NBER WORKING PAPER SERIES

WORKDAY, HOLIDAY AND CALENDAR ADJUSTMENT WITH 21ST CENTURY DATA: MONTHLY AGGREGATES FROM DAILY DIESEL FUEL PURCHASES

Edward E. Leamer

Working Paper 16897
http://www.nber.org/papers/w16897

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2011

Support provided by Ceridian. Comments from Brian Palmer are gratefully acknowledged. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Workday, Holiday and Calendar Adjustment with 21st Century Data: Monthly Aggregates from Daily Diesel Fuel Purchases
Edward E. Leamer
NBER Working Paper No. 16897
March 2011
JEL No. C81

ABSTRACT

This paper uses a Ceridian transaction-by-transaction data set on purchases of diesel fuel by over-the-road truckers to form a monthly diesel volume purchase index from 1999 to 2011, purged of weekday, holiday and calendar effects. These high-frequency data support a new and improved set of options to correct for (1) the variability in the weekday composition of months and (2) the drift of holiday effects between months. With only monthly data, Census seasonal adjustment methods are forced to make inferences about the effects of both weekday composition and holiday drift. With daily data, these can be directly observed, and removed from the data, if the patterns repeat. But the drift of holiday effects between December and January resists statistical treatment, leaving the December/January comparison the most noisy in a seasonally adjusted monthly series. This problem, and other issues of holiday drift, can be treated with an overhaul of the calendar to put all holidays but Easter firmly in one month or another. The bottom line here is that e-recording of transactions offers a new set of opportunities for studying the health of Main Street.

Edward E. Leamer
John E. Anderson Graduate School of Management
UCLA
Box 951481
Los Angeles, CA 90095-1481
and NBER
edward.leamer@anderson.ucla.edu

Introduction

Many of the data sets that economists work with are based on voluntary after-the-fact surveys conducted by trade groups and government agencies. Others are based on involuntary after-the-fact "surveys" conducted by the IRS and other tax-collecting agencies. The incentives for timeliness and accuracy are mixed, to say the least.

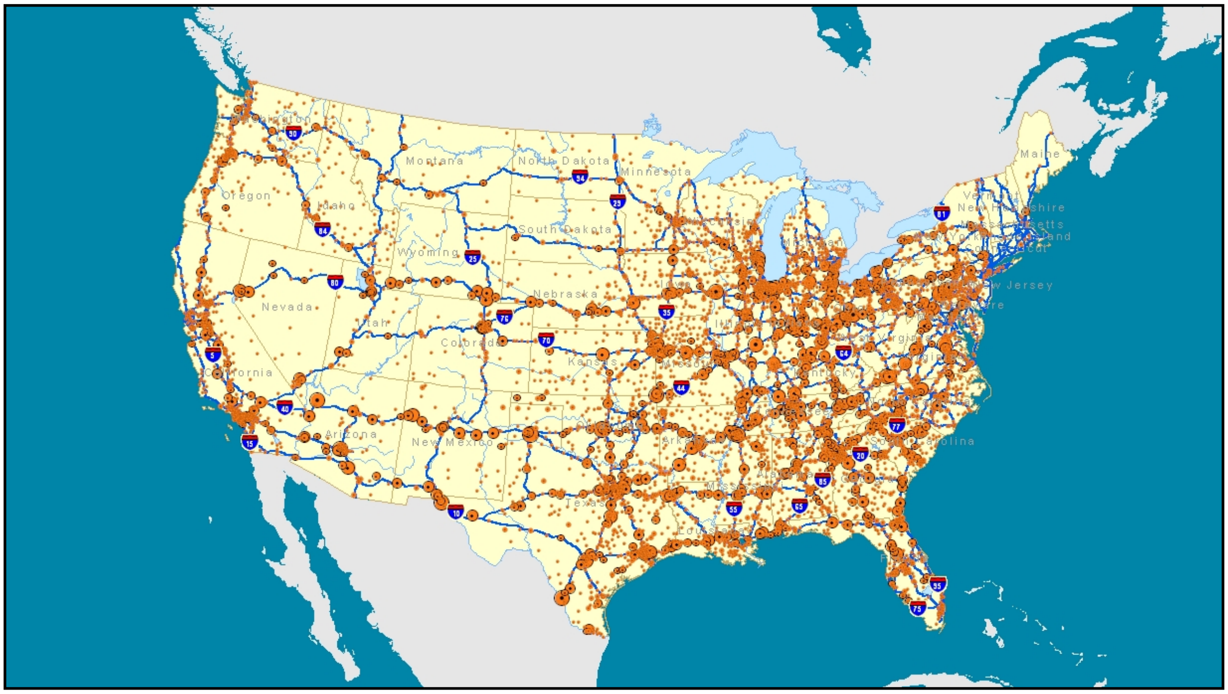
For items like retail sales and GDP, which are composed of individual transactions but treated by economists as "flows" per unit of time, the survey respondents are usually asked to perform the time aggregation themselves, reporting sales for a month or a quarter rather than individual transactions.

That's 20th Century data collection operating at a speed and accuracy worse than the US Postal Service. The 21st Century recording and communication technologies in principle allow each transaction to be reported instantaneously, something which is very familiar on Wall Street but less familiar on Main Street. Instantaneous recording of transactions creates new challenges and new opportunities for time series analysis designed to filter the noise from the signal, and thus to find the best possible answer to questions such as: How healthy was the economy?

I report here an analysis of diesel fuel purchases at 7,000 truck stops all over the United States recorded on a Ceridian transaction dataset. Large and small trucking firms employ Ceridian to provide credit cards to their employees, allowing the purchase of diesel fuel and other items, subject to the limits selected by the employer. Within a half-second after a credit card is swiped to pay for diesel fuel by over-the-road truckers, the transaction is recorded on Ceridian computers in Nashville, Tennessee.[1]

These transactions occur at distinct points in time but also at distinct points in space. The picture below locates the volume of diesel fuel purchases in the circles with area proportional to volume on a US map that includes the US Interstates. It is this image that led to the naming of the index of diesel fuel purchases: The Ceridian-UCLA Pulse of Commerce Index. The Interstates that crisscross the country are the arteries of the system and the product carried by trucks through those arteries are the life-blood of the system. Without the movement of goods, the economy "dies." Details can be found at www.ceridianindex.com.

[1] Though this is not a scientifically designed random sample, I am told that Ceridian has a large enough market share to be "representative," but in any case concerns about nonrepresentativeness are temporarily put aside as we focus here on seasonal and other related adjustments.



Since these transactions are sprinkled along the Interstates that crisscross the country, it is possible to zoom in on the manufacturing centers of the Midwest, or the ports on the East and West and Gulf Coasts, or on the Northeast when a snowstorm hits, and so on. While these data are capable of answering a wide variety of regional questions, in this manuscript I focus exclusively on the US overall and exclusively on US monthly aggregates built from daily data. I thus take the basic data to be daily aggregates of transactions for the US overall, and I use these daily data to improve upon the Census methods for removing "weekday" "holiday" and "calendar" effects from monthly aggregates. I realize that regional correlations could also play a role in filtering the noise from the monthly US aggregate. For example, there may be a different meaning when all regions experience the same change as opposed to a change in the US aggregate that is mostly due to the data from a single region. This is certainly something that should be explored, but it opens up a whole new set of challenges and opportunities in determining the best possible regional aggregates from data that are collected by position (latitude and longitude). Is there one truck stop out there that is the bellweather for the US economy? Is there one Interstate that is the critical artery? Where is the heart of America? Are there veins with empty trucks going back to load up again?

In addition to the decision to aggregate the transactions geographically, I also somewhat arbitrarily choose the task of creating monthly aggregates from daily aggregates. Why months, and why days? The survey methods traditionally used to collect economic data have probably influenced the choice of, for example, monthly frequencies for industrial production and quarterly frequencies for GDP, but with transaction-based data we are free to use any frequency that best fits our needs. I take the target to be monthly aggregates because of a personal judgment that the health of the economy changes slowly enough that variability of diesel fuel purchases within a month is mostly noise, and not useful for measuring economic illnesses. I take the daily aggregates to be the building blocks for forming monthly aggregates because of a personal judgment that variability within a day is not likely to be material to answering the question: how healthy was the economy last month? Neither of these judgments is a sure thing. It could be that weekly or bi-weekly reports would be better than monthly reports, for example, when the end of a month looks better than the beginning of a month, and it possible, but I think very unlikely, that unusual transactions within a day could signal a change in the healthiness of the economy.

The predictable changes between months in most economic time series is what allows calendar adjustment of monthly data, but calendar adjustment of weekly data or daily data remains an unsolved problem, the first step of which is to determine the day or week in one year that are the same as a day or week in other years.[2] Both disaggregation by time (more frequent than monthly) as well as disaggregation by space (regional groups) raise interesting but difficult unsolved problems that deserve attention and that could be explored with this Ceridian dataset, but these are beyond the scope of the present paper, which is targeted solely on creating monthly national aggregates.

The word "weekday" in the title of this paper refers to the fact that the volume of diesel transactions varies by day of week, more on weekdays and less on weekends Except for the 28 day (4-week) Februaries which have four of each weekday, the weekday composition of each month varies from year to year, causing problems for simple methods of calendar adjustment such as removal of monthly means, which adjusts for the fact that some months have 31 days and others have 30, but does not adjust for the variable number of weekends in any given month, depending, for example, on whether the first of the month is a Monday instead of a Saturday or Sunday. There is however a substantial though not well-known literature on this problem supported by the Census Bureau and already embedded as seasonal adjustment options in many computer packages under the title of "trading day" adjustment. The words "trading day" suggests some kind of Wall Street trading phenomenon and I think a better label is "weekday" effect. Speaking of language, for obvious reasons I will refer to "calendar" adjustment, not "seasonal" adjustment. There is hope with daily data actually to do seasonal adjustment, for example, allowing for the fact that as far as trucking is concerned the summer vacations begin the week after the 4th of July and ends in mid August. That's a season.

The literature on "trading day" adjustments dates at least to Bell and Hilmer(1983). With only monthly data these weekday effects must be inferred from unusual movements in the monthly data that are correlated with weekday composition. In contrast, the weekday effect is directly observed from our daily data and we should be able to do better than Census methods. One of the surprises for me discussed below is how well the Census has done in creating methods that accurately infer the weekday effect from the monthly diesel data, in the sense that the Census treatment for weekdays produces

[2] Weekly railroad loadings reported by Railfax are not seasonally adjusted. (http://railfax.transmatch.com/),

corrections to the monthly aggregates that are very similar to corrections based on the daily data.

After weekday variability it is holidays that have the most pronounced effect on sales of diesel fuel. The Census X12 seasonal adjustment allows corrections for Easter, Labor Day, Thanksgiving, and Christmas given an assumption of the distribution of the effect around the holiday. The daily trucking data indicate very substantial declines of sales both before and after these holidays but also around Memorial Day, the 4th of July and New Years. If these holiday-affected patterns of sales were confined to a single month or divided between two months the same way every year, then calendar adjustment of monthly data would automatically correct for the holiday. With the exception of Thanksgiving, every one of these holidays divides its impact on diesel fuel purchases between two different months in a way that varies from year to year, and holiday adjustment is needed in addition to calendar adjustment.

Although Census methods for inferring workday adjustments from monthly data works well, the same methods do not work well for inferring holiday adjustments. Our diesel data set extends for 145 months from 1999m01 to 2011m01. This turns out to be an adequate number of observations for inferring the weekday corrections, but for each holiday effect there are only 12 observations, which is not nearly enough to estimate the 7 or more parameters that describe the abnormal daily pattern of sales on and around each holiday. As it turns out, the problem is even greater than Census supposes, because the pattern of holiday effects depends on the weekday on which a holiday falls. This causes special problems for the June to July transitions which are affected by weekday of the 4th of July and for the December to January transitions which are affected by the weekday of New Years. Correction for the abnormalities in the data associated with these and other holidays substantially reduces the apparent noise in the diesel data, providing a more accurate indicator of the month-to-month changes in trucking.

Beyond the seasonal adjustment issues, a goal of this paper is to describe important features of the data set, including the very substantial variability by workday, the declines during the holidays, and also outliers that suggest weather or other events, including 9/11, that have affected the pattern of shipping. In that spirit, the first section of this paper offers a "first look at the data." Section 2 mines the daily data to determine the pattern of sales over the workweek – twice as much on Wednesday than on Saturday. Section 3 discusses regression methods for inferring the pattern of sales around the holidays. Section 4 describes the workday and holiday adjusted monthly index based on the daily data and compares it with Census methods applied to unadjusted monthly aggregates. It turns out that the daily data allow important holiday adjustments that the Census methods to not capture.

