the index of industrial production for US. See figures 2.11, 3.4A, 3.4B and 3.5.

(c) At least for forecasting purposes stochastic seasonal formulations are in many cases preferable to deterministic schemes.

But for short time series, less than five years, deterministic schemes are advisable. Evidence in favour of stochastic seasonality with respect deterministic seasonality can be obtained analyzing the residuals of the corresponding models, as it will be seen in the next chapter.

Table 3.2 **MODELS FOR ECONOMIC TIME SERIES WITH TREND AND SEASONALITY.**

TREND **I (d,m)**	**TREND PLUS DETERMINISTIC SEASONALITY** **(DS)**	**TREND PLUS STOCHASTIC SEASONALITY** **(SS)**
	A. DETERMINISTIC TRENDS	
A.1. **(1)** WITH LOCAL OSCILLATIONS IN LEVEL I (0, 1^s) without seasonality. Models (2.29) or (2.37)	**(2)** Local oscillations in level with deterministic seasonality: Like models (1) but enlarged with $\sum_{j=1}^{s-1} a_j^* S_{jt}^*$	
A.2. WITH SYSTEMATIC GROWTH without seasonality: **(3)** Without segmentation I (0, 2). Model (2.5).	**(4)** Systematic growth without segmentation and with deterministic seasonality: I (0,2) (DS). Model (2.47).	
(5) With segmentation I (0, 2^s). Model (2.33).	**(6)** Systematic growth with segmentation and deterministic seasonality I (0, 2^s) (DS) Like model (5) but enlarged with $\sum_{j=1}^{s-1} a_j^* S_{jt}^*$	
	B. STOCHASTIC TRENDS.	
B.1. **(7)** WITH LOCAL OSCILLATIONS IN LEVEL without seasonality I (1, 0). Model (3.3)	**(8)** Local oscillations in level with deterministic seasonality I (1, 0) (DS). Model (3.18).	**(9)** Local oscillations in level with stochastic seasonality I (1, 0) (SS). Model (3.21) or (3.23).
B.2.SYSTEMATIC GROWTH WITH DETERMINISTIC MEAN GROWTH without seasonality: **(10)** Without segmentation I (1, 1). Model (3.9).	**(11)** Systematic growth with deterministic mean growth without segmentation and with deterministic seasonality. I (1,1) (DS). Model (3.28).	**(12)** Systemtic growth with deterministic mean growth without segmentation and with stochastic seasonality I (1, 1) (SS). Model (3.30) or (3.31).
(13) With segmentation I (1, 1^s). Model (3.15).	**(14)** Systematic growth with segmented mean growth and with deterministic seasonality I (1, 1) (DS). Like model (13) enlarged with $\sum_{h=1}^{s-1} b_h^* S_{ht}^*$	**(15)** Systematic growth with segmented mean growth and stochastic seasonality I (1, 1^s) (SS). Like model (3.15) but substituting x_{t-1} by x_{t-s}.
B.3. **(16)** SYSTEMATIC GROWTH FULLY STOCHASTIC without seasonality. I (2, 0). Model (3.13).	**(17)** Systematic growth fully stochastic with deterministic seasonality: I (2, 0) (DS). Model (3.32).	**(18)** Systematic growth fully stochastic with stochastic seasonality: I (2, 0) (SS). Model (3.33) or (3.34) or (3.35).

3.3. REMOVING TRENDS AND SEASONALITY.

In section 3.1. stochastic trends based on unit roots were discussed. They imply that a model explaining x_t includes the previous level, and perhaps also the previous increment of x_t, with coefficient values of one. Then it becomes obvious that by differencing once or twice the trend disapperars. Thus the trend in a model I (1, 0), see equation (3.3), disappears after differencing once and the same happens with models I (1, 1) as in equation (3.9). This is so because in both cases the models contain only one unit root. For models I (2, 0), i.e. with two roots, like model (3.13), the trend is eliminated after two differences. In general for models I (d, m), $d \neq 0$ and $m = 0$ or 1, after differencing d times the trend is removed. This is so by construction. In fact unit-root trends were proposed after observing that in many instances differencing non-stationary data, once or twice, the resulting time series could be considered as stationary. Then the original data is a fortiori characterized by unit-root dependency of its own past. Consequently proposing unit-root formulations for non-stationary time series as it has been done in previous sections, implies that stationarity is necessarily obtained by differencing.

