Seasonal Adjustment of Housing Data

Michael Carliner

Each month data on housing activity are reported on a seasonally-adjusted basis. The seasonal adjustment procedure attempts to correct for *normal* seasonal variation and provides a basis for comparing different months on a more meaningful basis than the actual unadjusted data. However, the data remain subject to distortions as a result of *abnormal* seasonal influences or random events. The computer programs that estimate the seasonally-adjusted values attempt to distinguish normal from abnormal seasonal variation, by averaging the seasonal deviations of several years and by giving small weights to extreme values.

It should be emphasized that seasonal adjustment accounts only for normal seasonal variation, based on historical experience. There is no adjustment made for an unusual amount of snowfall, for example, in a particular month and year.

The preliminary housing starts data for December 1987 showed a sharp 16% decline from the seasonally adjusted annual rate for November 1987. For units in structures of 5 or more units, the reported decline in the seasonally adjusted rate was 37%. How much of that was an artifact of the seasonal adjustment process? If starts had been weaker in December of 1986 or 1985, would the seasonally-adjusted values for December 1987 have been significantly affected?

How much of the December 1987 drop in starts was an artifact of seasonal adjustment?

In order to evaluate the effect of recent history on the seasonally-adjusted data, we arranged an experiment, which is described below. The results indicate that the seasonally adjusted values are indeed sensitive--perhaps too sensitive--to recent history.

Seasonal Pattern of Housing Starts

The annual ebb and flow of construction activity reflects seasonal changes in weather conditions, but is also influenced by institutional factors--such as the school year schedule--that affect housing demand patterns.

Table 1 shows the percentage of annual starts that occurred in each month, on average, over the period 1983 to 1987. The strongest months for single family starts are April, May and June, a period which not only has favorable weather but which also allows homes to be completed in time for the peak demand season in September.

Table 1

Monthly Starts As a Share of Annual Starts
1983 - 1987 Average

	Total	Single Family	2 to 4 Units	5+ Units
January	6.1%	5.8%	7.2%	6.6%
February	6.2%	6.2%	6.3%	6.2%
March	8.2%	8.4%	8.9%	7.9%
April	9.7%	9.8%	10.0%	9.2%
May	10.1%	10.4%	9.5%	9.5%
June	10.1%	10.2%	9.4%	9.9%
July	9.4%	9.5%	9.8%	9.0%
August	9.2%	9.2%	9.0%	9.2%
September	8.8%	8.8%	7.6%	8.9%
October	9.0%	8.8%	8.7%	9.7%
November	7.5%	7.1%	7.3%	7.5%
December	6.1%	5.8%	6.4%	6.5%

Multifamily starts are less heavily concentrated in the Spring, in part because a Spring start for large multifamily structures does not guarantee completion in time for occupancy during the peak demand period.

In addition to weather and demand patterns, the distribution among months also reflects the number of work days available. The relatively weak shares in November through February are partly due to major holidays in that period and the fact that February has only 28 or 29 days. March, August, and October, on the other hand, have 31 days each and no major holidays.

Seasonal Adjustment - Simplified

Most government and private seasonally-adjusted data are calculated using the "X-11 Variant of Census Method II", a computer program developed in 1965 by the Census Bureau and known commonly as X-11. Stated simply, it is based on an assumption that any time series can be decomposed into a "trend-cycle" component, an "irregular" component, and a "seasonal factor" that constitutes the normal seasonal influence.

Thus,

$O_t = C_t \times S_t \times I_t$

where

O_t is the original series
C_t is the trend-cycle component
I_t is the irregular component
S_t is the seasonal factor

The attempt to distinguish the three components basically depends on constructing a series of moving averages. First a moving average of the original series is calculated. That is assumed to represent the trend-cycle component. What's left (after dividing the original series by the trend-cycle moving average) is thus the seasonal factor times the irregular,

$$\frac{O_t}{C_t} = S_t \times I_t$$

The seasonal factor is then estimated by taking a moving average across the same period in successive years, i.e., the average of the Januaries, the Februaries, etc.