Last, in section 5, I offer the suggestion that all these problems with holidays would disappear if we adopted a different calendar with December extended into January to keep the New Years effect strictly in the December data, and with the first week of June given to May to keep the Memorial Day effect strictly in May, and with the end of June given to July, to keep the July 4th effect strictly in July, and with the last days of August moved to September to keep the Labor Day effect strictly in September. A horse race between an index based on this new calendar and the best monthly index from the daily data doesn't produce a clear winner. Parenthetically, I bring to your attention the possibility of a more radical revision of the calendar designed to minimize the information loss caused by the time aggregation. The goal might be to predict the next month's daily average given data on previous months' daily averages, with the 12 months designed to accomplish this with the highest level of average accuracy, given the statistical properties of the daily data. This would produce longer months when little is happening and shorter months where significant changes in the economy are occurring.

Bottom line here: If you desire to know how healthy trucking is at any point in time, it is essential to get the workday and holiday adjustments right, and the daily data are very valuable in accomplishing that task. And you should want to know how healthy trucking is. Healthy trucking is an essential symptom of a healthy modern economy, and the Ceridian-UCLA Pulse of Commerce helps track and forecast a number of other important monthly indicators including industrial production. Since trucking is a symptom of "inventories in motion" the PCI tracks and forecasts the volatile inventory component of GDP and also imports. Best of all, the Ceridian data are actual transactions recorded instantaneously, and all this good stuff is available for immediate gratification – no need to wait until someone fills in a survey form next month.

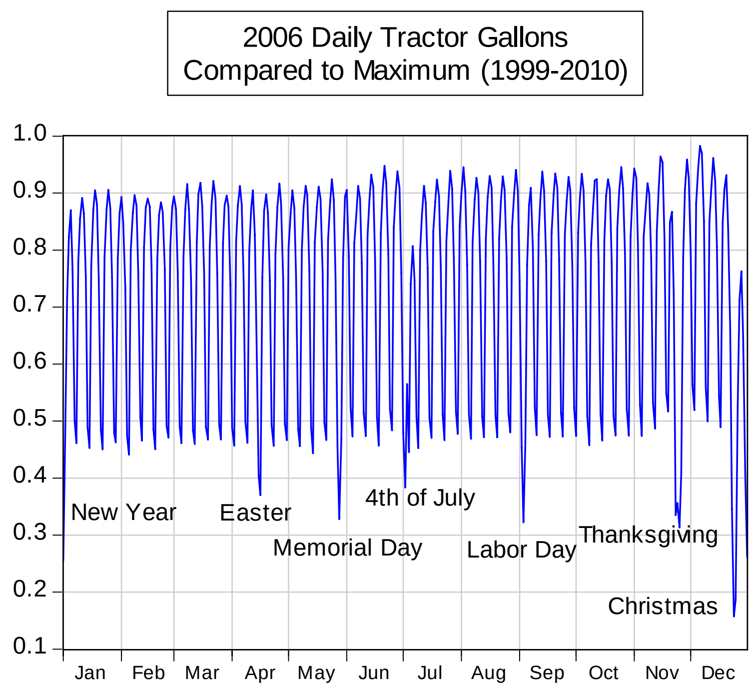
1. First Look at the Data

The first step in the analysis is to take a close look at the data to identify what are the biggest adjustment issues.

Weekday Effect

Daily diesel fuel purchases in 2006 divided by the series maximum which occurred on March 5, 2008 are illustrated in Figure 1. For economists used to looking at images of monthly data this is an astonishing display. By far the most salient feature is the rhythmic swing up and down by 50%, 52 times in 2006. This is the variability within the week: diesel fuel purchases are twice as high on Wednesday and as on Sunday. This matters for the monthly data because the number of Sundays (or other days) can vary for the same month from year to year and cause variability in monthly totals that is not picked up by seasonal adjustment methods that do not include weekday corrections.

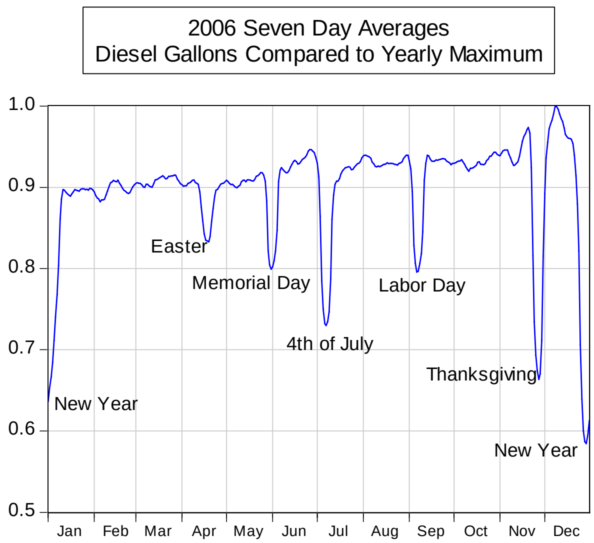
Figure 1 **2006 Tractor Gallons**



Holiday Effects

After the weekly swings, the next most obvious features of Figure 1 are the extreme dips that correspond with seven major holidays. Though extreme, these dips are disguised by the similarly extreme swings in sales over the weekdays. To make the holidays transparent, Figure 2 illustrates the same 2006 data but with the weekday effect removed via 7-day averages, each of which includes one and only one of each weekday. Here we see that the holiday with the greatest effect on diesel fuel purchases is the combined Christmas/New Years week, with sales about 40% below normal. After that come Thanksgiving, the 4th of July, Labor Day, Memorial Day and Easter, in that order.

Figure 2 2006 Seven-Day Averages



Simple seasonal adjustment can deal with these holiday effects if they are confined to a single month, or if the effects are divided between months in the same way, from year to year. Easter is an obvious problem since it varies between March and April. Census seasonal adjustment routines, X11 and X12, include both weekday and holiday adjustments. Here is what Census says about the holidays, thinking I suppose of retail sales, and not necessarily trucking:

> US Census(2009) pages 32-33
>
> Holiday effects (in a monthly flow series) arise from holidays whose dates vary over time if (i) the activity measured by the series regularly increases or decreases around the date of the holiday, and (ii) this differentially affects two (or more) months depending on the date the holiday occurs each year. (Effects of holidays with a fixed date, such as Christmas, are indistinguishable from fixed calendar effects.) Easter effects are the most frequently found holiday effects in U.S. economic time series, since the date of Easter Sunday varies between March 22 and April 25. Labor Day and Thanksgiving also are potential, though less common, sources of holiday effects.

The moving Memorial Day effect

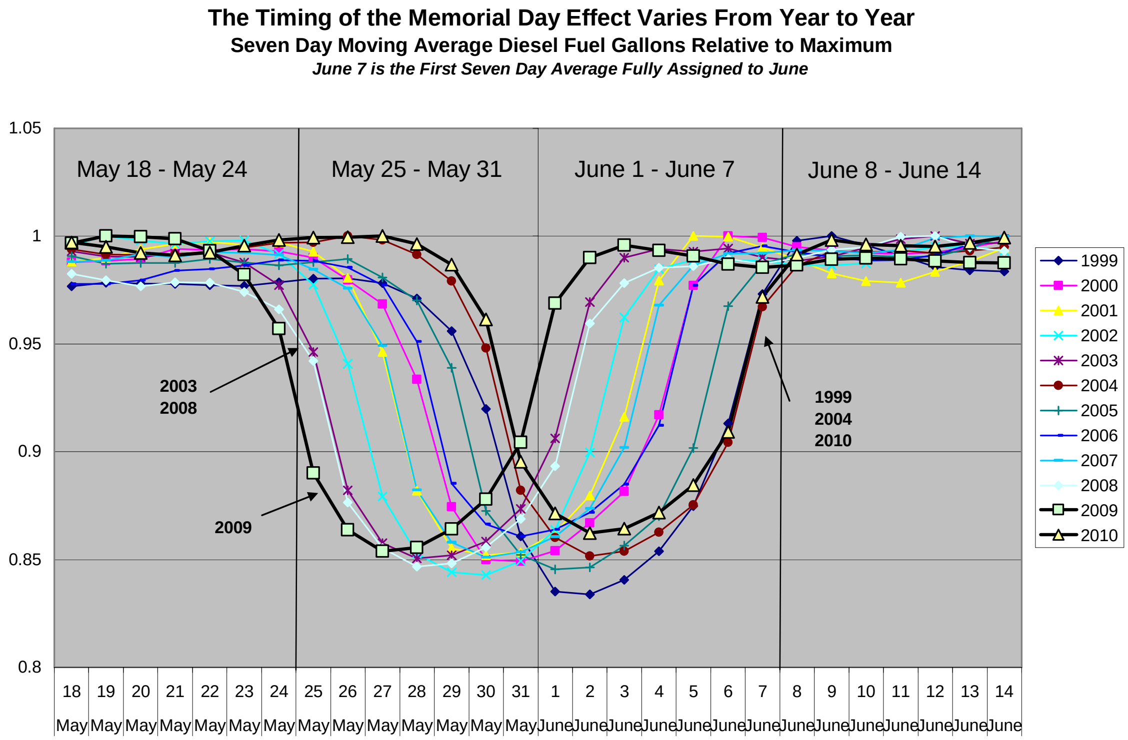
Though Census statisticians have concentrated their attention mostly on Easter, and somewhat on Labor Day and Thanksgiving, there are similar problems for other holidays that are sometimes near the transitions between months and sometimes in the middle of a month. For example Memorial Day has an depressing effect on diesel fuel purchases that is usually confined to May but leaks into June when the Monday Memorial Day holiday is the last day of May, as it was in 1999, 2004 and 2010, or next-to-last day, as it was in 2005.

Table 1 Date of Memorial Day

Memorial Day	Year
25-May	2009
26-May	2003
26-May	2008
27-May	2002
28-May	2001
28-May	2007
29-May	2000
29-May	2006
30-May	2005
31-May	1999
31-May	2004
31-May	2010

Table 1 indicates dates of Memorial Day for each year in the sample, sorted by the date, from earliest to latest. It is a late Memorial Day that allows the Memorial Day effect to leak into June. Figure 3 illustrates seven-day average sales for the four weeks from May 18 until June 14. The timing of these Memorial Day waves coincides exactly with the timing of Memorial Day reported in Table 1, first 2009, then 2003 and 2008, and so on. These weekly values bottom out about 15% lower at the minimum around Memorial Day then at the normal periods a week or two later, or earlier. These minima sometimes occur in the last week of May and sometimes in the first week of June. The first seven-day average that is exclusively in June occurs on June 7, which is where to look for a Memorial Day effect that is leaking into June. This is clearly the case for 1999, 2004 and 2010, the three instances when Memorial Day was Monday, May 31st. The value of this weekly average is about 3% lower than the subsequent data and the corresponding value on May 31 (the seven-day average of the last week of May) is about 3% higher. Thus absent adjustment for the timing of Memorial day, the May data are about 3%/4 = 0.75% underestimated and the June data are 0.75% overestimated, making the May-June growth overstated by 1.5%, which is an intolerably large error for a series with average monthly growth of only 3%/12 = .25% and a standard error of 2.8% after correction for month but not workdays or holidays.

Figure 3 **Moving Memorial Day Effect**



Easter, the 4th of July, Labor Day, Thanksgiving, and New Years all need attention too

Easter creates an obvious problem for seasonal adjustment since it is sometimes in March and sometimes in April. We have just discovered that movement of a holiday within a month (Memorial Day) creates problems that need treatment also Labor day and Thanksgiving also drift within their respective months.. Furthermore, the Census assertion that "Effects of holidays with a fixed date, such as Christmas, are indistinguishable from fixed seasonal effects" ignores the fact that we will make clear below that the weekday on which the 1st of January or the 4th of July falls affects the pattern of sales around these holidays, and the division of sales between December/January and June/July.

2. Weekday Adjustments

Census statisticians have devised methods using monthly data to infer the weekday effects from annual differences in monthly values when these differences are correlated with the annual changes in the weekday composition of the months. But our daily data allow us a direct measure of the weekday effect. The workday correction used here is based on 52-week centered averages for each weekday. Although almost a full year of data, the holidays included in any 52 week average vary over time and these create unacceptable jumps in the 52-week average. This is cured by eliminating from the averages all the major holidays and also days within three of a holiday, thus eliminating most of the days affected by holidays. There are also some outliers that are great enough to affect these averages. These outliers are of interest beyond the problems they cause for workday corrections. We turn first to them.