For deterministic trends like I (0, 2)

$$x_t = a + bt + w_t \tag{3.38}$$

after differencing the following expression

$$\Delta x_t = b + (w_t - w_{t-1}) \tag{3.39}$$

is obtained and in Δx_t there is not a trend any more. In general for models I (0, m), $m \geq 2$, after differencing (m-1) times the trend disapperars. So it can be stated that differencing also remove trends in models I (o, m). Nevertheless I (0, m) models include a specific formulation of trends different from unit roots and certainly the optimal way to remove the trend in them is not by differencing, but by estimating the trend coefficients through a regression and taking the residuals as the deviations of x_t from the trend, which by definition constitute the stationary transformation of x_t. Thus in case of model (3.38) estimating the parameters a and b the trend is removed as

$$x_t - a - bt = w_t \,. \tag{3.40}$$

In this case the differenced data, ignoring the mean b, is $(w_t - w_{t-1})$, i.e. the first differences of the w_t residuals, which constitute the proper stationary transformation of x_t. It will be seen later that transformed data of this type - $(w_t - w_{t-1})$- have undesirable statistical properties. In particular they cannot be formulated in terms of a convergent sequence of its own past values plus a random shock.

Similar results are obtained when differencing other I (o, m) models. The conclusion is that by differencing, polynomial trends are removed, but the proper way of doing it is by regression analysis and consequently after differencing in those cases one obtains a transformed time series with bad properties.

Models with segmented trends or segmented means of the type I (o, m^s), m > o, capture a non-linear trend and the proper way to remove the trend is by applying the corresponding regression and taken the residuals. In those cases, for instance model (2.31) with just one break at time t^* which we write here again

$$x_t = a + bt + a_1\, \zeta_{1t} + b_1\, \xi_{1t} + w_t, \qquad (3.41)$$

applying first differences most of the trend is removed but not all and these remaining trends parts could distort the posterior analysis. Applying first differences to (3.41) one gets

$$\Delta x_t = b + a_1\, D_{1t} + b_1\, \zeta_{1t} + (w_t - w_{t-1}). \qquad (3.42)$$

These differenced data contain, appart from the constant mean b, a residual $(w_t - w_{t-1})$ with the bad properties mentioned before and the remainings of the original trend given by

$$a_1\, D_{1t} + b_1\, \zeta_{1t}, \qquad (3.43)$$

where D_{1t} takes value one at observation t^* and zero otherwise and ζ_{1t} has already been defined and takes value one from t^* onwards and zero otherwise.

For models I (d, m), $d \neq 0$ and m>1, after differencing d times one still has a polynomial trend with m factors. The proper way to remove trend in this case is by differencing d times and run the appropiate regression on $\Delta^d\, x_t$ in order to eliminate the remaining deterministic trend. In this case trend can be eliminated by differencing [d + (m-1)] times, but doing that one ends up with residuals with the undesirable properties mentioned above.

Finally for models I (d, m^s), $d \neq 0$ and m> 0, the trend is adequately removed by differencing first and then running the corresponding regression on $\Delta^d\, x_t$. If in this case [d + (m-1)] differences are applied the trend is not fully removed and the residuals also have the inconvenient of containing elements similar to those in (3.43).

In table 3.3 we summarize the above discussion showing the appropiate trend removing procedure for each model and comment on the effects of removing it by differencing. These results indicate that for linear models, i.e. with no segmentation in them, by differencing one removes trends even when in the cases which include deterministic components the resulting transformed data have not desirable properties for subsequent analysis.

Differencing in the case of models with segmentation one obtains residuals with the above properties and besides they also include a remaining part of the trend.

In order to discuss how to remove seasonality consider first a time series which has seasonal evolutivity but no trend. In this case if seasonality is deterministic a possible model is

$$x_t = \sum_{j=1}^{s} a_j^*\, S_{jt} + w_t \qquad (3.44)$$

and we remove it by running regression (3.44) and taken the residuals. This correspond to the general principle that deterministic components are remove by regression methods. If in the case under consideration seasonality is stochastic a model for x_t is

$$x_t = -\sum_{j=1}^{s-1} x_{t-j} + w_t, \qquad (3.45)$$

therefore replacing x_t by the sum of x_t values corresponding to the seasonal cycle ending at t seasonality is removed. Thus from (3.45)

$$\sum_{j=0}^{s-1} x_{t-j} = U_{s-1}(B)\, x_t = w_t. \qquad (3.46)$$

Therefore the unit-root stochastic seasonality is removed by summing over the seasonal cycle.

Seasonal sums also removed seasonality in models with determinsitic seasonality but then the resulting transformed data has bad properties. To ilustrate this result take model (3.44) for s = 2. Then $a^*_2 = -a^*_1$ and

$x_t = a^*_1 S_{1t} - a^*_1 S_{2t} + w_t$

and summing over the seasonal cycle

$$U_1(B)\, x_t = (1 + B)\, x_t = w_t + w_{t-1}. \qquad (3.47)$$

The transformed variable $(1 + L)\, x_t$ in (3.47) has no seasonality but is composed by a scheme of residuals w_t with seasonal unit roots and, as we will see later, this type of structure has bad properties for subsequent analysis.