With these estimates of the trend-cycle and seasonal factor components, and with the estimate of the irregular component that is left after dividing by the seasonal factor, the seasonally adjusted series can be calculated as the trend-cycle component times the irregular component or, equivalently, as the original series divided by the seasonal factor

$$A_t = C_t \times I_t = \frac{O_t}{S_t}$$

Where A_t is the seasonally-adjusted estimate.

Seasonal Adjustment - Complicated

The actual X-11 program includes a number of refinements, options, complications, diagnostic measures, and other assorted bells and whistles that go beyond the simplified procedure described here, but the basic idea is no different. The complications include:

1. An additive, rather that multiplicative, version is available. That option is useful if seasonal influences are assumed to be additive rather than proportional, and is necessary if the series includes zeroes or negative numbers.

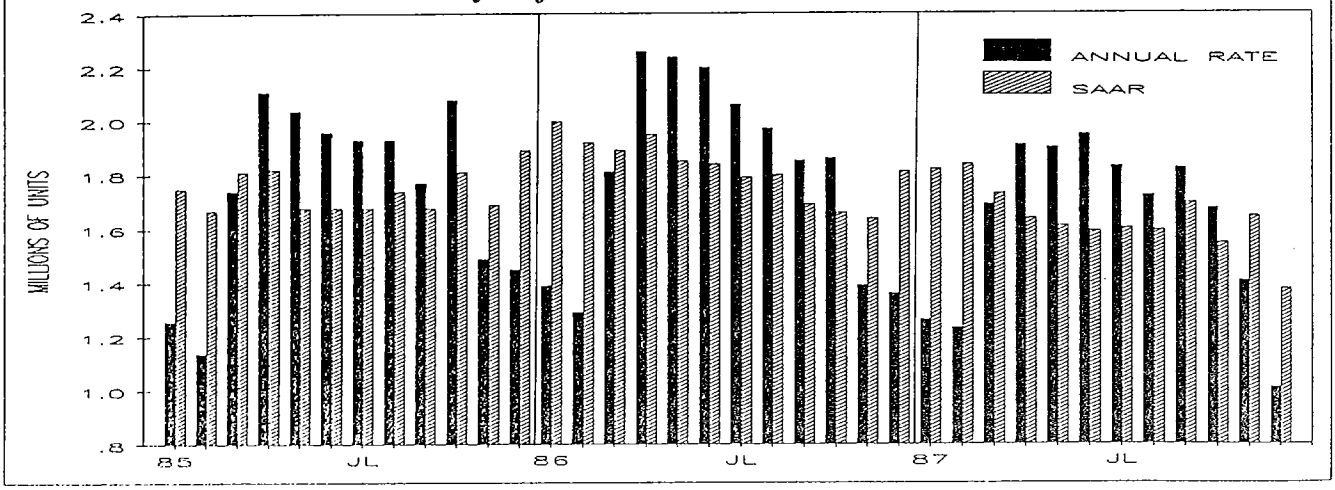
2. The moving averages used to calculate the components are centered around each month and weighted, with the greatest weight given to the center month. Since a centered average cannot be calculated for the months at the beginning and end of the series, a special weighting scheme for handling those values is included.

3. Instead of calculating the trend-cycle, seasonal factor, and irregular components in one sequence, a revised set of moving averages is calculated after the first round, in which extreme values are given reduced or zero weights so that they do not distort the estimate of normal seasonality. Whether values are considered "extreme" depends on how the irregular component for the month compares to the standard deviation of the irregular components.

4. The length of the moving average used in calculating the final estimate of the trend-cycle component depends on how smooth or irregular the series is.