Non-holiday Daily Outliers

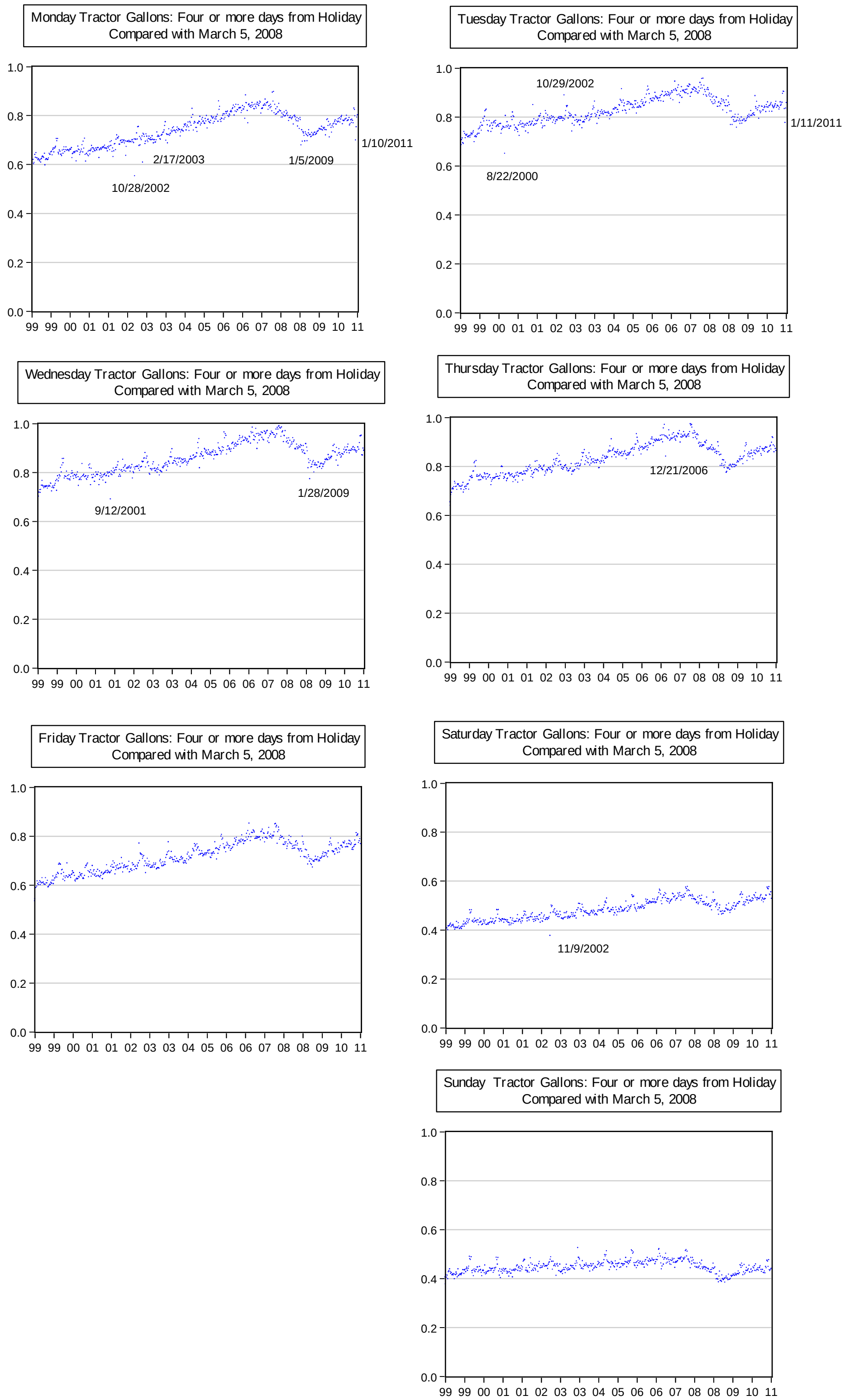
Figure 4 illustrates the gallons data *separately* for each day of the week, with the seven major holidays and three days on each side excluded. The shape swept out from 1999 to 2010 by the preponderance of the data is the same for each of the days. Wednesday is when the diesel fuel purchases are greatest. Sunday has the least gallons purchased.

Although most of these data hug together, there are some extreme outliers for most days of the week, with sales differing by 10% or more compared with the same day a week earlier or later. These nonholiday extreme outliers are listed in the table below. We will return to the issue of outliers after further statistical adjustments.

Table 2 Non-holiday Outliers

Negative Outliers		
Monday	10/28/2002	Offset the next day
	2/17/2003	President's Day Category 4 Snowstorm Weekend: 2/15 – 2/18
Tuesday	8/22/2000	
Wednesday	9/12/2001	The day after 9/11
Thursday	12/21/2006	
Saturday	11/9/2002	
Positive Outliers		
Tuesday	10/29/2002	Combines with 10/28/2002

Figure 4 Tractor Gallons by Day of Week, Days within 3 of Holiday Excluded



Weekday Adjustment Factors

The Census weekday adjustments are inferred from the correlation of monthly sales with monthly weekday composition. Since we have the direct daily sales, we don't need to infer the weekday effect. We can directly observe it. Our only problems are how to smooth the data and how to deal with holidays and outliers. We will simply omit the holidays and outliers, and smooth using a 52 week moving average.

Weekday adjustment factors from January 1999 to June 2010 are illustrated in Figure 5. These are 52 week centered averages of diesel sales for each weekday relative to the 52 week moving average of Wednesday sales, with holidays and outliers omitted. The logic for the 52 week average is that this covers a whole year and is not contaminated by patterns that are confined to specific months, and with 52 observations in each average we obtain the kind of slow-moving adjustment appropriate to something like the weekly pattern of sales. To maintain 52 weeks in every average, the first six months and the last six months are held flat, equal to the 1999 and 2010 averages respectively.

Notice in Figure 5 that Tuesday and Thursday have been competing for second place in the race for the most diesel sales, with Thursday clearly winning out at the end of the series. Next in the race at positions 4 and 5 are the weekdays Monday and Friday. These four days have sales that are all over 80% of Wednesday sales. The very slow sales days are Saturday (60% of Wednesday in 2010) and Sunday(50% of Wednesday in 2010). Though Saturday and Sunday sales were essentially the same in 1999, a substantial drift downward of the Sunday sales begins in 2003, and is offset by increases in Monday, Friday and Saturday sales. I am tempted to suggest this is the rising influence of Christianity.

We are now in a position to do a weekday adjustment by dividing the data by the weekday factors illustrated in Figure 5. With this correction, the unadjusted tractor gallons illustrated in Figure 1 are turned into the weekday-adjusted data illustrated in Figure 6 where the seven holidays are clearly present, but otherwise the daily data seem pretty smooth, almost outlier free. But this figure disguises whatever outliers may be present by choice of scale that is coarse enough to allow the holidays to be displayed. More on outliers below.

Figure 5 **Weekday Adjustment Factors**

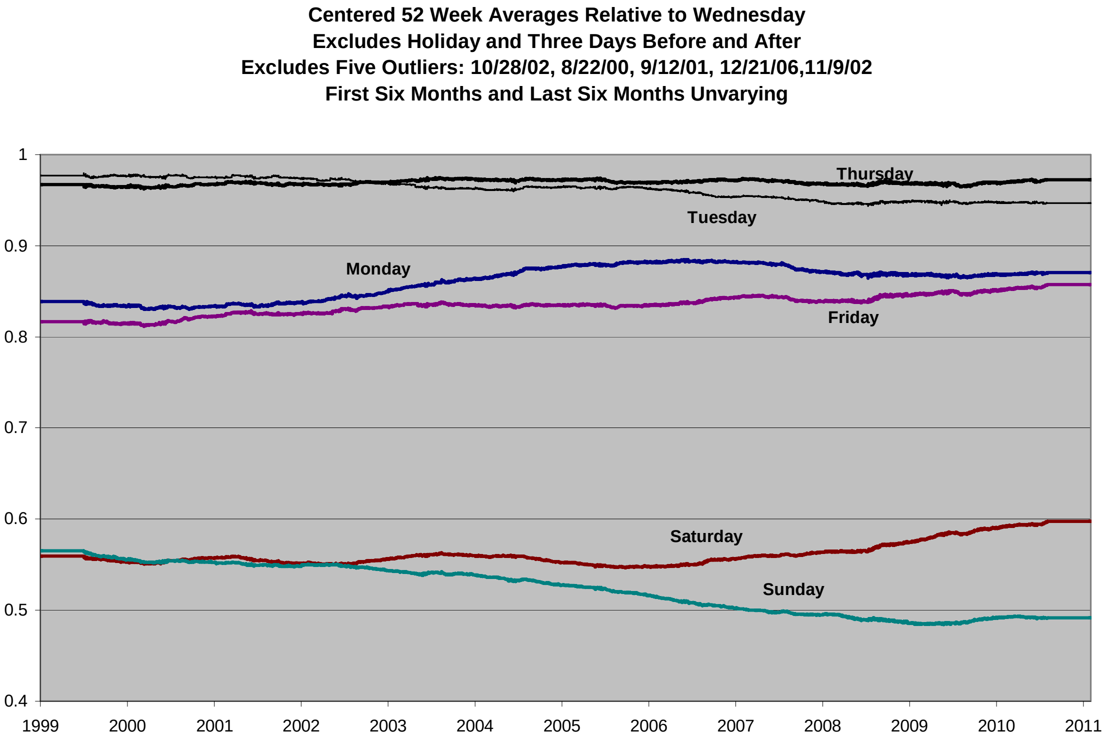
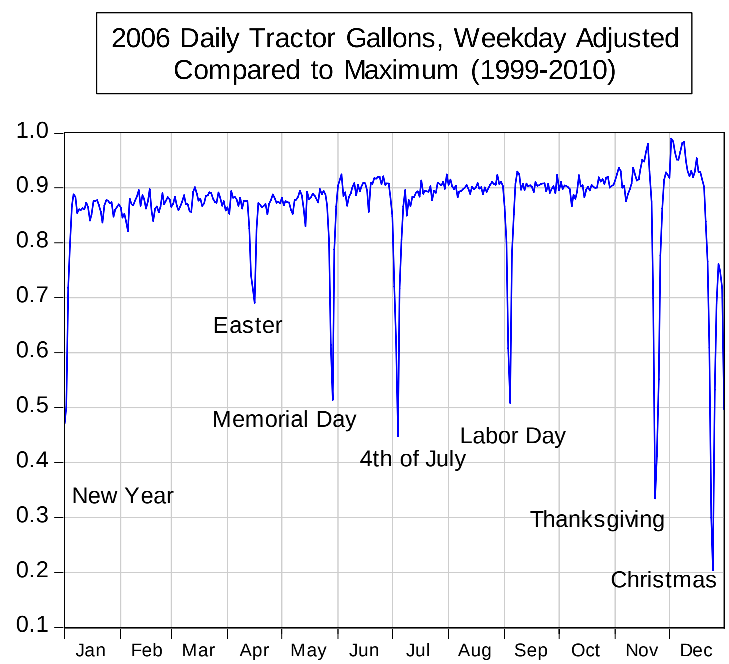


Figure 6 **2006 Daily Data, Workday Adjusted**



3. Holiday Adjustments

Now that the data have been purged of the weekday effects, the next step is to remove the holiday effects. Figure 7 through Figure 12 illustrate the weekday adjusted data around the holiday periods. Specific comments on each holiday are provided below. These figures make it seem that holiday adjustments will have large effects on the Easter-affected months, March and April, and on pairs of months surrounding the holidays: May and June because of Memorial Day, August and September because of Labor Day, November and December because of Thanksgiving, and December and January because of New Years.