In general time series with seasonality will also have trends. In those cases one needs to remove trend and seasonality. If both components are stochastic one needs to apply d times first differences and then a seasonal sum on the difference data. But this procedure is equivalent to apply (d-1) first or regular differences and one aditional seasonal difference. When trend or seasonality follow deterministic schemes applying regular and seasonal differences one removes them but end up with a transformed time series with bad properties. The question is that if trend and seasonality are fully deterministic one should apply the appropiate regression to eliminate both elements. For the general model I (d, m^s) with seasonality, deterministic or stochastic, one should apply first the appropiate differences and then run the corresponding regression on the difference data. See table 3.3. for a summary.

In conclusion, the general rules for removing the types of trends and seasonality discussed in previous sections are:

a). Trend and seasonality are fully stochastic with unit roots.

The stationary transformation is obtained by differencing, noting that if seasonality is present one of the differences must be seasonal.

b). Trend and seasonality are fully deterministic.

The stationary transformation is obtained by running the corresponding regression and taking the residuals.

c). Trend and seasonality have stochastic unit-root properties and deterministic components.

The stationary transformation is obtained by differencing first and then running the appropriate regression and taking the residuals.

Table 3.3

REMOVING TRENDS AND SEASONALITY. DIFFERENCING.

MODEL	PROCEDURE AND COMMENTS
(a) I (d,m), d $\neq$ 0 and m= 0 or 1	The optimal way of removing trend is by differencing d times. Examples: Models (3.3), (3.9) and (3.13).
(b) I (d,m) (SS), d $\neq$ 0 and m = 0 or 1	Proceed as case (a) but one of the differences must be seasonal. Examples models (3.23), (3.31) and (3.34).
(c) I (d, m) (DS), d $\neq$ 0 and m = 0 or 1	Proceed as case (a) and then with $\Delta^d x_t$ run a regression to remove the deterministic seasonality. Examples: models (3.18) and (3.32).
(d) I (0, m), m $\geq$ 2	The optimal way of removing trend is by running for x_t the appropiate regression and take the residuals. Example: model (2.5) Differencing (m-1) times also removes trend but leaving residuals with undesirable statistical properties.
(e) I (0, m) (DS), m $\geq$ 2	Proceed as case (d) but now the regression should include seasonal dummies. Applying (m-1) differences, one of them seasonal, trend and seasonality are removed but leaving residuals with undesirable properties.

Table 3.3 (Continuation)

REMOVING TRENDS AND SEASONALITY. DIFFERENCING.

MODEL	PROCEDURE AND COMMENTS
(f)I $(0, m^s)$, $m > 0$	The optimal way of removing trend is by running for x_t the appropiate regression and take the residuals. Example: Model (2.33). Differencing leaves residuals with similar properties as case (d) and also, and this is much worst, these residuals contain part of the trend which is not fully removed by differencing.
(g) I $(0, m^s)$ (DS), $m > 0$	Proceed as in case (f) but now the regression should include seasonal dummies. Applying differences, one of them seasonal, has all the inconvenients mentioned in case (f)
(h)I (d, m), $d \neq 0$ and $m > 1$	The proper way of removing trend is by differencing d times and running for $\Delta^d x_t$ a regression to eliminate the deterministic trend. Applying [d +(m-1)] differences eliminates trend but leaving residuals with undesirable properties.
(i)I (d, m) (SS), $d \neq 0$ and $m > 1$	Similar comments to (h) apply here, but now one of the differences is seasonal.
(j) I (d, m) (DS), $d \neq 0$ and $m > 1$	The proper way of removing trend and seasonality is by differencing d times and running for $\Delta^d x_t$ an appropiate regression to eliminate trend and seasonality.

Table 3.3 (Continuation)

REMOVING TRENDS AND SEASONALITY. DIFFERENCING.

MODEL	PROCEDURE AND COMMENTS
(k) I (d, m^s), d ≠ 0 and m> 0	The proper way of removing trend is by differencing d times and running for $\Delta^d x_t$ an appropiate regression to eliminate the deterministic trend. Appling [d + (m-1)] differences, one of them seasonal, eliminates trend and seasonality but leaving a transformed time series with bad properties for subsequent analysis. Similar comments to case (h) apply here.
(l) I (d, m^s), (SS), d ≠ 0 and m > 0	Similar comments to case (i) apply here.
(m) I (d, m^s) (DS), d ≠ 0 and m > 0	Similar comments to case (j) apply here.