5. An option is available to adjust for "trading day" variation. This refers to the fact that months don't start on the same day of the week each year, and the structure of a month, such as the number of weekends in the month, may affect activity during the month. Using regression techniques coefficients measuring the effects of each of the seven days of the week are

Private Housing Starts, Annual Rate and Seasonally Adjusted Annual Rate (SAAR), 85-87



estimated, and these coefficients are then multiplied by the number of Mondays, Tuesdays, etc. to get the trading day factor for each month.

The availability of various user-controlled options in the computer program means that the process is not strictly mechanical. Moreover, the choice of whether to adjust an aggregate (i.e., national, total) series or to adjust the components (e.g., regions or sectors), as well as other choices, introduces judgement into the process.

Seasonal Adjustment of Housing Start Data

Generally, the Census Bureau uses the standard X-11 options with the trading day adjustment to adjust starts data. Among the specific procedures used by the Bureau in producing the monthly data are:

1. The national single family starts data are created by adjusting (unpublished) regional monthly single family starts and adding together the adjusted values for the four regions.

2. Data for 2 to 4 unit structures and 5 or more unit structures are adjusted at the national level.

3. Monthly seasonally-adjusted regional total starts are created by adjusting monthly total starts in each region and then adjusting the regional values so that they add up to the national total.

4. The seasonal factors are recalculated each month, so that the factor for the current month reflects that month's experience as well as the same month in earlier years. An 11-year centered moving average is used, meaning that the latest value is affected by the same month up to 5 years earlier.

5. Unlike the calculation of the seasonal factors, the trading day factors are not recalculated every month, but only when it appears there has been a significant change in the relationship between days of the week and construction activity.

Sensitivity to Recent History

The seasonal factors for each month are calculated as a weighted moving average of the ratio of the original series to the trend-cycle component times the trading day component. The seasonal factor moving average is calculated for the current year using data for the preceding five years, with weights as follows:

Current Year (t)	.246
Year t-1	.221
Year t-2	.197
Year t-3	.173
Year t-4	.112
Year t-5	.051

The influence of each year is complicated by the fact that extreme values get reduced or zero weights, so that the remaining years become more important.

What if last year's value had been different? To see how sensitive the current value is to recent history, with the cooperation of the Census Bureau we ran an experiment using starts for 5 or more unit structures.

The actual value for December 1986 was reduced by 20% and the seasonal adjustment calculations were rerun. The result was that the seasonally adjusted rate for December 1987 was increased by 10.3%, from 303,000 to 334,000, a non-trivial impact. Moreover, although the value of actual starts for December 1986 had been reduced by 20%, the calculated seasonally-adjusted rate was reduced by only 12.9%.

One of the reasons that the experimental change in the December 1986 value of starts in structures of five or more units had such a substantial impact on the seasonally-adjusted value for December 1987 was that the changes caused the value for December 1985, which was unusually high, to be assessed as more extreme and to get reduced weight. Conversely, the change caused the value for December 1987, which was unusually low, to be assessed as less extreme and to get greater weight. Thus the lowered value for December 1986 affected not only the value for that month used in the calculation of the seasonal factor, but also the weights given values in other years.

These results illustrate the fact that the seasonal adjustment process is indeed sensitive to recent history, although the sensitivity for a particular time period depends on the complex internal operation of the X-11 program as well as the pattern of activity. The fact that the current period gets heavy weight in determining the seasonal factor means that large changes in actual starts get partly smoothed out in the seasonally-adjusted rate, as the muted impact of the experimental change in actual December 1986 starts on the December 1986 seasonally adjusted rate demonstrated.

The fact that seasonal factors are sensitive to recent history does not necessarily imply that the latest data on housing starts are distorted or meaningless. However, this analysis suggests yet another reason why the data for a single month often show erratic changes and should not be too heavily relied upon in making business or policy decisions.

If seasonal adjustment is indeed to correct for normal, consistent seasonal patterns, perhaps an average over a longer period of time, with less weight on recent history, ought to be used. The current procedure does seem a little too responsive to short-term fluctuations.