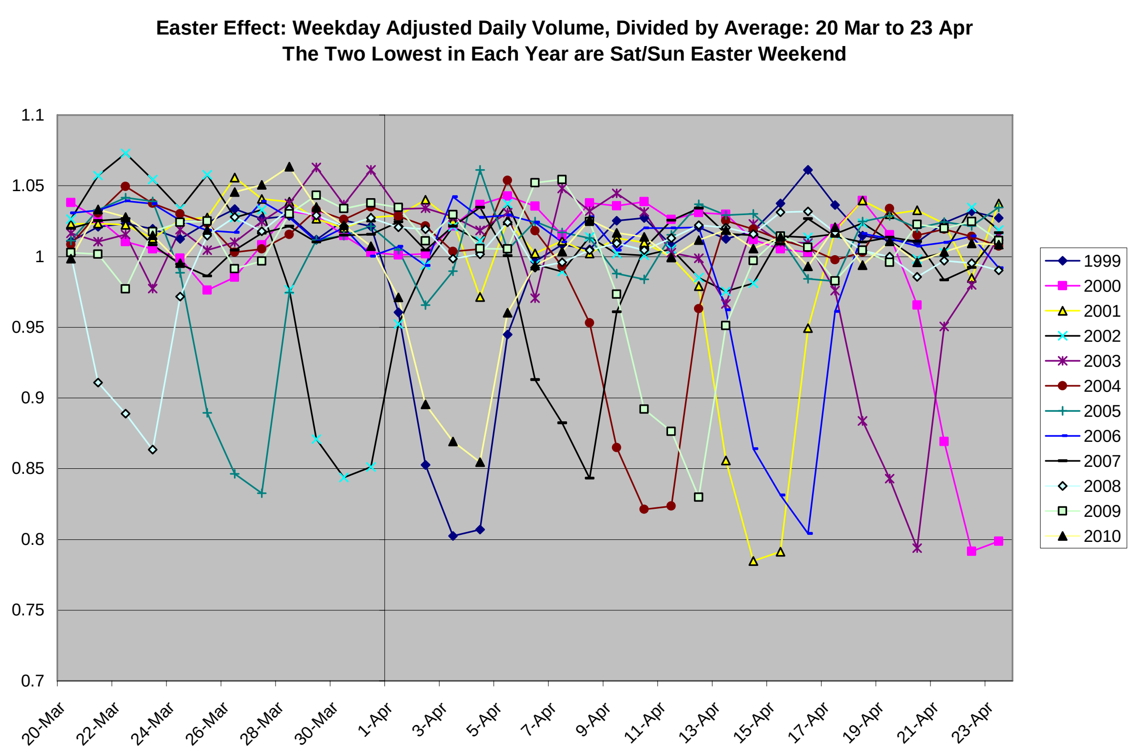
Images of Weekday Adjusted Sales Around Holidays

Easter Weekday Adjusted Daily Sales

As illustrated in Figure 7 the effect of Easter on diesel volumes wanders back and forth between February and March, which is well understood. But there are at least a couple of other interesting features evident in the display.

- The Friday before Easter is more depressed than the Monday afterward, though this is after accounting for the fact that Friday is normally a bit lower than Monday.
- The effect of Easter seems to increase with the calendar date, more in April than in March.

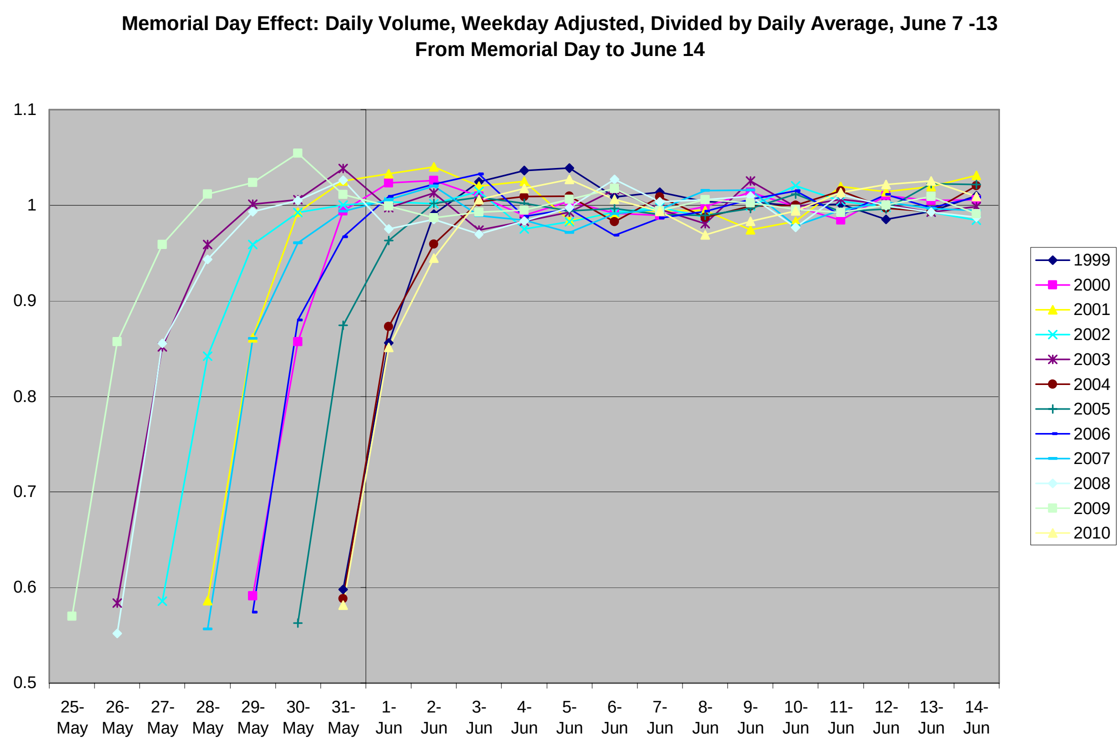
Figure 7 Easter: Weekday Adjusted Daily Sales



Memorial Day Weekday Adjusted Daily Sales

When Memorial Day Monday is May 31, diesel purchases are depressed by over 10% on the Tuesday, June 1, and depressed about 4% on June 2. When Memorial Day is May 29, there is also an effect on sales on Wednesday June 1. For earlier Memorial Days, there is no apparent leakage into June. The similarity in the pattern across years is what allows straightforward and effective adjustment for this holiday.

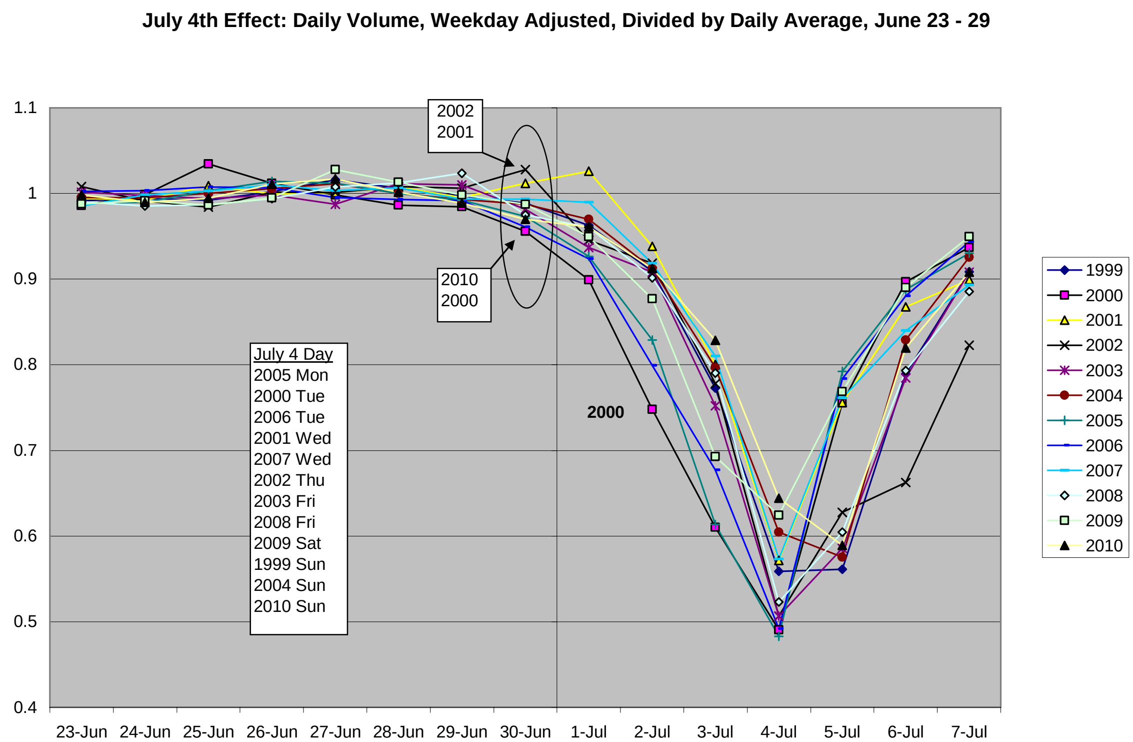
Figure 8 Memorial Day: Weekday Adjusted Daily Sales



July 4: Weekday Adjusted Daily Sales

Sales on the 4th of July illustrated in Figure 9 dip about 50% from their late June levels but this varies across years by about 10%. This noise doesn't go deep into June, but the circled date of .June 30 has quite a bit of variability, suggesting some leakage of the July 4th effect into June. If 5% of June 30 sales should have been booked in June, this is making the June number only 5%/30 days= 0.17 % higher than it should have been, which is not a large number compared with the other unexplained noise. Furthermore, the pattern on June 30 across years doesn't very clearly associate with the weekday of July 4. But the figure does reveal that a Sunday 4th of July is followed by an equally depressed level of Monday sales. While that might matter for some calculations, these event is restricted to July and thus is not a cause of leakage into another month.

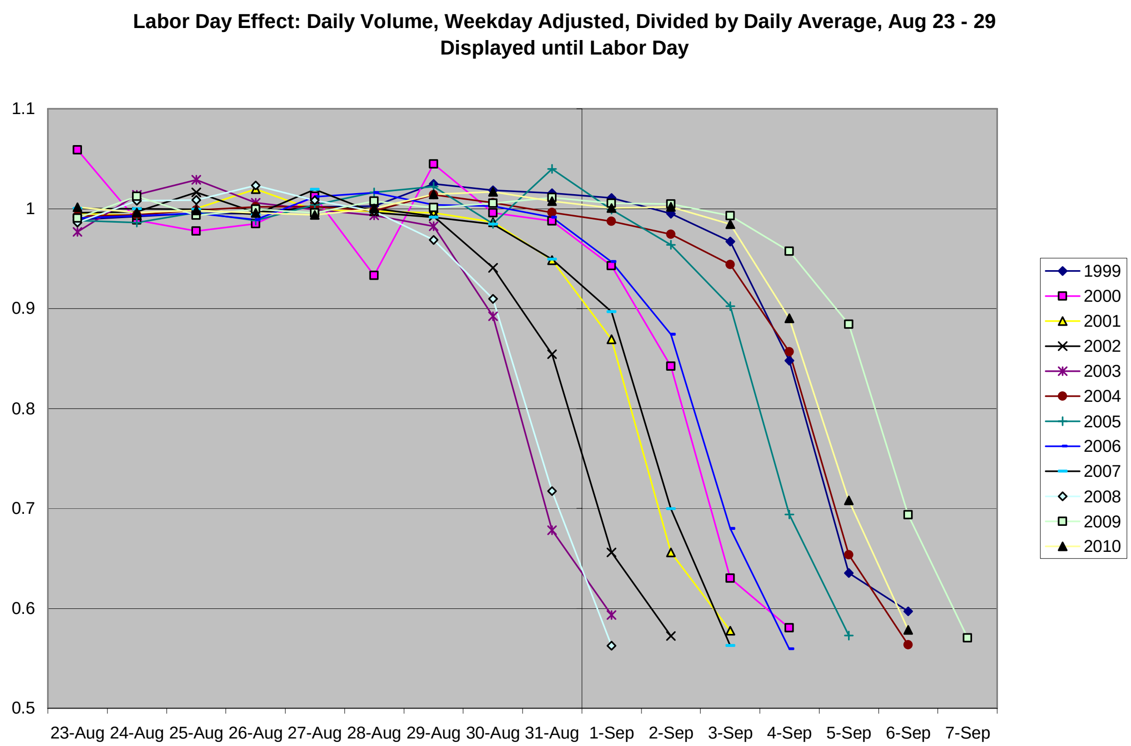
Figure 9 July 4: Weekday Adjusted Daily Sales



Labor Day: Weekday Adjusted Daily Sales

The basic pattern of sales before Labor Day appears very similar regardless of the date of Labor Monday, but when Labor Day is early in September, the effect clearly leaks into August.

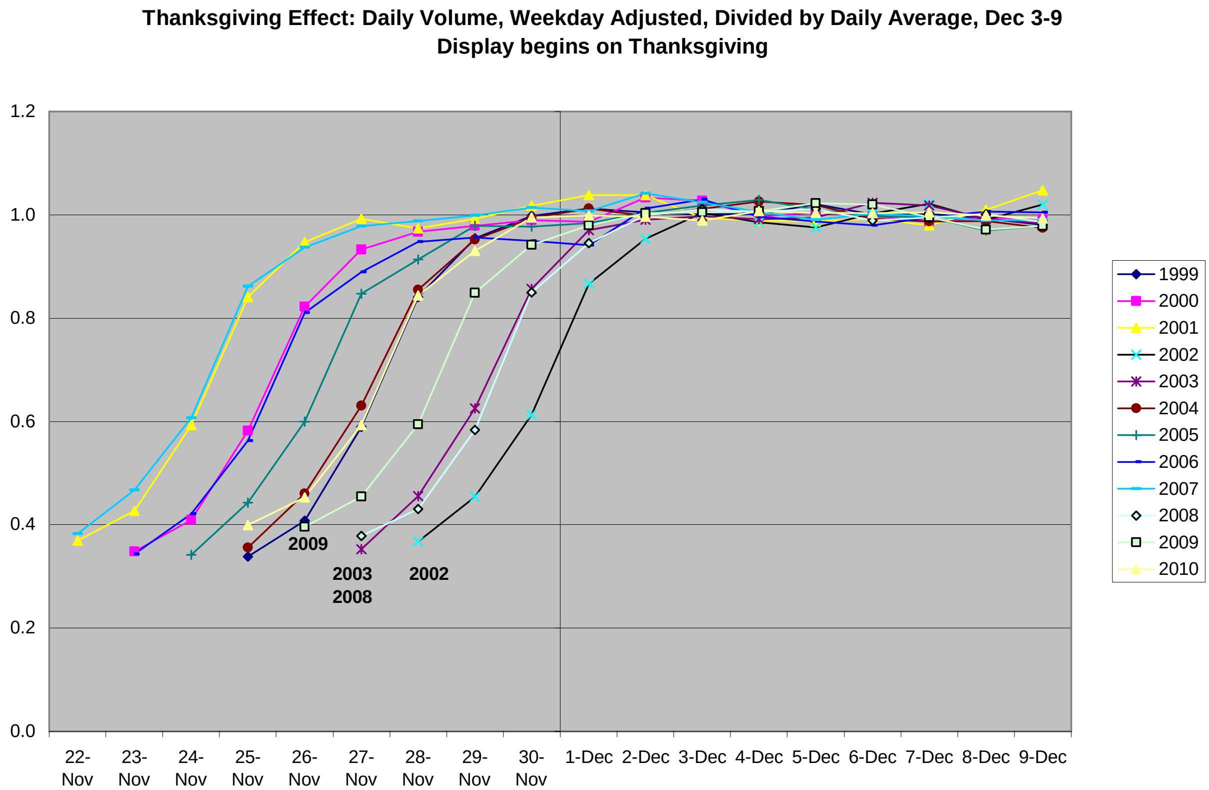
Figure 10 **Labor Day: Weekday Adjusted Daily Sales**



Thanksgiving, Weekday Adjusted Daily Sales

The pattern of diesel sales following Thanksgiving seems not affected by the date on which Thanksgiving occurs, but when Thanksgiving is late enough in November, the effect leaks into the first couple of days in December. This is most evident in 2002, less so in 2003/2008 and not much otherwise.

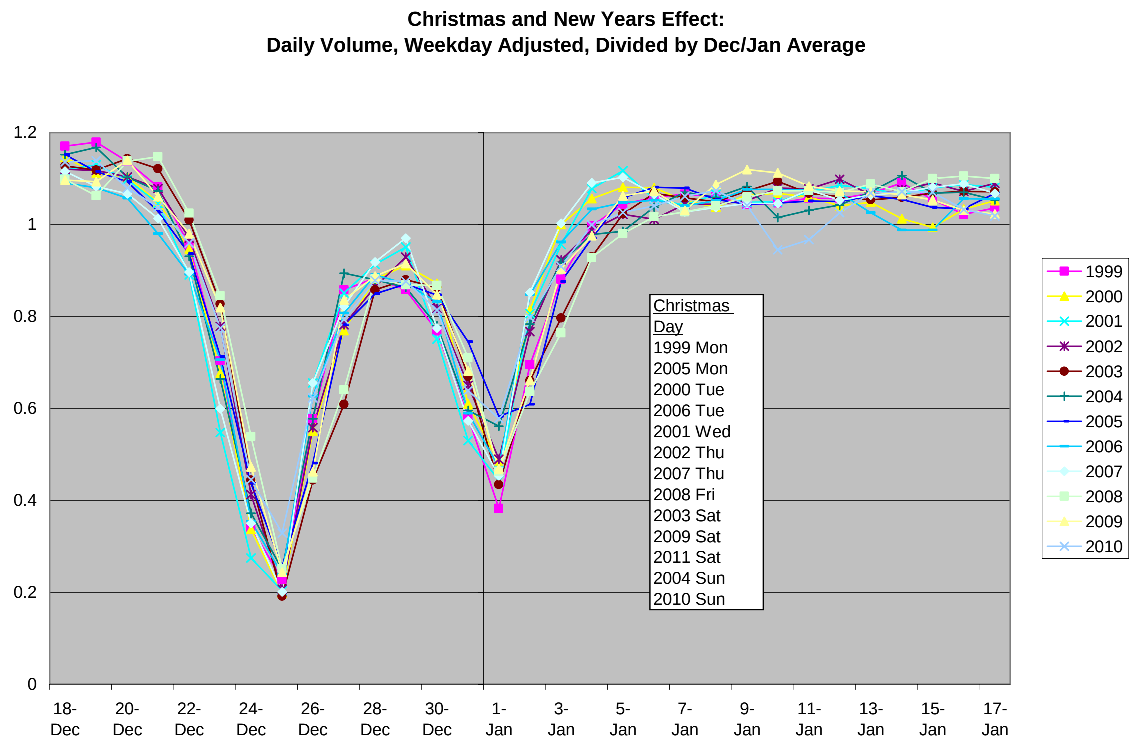
Figure 11 Thanksgiving, Weekday Adjusted Daily Sales



Christmas and New Years, Weekday Adjusted Daily Sales

There is a substantial amount of variability of diesel sales volumes between Christmas and New Years and in the first week of January, but no weekday effect is apparent in the figure.

Figure 12 Christmas and New Years, Weekday Adjusted Daily Sales



Linear Regression for Holiday Adjustment

Linear regression is the standard tool for doing both weekday and holiday adjustments. The models estimated with the daily data take the form

$$\log(gallons_t / weekday\_adj_t) = \alpha + \sum_{i=-5}^{5} \beta_i D_{it} + \varepsilon_{it}$$

where the variable *weekday_adj* is illustrated in Figure 5, where the coefficients $\beta_{-5}, \beta_{-4}, \ldots, \beta_5$ are holiday effects starting 5 days before the holiday and ending five days after, and D_{it} are binary variables equal to 1 if t-i is the holiday and equal to zero otherwise. While this is entirely conventional way to deal with holidays and weekdays, treatment of the nonholiday part of the variability in gallons, ε_{it}, is more a matter of discretion. Census seasonal methods include options that correct for ARIMA structures on these residuals within the context of seasonal adjustment of monthly data (or other frequencies). Here we pursue a two-step approach made possible by the availability of daily data – we first remove the workday and holiday effects from the daily data and then do the seasonal adjustment on monthly aggregates.

Given the regression coefficients, the data purged of weekday and holiday effects take the form

$$(gallons_t / weekday\_adj_t) / \exp(\sum_{i=-5}^{5} \beta_i D_{it})$$
$$= \begin{cases} (gallons_t / weekday\_adj_t) & \text{for nonholidays} \\ (gallons_t / weekday\_adj_t) / \exp(\beta_i) & \text{for holidays} \end{cases}$$

Once these daily data are purged of weekday and holiday effects, they are aggregated into monthly averages, and then subjected to Census X12 to remove the calendar effects.[3]

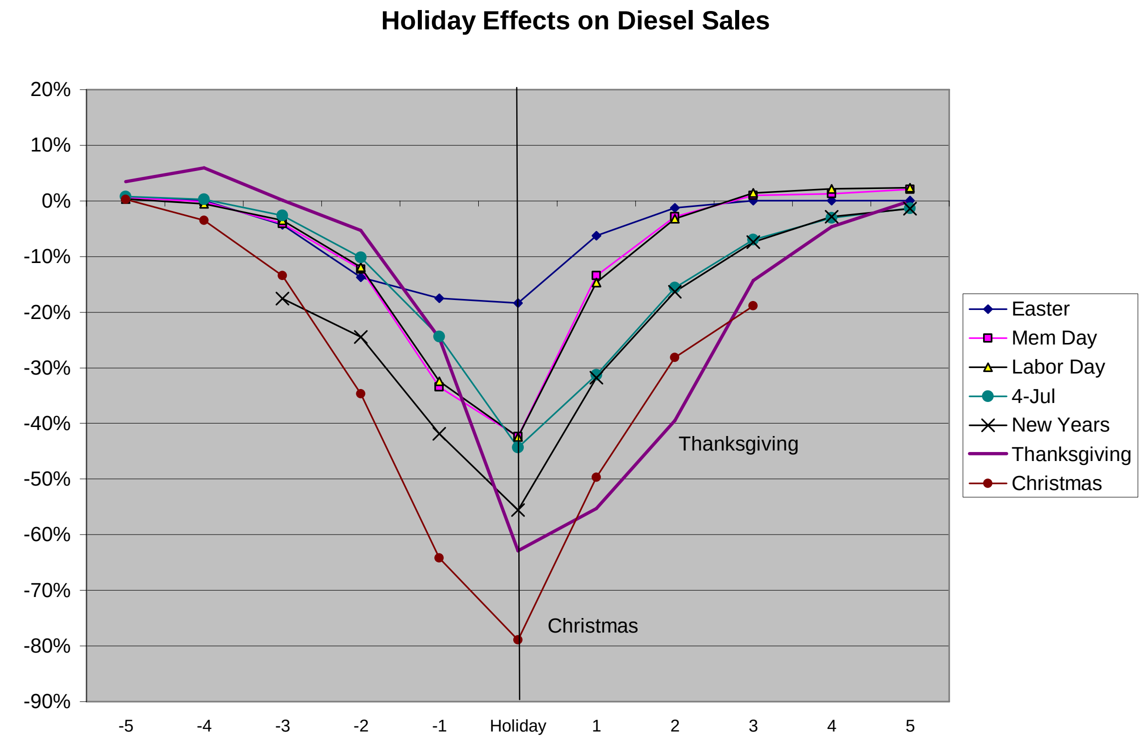
First Attempt: EQ00

Table 13 in an appendix reports a regression with ARMA residuals explaining the logarithm of weekday-adjusted daily data as a function of indicators selecting each of the prominent holidays and five days on either side of each. The Christmas and New Years indicators do not extend the full range because there are only 6 days between these two holidays. The estimated coefficients on the holiday indicators are translated into holiday percentage adjustment factors using the formula exp(coeff)-1. These are all illustrated in Figure 13, which reveals that Christmas has the largest effect (-80%), followed by Thanksgiving(-63%) which has a distinct asymmetric effect, more after than before. Probably this is trucking preparing for Black Friday sales and taking it a bit easy in the

[3] Incidentally, if x is normally distributed with mean μ and variance σ^2 then the expected value of exp(-x) is $\exp(-\mu+\sigma^2/2)$. This could serve as a basis for shrinkage of the holiday effect when the uncertainty in the coefficient is large.

week after Thanksgiving.   Next in magnitude is New Years (-56%), then Memorial Day and Labor Day which have almost exactly the same pattern of effects. Last in terms of its effect on trucking is Easter(-18%), though remember this is on top of a 60% Sunday effect.

Figure 13 **EQ00: Estimated Holiday Effects on Weekday Adjusted Diesel Sales**



EQ00 Residuals

The holiday part of the regression equation can be used to remove the holiday effects from the weekday-adjusted daily data by dividing the daily data by the exponent of the holiday coefficient, a divisor which is constant beyond five days from the holidays. This works well for the holidays that are always on the same weekday but not so well for the 4th of July, Christmas and New Years which fall on different weekdays in different years. This section lays the groundwork for equations that allow the holiday effect to be weekday dependent, first by exploring the outliers in the EQ00 residuals and second by exploring graphically the weekday patterns of the adjusted data around the holidays.

Outliers in EQ00 Residuals

Table 3 reports the residuals from EQ00 that are larger than 0.1 (10%) in absolute value. Four of these outliers occur during high-impact Northeast snowstorms listed in Table 4. One is on the Wednesday after the terrorist attack on 9/11/2001. But most of these outliers occur late in December, early in January and early in July, close to Christmas, New Years and the 4th of July. These are the problem holidays that are celebrated on varying days of the week.

Our daily data allow us to explore the impact of snowstorms on trucking, and in principal to add weather adjustments to the equation. Figure 14 illustrates the NESIS values per NOAA for each of the greatest Northeast snowstorms, and the corresponding residuals from EQ00 during the snowstorms and one day extra at each end, possibly capturing preparations or postponed travel.

Table 3 EQ00 Large Residuals

EQ00 Residuals greater than 0.1 in absolute value

December/January Transitions

	Weekday	Value	
12/24/1999	Fri	-0.146	
1/1/2000	Sat	-0.211	
12/24/2000	Sun	-0.141	
12/23/2001	Sun	-0.248	
12/24/2001	Mon	-0.218	
12/26/2001	Wed	0.217	
12/25/2002	Wed	-0.142	
12/25/2003	Thu	-0.265	
12/26/2003	Fri	-0.120	
12/27/2003	Sat	-0.157	
12/27/2004	Mon	0.104	
1/1/2005	Sat	0.152	
12/26/2005	Mon	-0.198	
12/31/2005	Sat	0.111	
1/2/2006	Mon	-0.262	
12/26/2006	Tue	0.112	
1/2/2007	Tue	0.121	
12/23/2007	Sun	-0.153	
12/26/2007	Wed	0.249	
1/2/2008	Wed	0.132	
12/23/2008	Tue	0.120	
12/24/2008	Wed	0.215	
12/26/2008	Fri	-0.250	
12/27/2008	Sat	-0.109	
1/1/2009	Thu	-0.107	
12/23/2009	Wed	0.113	1
12/24/2009	Thu	0.109	
12/26/2009	Sat	-0.202	
12/27/2009	Sun	0.147	
12/25/2010	Sat	0.275	2

July 4 Problems

	Weekday	Value
7/5/1999	Mon	-0.20
7/2/2000	Sun	-0.15
7/3/2000	Mon	-0.12
7/5/2000	Wed	0.16
7/4/2002	Thu	-0.11
7/6/2002	Sat	-0.19
7/5/2003	Sat	-0.11
7/5/2004	Mon	-0.22
7/3/2005	Sun	-0.16
7/5/2005	Tue	0.22
7/5/2006	Wed	0.20
7/4/2009	Sat	0.16
7/5/2010	Mon	-0.23

Other Problems

8/22/2000	Tue	-0.16	
8/23/2000	Wed	0.16	
8/29/2000	Tue	0.12	
12/13/2000	Wed	-0.10	
9/12/2001	Wed	-0.19	9/11
10/28/2002	Mon	-0.24	
10/29/2002	Tue	0.24	
11/8/2002	Fri	0.10	
11/9/2002	Sat	-0.24	
11/10/2002	Sun	0.13	
2/17/2003	Mon	-0.11	3
11/17/2009	Tue	0.10	
1/10/2011	Mon	-0.12	4

High Impact Northeast Snowstorms

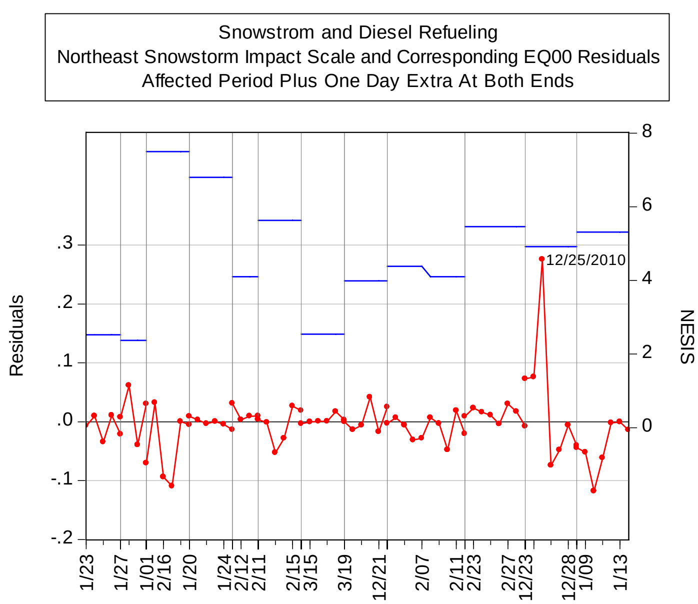
1 18-21 December 2009
2 24-28 December 2010 (preliminary)
3 15-18 February 2003
4 9-13 January 2011 (preliminary)

Table 4 **Northeast Snowstorms (NOAA)**

High-impact snowstorms that affected the Northeast urban corridor, 1999-2011
NESIS = NorthEast Storm Impact Scale

			NESIS	Category	Description
2000	January	24-26	2.52	2	Significant
2000	December	30-31	2.37	1	Notable
2003	February	15-18	7.5	4	Crippling
2005	January	21-24	6.8	4	Crippling
2006	February	12-13	4.1	3	Major
2007	February	12-15	5.63	3	Major
2007	March	15-18	2.54	2	Significant
2009	March	1-3	1.59	1	Notable
2009	December	18-21	3.99	2	Significant
2010	February	23-28	5.46	3	Major
2010	February	4-7	4.38	3	Major
2010	February	9-11	4.1	3	Major
2010	December	24-28	4.92	3	Major
2011	January	9-13	5.31	3	Major

Figure 14 **EQ00 residuals during High-Impact Northeast Snowstorms**



Two Problematic Monthly Transitions: Dec/Jan and June/July

As a first step in exploring the possibility that the noise around the holidays can be understood, Figure 15 displays January seven day averages of the EQ00 workday and holiday adjusted data, beginning on the 7th of January and ending on January 31. The 7th is the first seven-day average that uses data exclusively for the month displayed. Transitional problems between the months are thus suggested when either the 7th is an apparent outlier or the last day of the month is an apparent outlier. Figure 15 has apparent January 7 positive outliers for 2008, 2011 and other years, suggesting that some of sales that traditionally would have been booked in December were at those times booked in January. If so, it will make January appear stronger and December appear weaker. This can matter greatly. January 2008 was the first NBER-official recession month of the Great Recession. If we want the recession alarm to ring early and loud, it should be ringing about that time. We need to get this one right.

The same kind of graphs for all 12 months are collected together in Figure 16. Visually, it appears that January, July and December are the problem months.

Figure 15 January Seven-Day Moving Average of EQ00 Adjusted Data

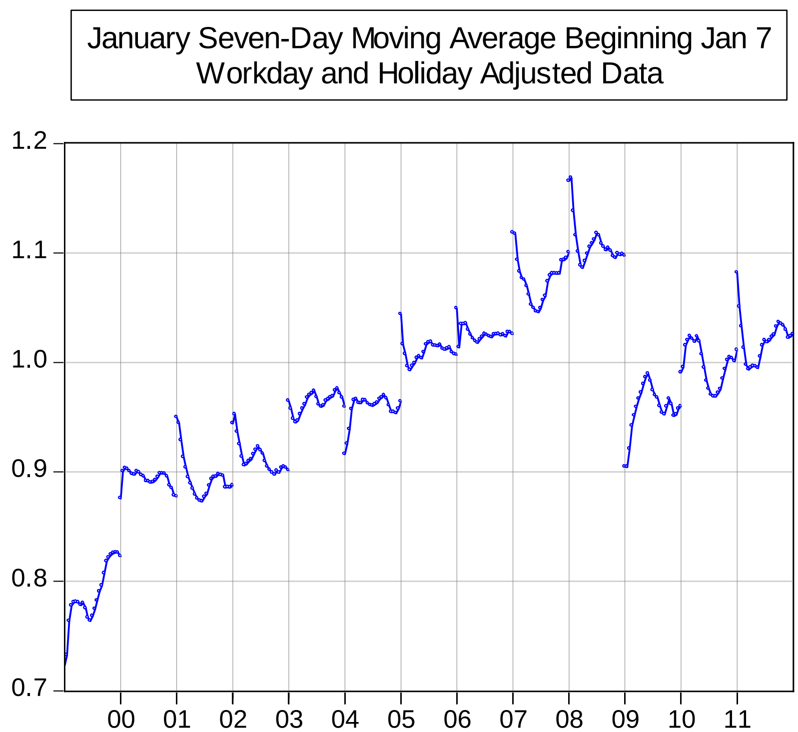
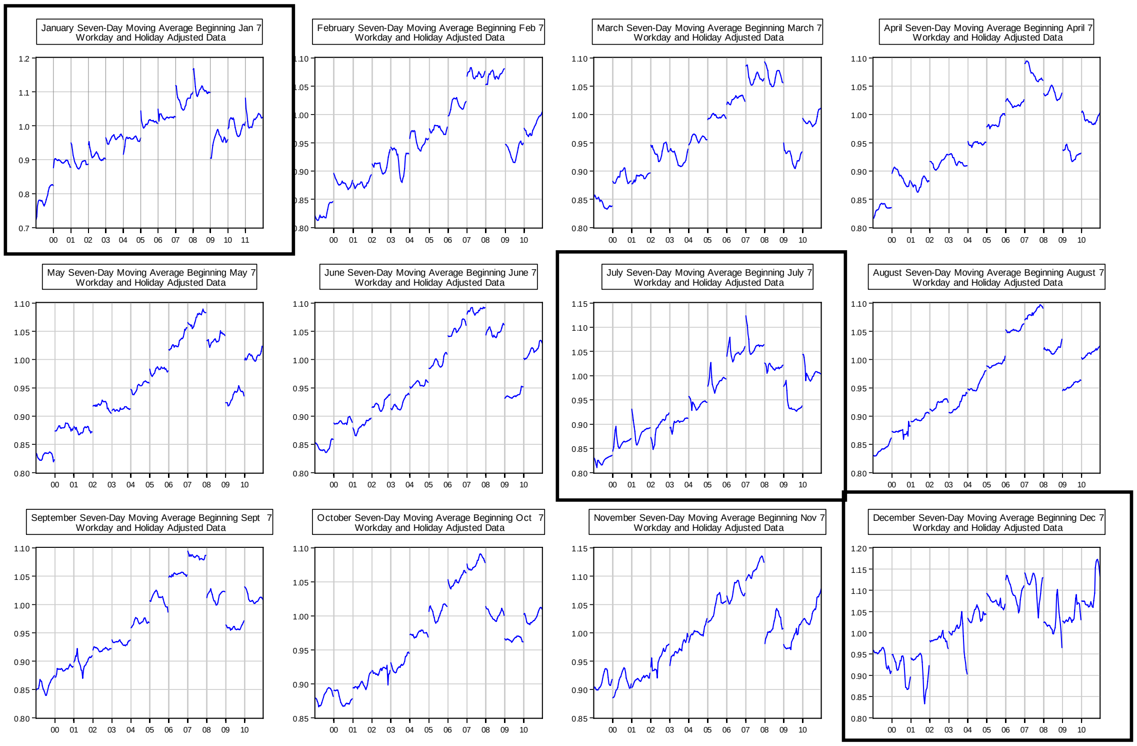


Figure 16 Seven-day Moving Averages of Workday and Holiday Adjusted Data, from the 7th to the last day of each month, January, July and December HIGHLIGHTED



The standard errors of EQ00 daily residuals during the monthly transitions reported in Table 5 further confirm that the problem transition is Dec/Jan, and to a lesser extent June/July. These summary statistics apply to the days at the ends and beginning of months, first seven days in each, and then two days in each. The biggest standard error for the 14 day transition is in Dec/Jan (7.7%) followed by June/July (5.8%). The other months all have substantially smaller standard errors of these peculiarity measures. When attention is restricted to two days from each month, it is the Dec/Jan item that is extreme by a factor of two, as measured by the standard deviation.

Table 5 **Transition Means and Standard Deviations of EQ00 daily residuals**

EQ00 Residuals
During Monthly Transitions

		14 Days per Transition			4 Days per Transition		
FROM	**TO**	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.
Jan	Feb	-0.001	0.023	175	0.004	0.026	50
Feb	Mar	0.006	0.018	168	0.008	0.015	48
Mar	Apr	-0.001	0.017	168	-0.003	0.013	48
Apr	May	-0.001	0.013	168	0.000	0.011	48
May	Jun	0.000	0.020	168	0.003	0.021	48
Jun	Jul	-0.001	**0.058**	168	-0.007	**0.033**	48
Jul	Aug	0.001	0.010	168	0.003	0.009	48
Aug	Sep	0.003	0.021	168	0.001	0.020	48
Sep	Oct	0.000	0.012	168	0.002	0.014	48
Oct	Nov	0.003	0.030	168	-0.003	0.017	48
Nov	Dec	0.010	0.022	168	0.017	0.017	48
Dec	Jan	-0.011	**0.077**	168	-0.012	**0.075**	53
Other		0.000	0.022	2384	0.000	0.028	3824

June/July Transitions

The goal of this paper is to create monthly aggregates from daily data with adjustments for weekday and holiday effects applied to the daily data, and with calendar adjustments applied to monthly aggregates of the adjusted daily data. For that purpose, it is more revealing to look at weekly data around the troubling holidays rather than the noisier daily data. Figure 17 illustrates the weekly totals for the first week in July, ending July 7, and for the last week of June, ending June 30, and for three previous and three subsequent weeks. The June and July aggregates are basically averages of these weekday totals, with a couple of extra days. The 4th of July would create problems if some years there is substantial leakage into June and other years not so much. Figure 17 clearly reveals the July 7 dip in weekly sales but doesn't suggest that there is a big problem with leakage into June.

Though the July 4th effect seems very similar when the data are displayed in Figure 17, Figure 18 tells a different story. Here the 4th of July effects have been removed from the daily data per EQ00, and the year level-effects have been removed by dividing by yearly averages. All that remains are the June/July patterns not removed by the EQ00 July dummy variables. There is clearly a lot of noise remaining in the first week of July, but not much in the last week of June. This is supporting the conclusion that the volatility in the June/July comparison is not due to July 4th effects that sometimes leak into June but rather to the noise strictly within the data during the first week of July.

Figure 17 **Weekly Diesel Fuel Volumes, June - July**

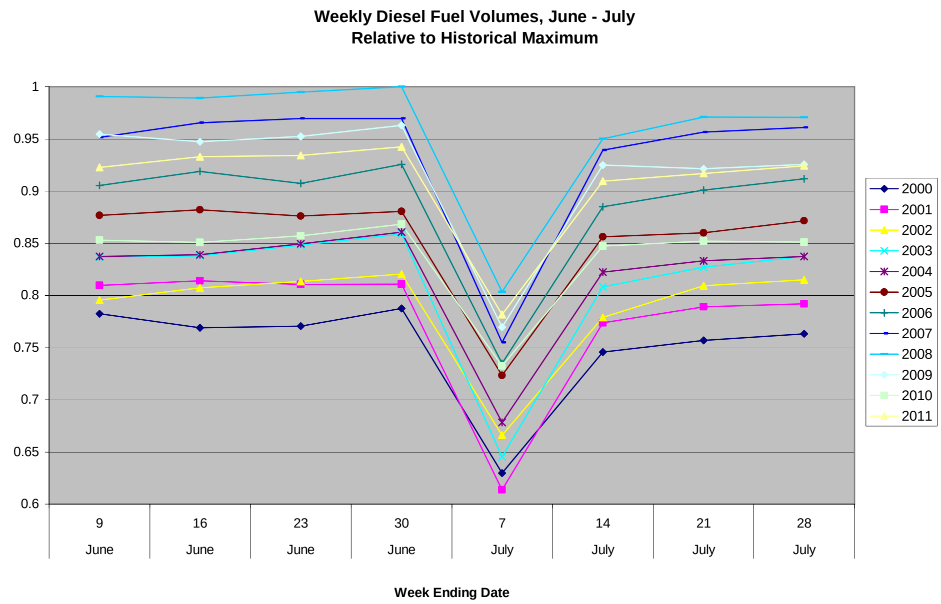
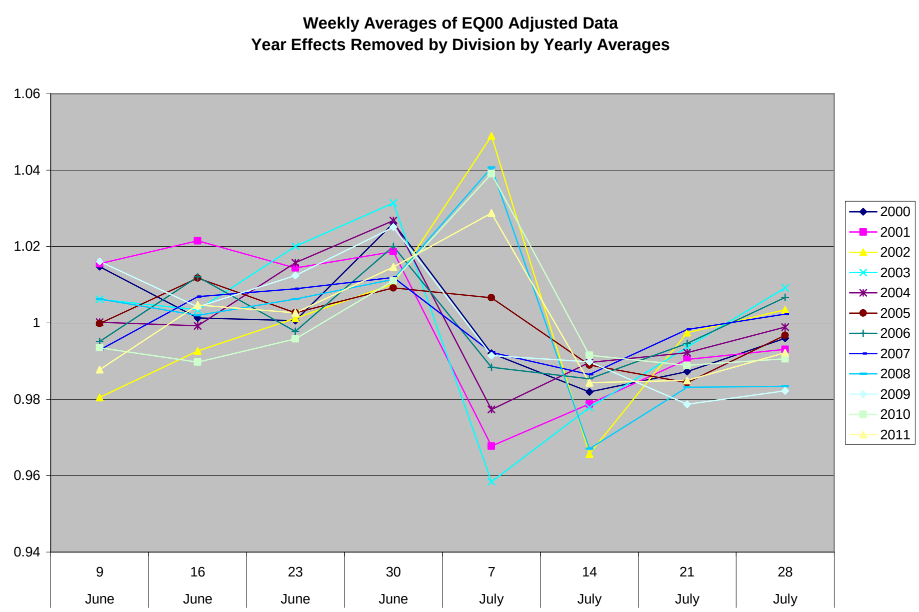
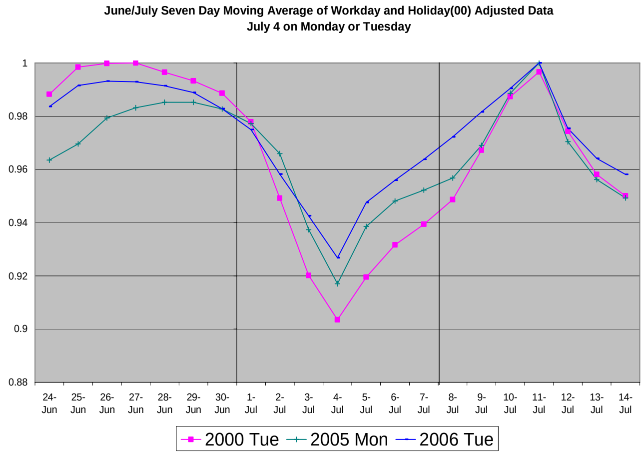


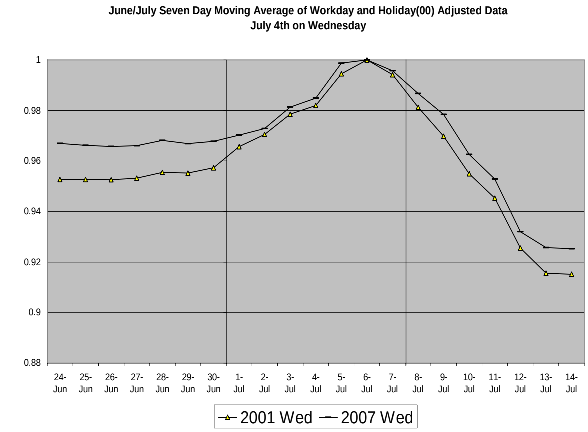
Figure 18 **Weekly Diesel Fuel Averages of Adjusted Daily Data, EQ00**

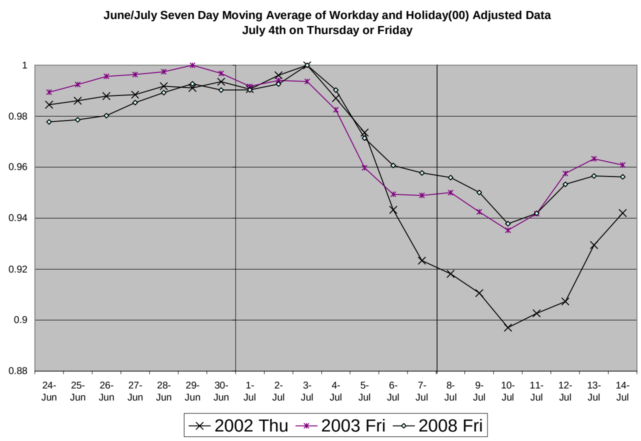


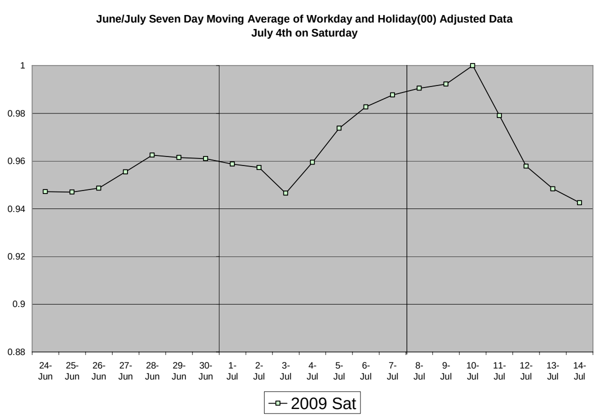
Though these initial displays do not suggest that the noise in the June/July transition are susceptible to statistical correction, the pattern during the first week of July could still be dependent on the weekday on which the 4^{th} of July falls. The problems with the June/July transition are further revealed in Figure 19 which contains the seven-day averages of the adjusted data clustered by weekday. Remember if the weekday and holiday adjustments are working correctly these averages should be generally flat or trending. Clearly they are not. The patterns for (Monday, Tuesday) (Wednesday) (Thursday, Friday), Saturday and Sunday are distinctly different, which means that the 4^{th} of July effect is spread over time depending on the day of the week on which the 4^{th} falls. Some clustering of these weekdays is required because with only ten years of data there are not enough instances to allow each weekday to have a distinct pattern of effects. The figure however supports a very natural clustering: (Mon, Tues), Wednesday, (Thurs, Fri), Sat, Sun. While I will use dummy variables to allow these patterns to differ, for the single year of data with a Saturday July 4^{th} (2009) this amounts to discarding these data, which is a symptom of the overfitting that I am now embarked on. More data, especially more Saturdays are going to help.

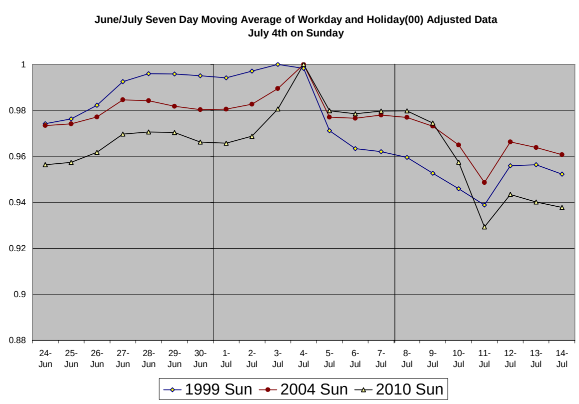
Figure 19 June/July Seven Day Moving Average of Workday and Holiday(00) Adjusted Data











December/January Transitions

The December/January transitions are quite a bit noisier than the June/July transitions, more dependent on the weekday of the holidays, and more problematic for forming monthly aggregates. Figure 21 illustrates the weekly data during December and January, divided by year and week averages. The horizontal axis is the year of the data, each with its own weekday for Jan 1. For each of these years, the data for eight weeks are displayed. These weeks are therefore the different series, and the critical weeks ending December 31 and January 7 are highlighted. Here we see that for the two Thursdays in the data, both of these weeks are abnormally low. The leakage hypothesis would suggest that one would be low and one would be high.

Figure 20 Doubly Normed Weekly Averages, December/January

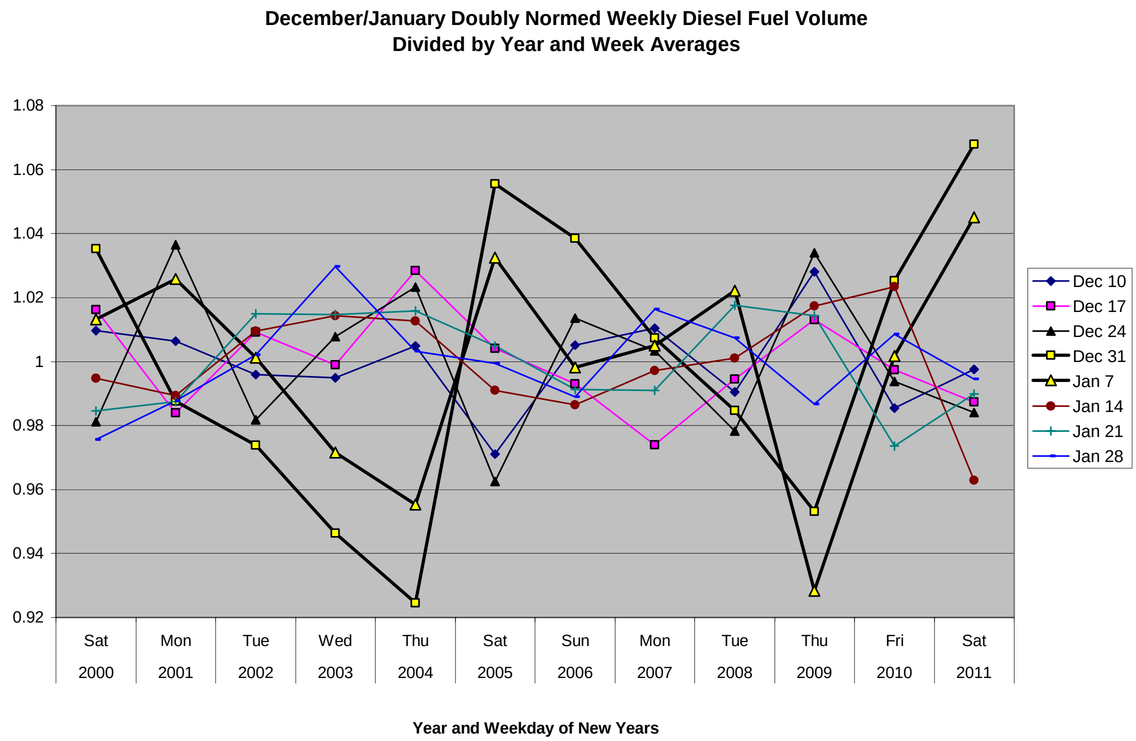
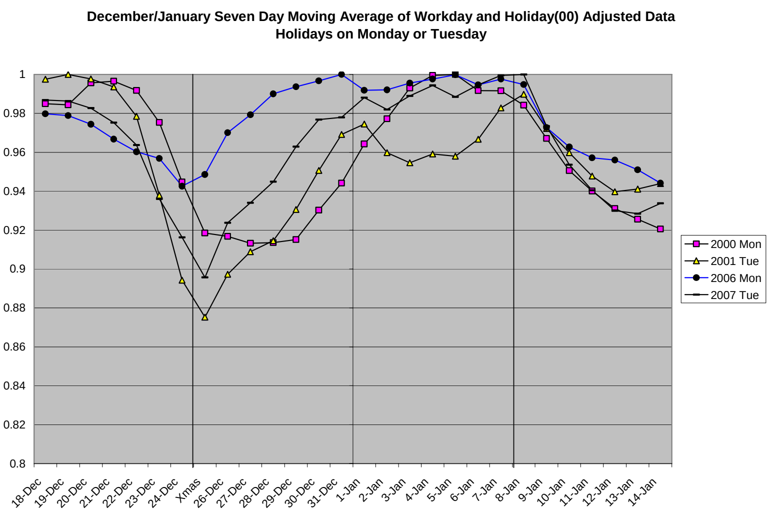
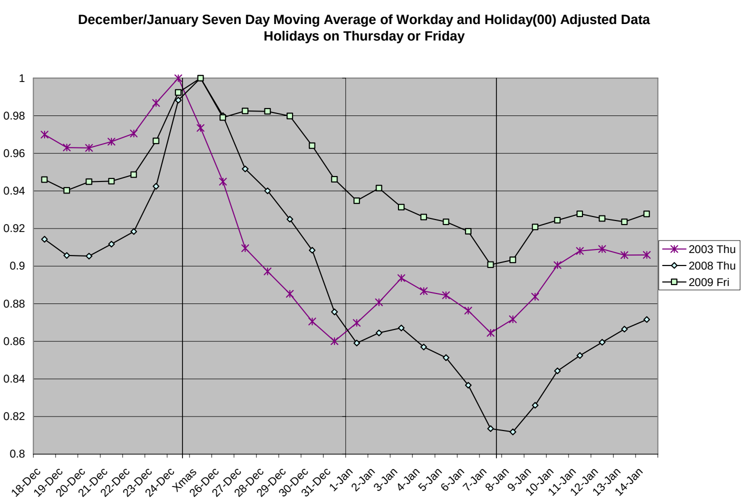
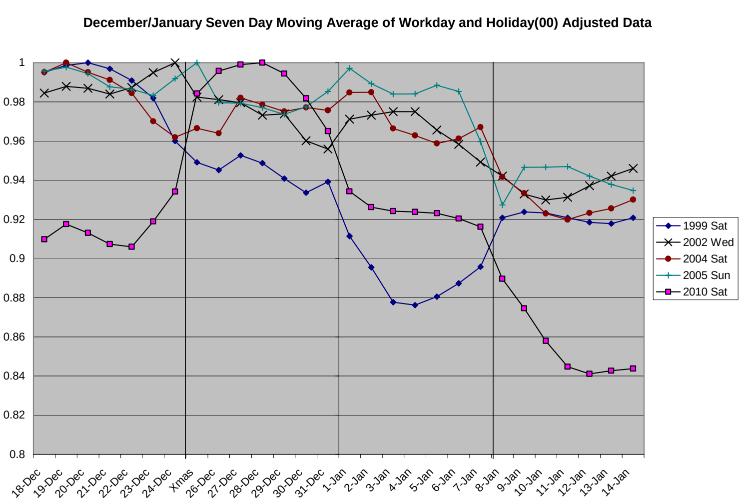


Figure 21 displays the seven-day averages of the EQ00 adjusted data clustered by weekday for the December/ January transition. As is the case with the July 4^{th} holiday, the beginning of the workweek (Mon, Tues) and the end of the workweek (Thurs, Fri) form natural groups. The shape of the Wednesday effect seems similar to Saturday and Sunday, though the 1999 Saturday profile is distinct.

Figure 21 Jan/Dec Seven Day Moving Average of Workday and Holiday(00) Adjusted Data







Fully Adjusted Data: EQ03

Based on the displays and discussion above, a EQ02 is estimated that allows the holiday effects to vary by workday per the groupings in Table 6. EQ03in addition allows the Dec/Jan pattern in 1999/2000 to be distinct, something that is suggested by the different timing of the Sunday and Saturday pattern in Figure 21. The estimated regression for EQ02 is not reported. EQ03 is reported in Table 14 in the appendix.

Table 6 Clustering of July 4 and Holiday Effects by Weekdays

July 4 Clustering

2005	Mon
2000	Tue
2006	Tue
2001	Wed
2007	Wed
2002	Thu
2003	Fri
2008	Fri
2009	Sat
1999	Sun
2004	Sun
2010	Sun

Holiday Clustering

Dec	Jan	
2000	2001	Mon
2006	2007	Mon
2001	2002	Tue
2007	2008	Tue
2002	2003	Wed
2003	2004	Thu
2008	2009	Thu
	1999	Fri
2009	2010	Fri
1999	2000	Sat
2004	2005	Sat
2010	2011	Sat
2005	2006	Sun

Holiday Patterns per EQ03

The patterns of the Christmas/New Years effect that come from EQ03 are displayed in Figure 22 and the corresponding patterns for the July 4^{th} effects are displayed in Figure 23. The adverse effects on trucking "lean left" and the Thurs/Fri effects "lean right". In both cases, the holiday is extended into the adjacent weekend: the previous weekend for the Mon/Tues holidays and the subsequent weekend for the Thurs/Fri holidays.

For the first few days of January, the difference between the Mon/Tues effects and the Thursday/Friday effects are more than 10%, which can wreak havoc with seasonal adjustment methods that do not account for this effect. For June, however the difference among the corrections is mostly confined to June 30 and is not very large. This suggests that the weekday controls in EQ03 are most important for month totals for Dec/Jan and not very important for the June/July transitions.

Figure 22 **EQ03: Christmas and New Years Effects by Weekday of Holiday**

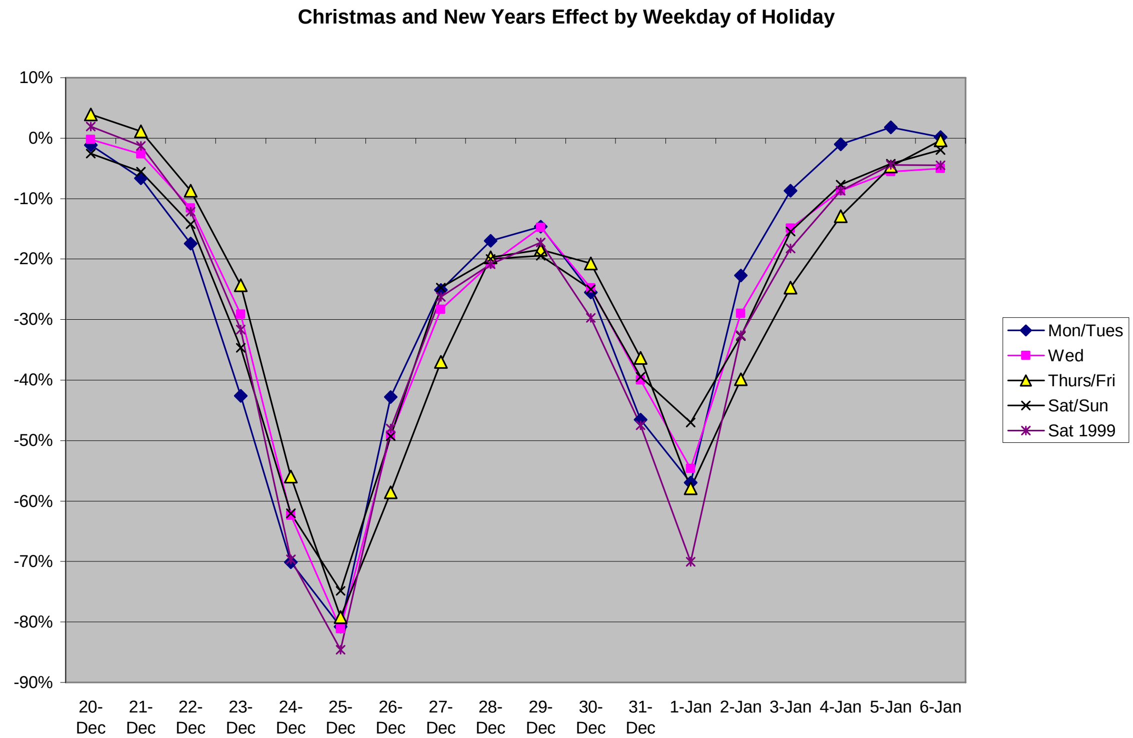
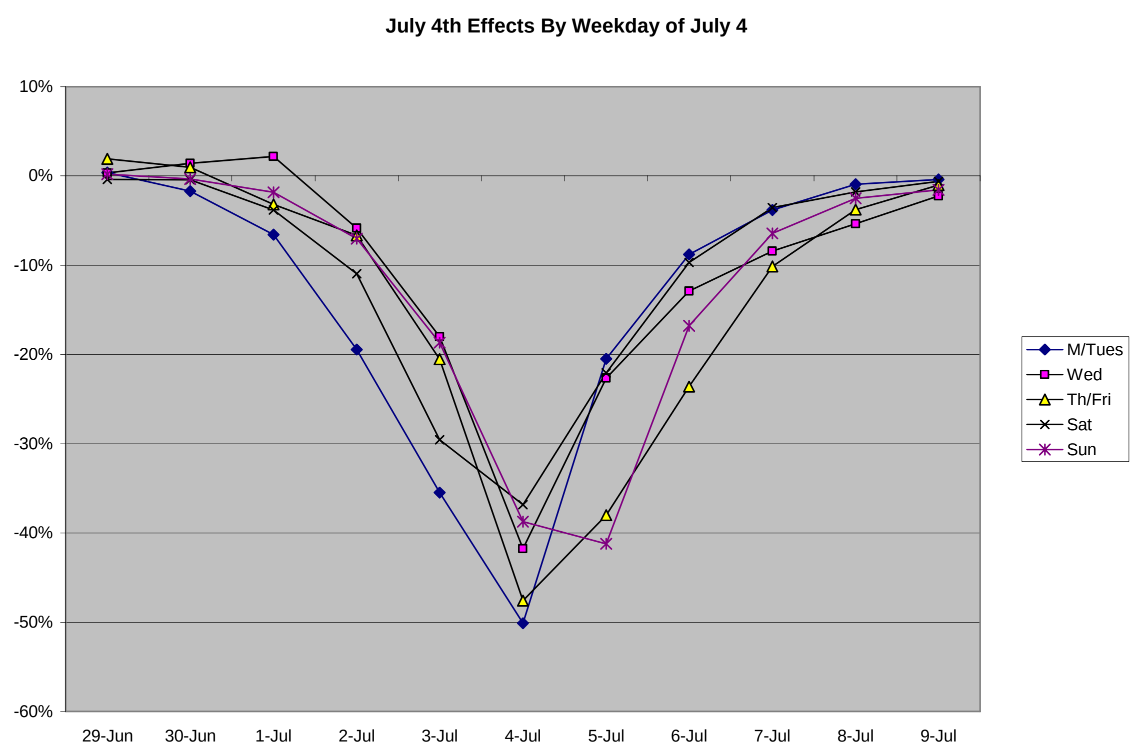


Figure 23 **EQ03: July 4th Effects by Weekday of July 4**



Residuals during Dec/Jan and June/July Transitions: Summary Stats

The effectiveness of the equations EQ02 and EQ03 that allow holiday effects to vary by holiday weekday compared with EQ00 which does not is summarized numerically by standard deviations of residuals during each of the Dec/Jan sequences in Table 7 and likewise for each June/July pair in Table 8   The December/January transitions include the last week of December and the first week of January, and the June/July transitions include the last week of June and the first week of July.

The overall standard deviation of the Dec/Jan transition residuals is reduced 33% when the holiday drift is allowed to be weekday dependent per the groupings in the table which is the basis for EQ02. The greatest improvements comes for the 2003-Wednesday pattern, though this is a consequence of overfitting, since there is only one Wednesday New Year in the sample. The holiday drift per EQ02 actually makes the 2000 data worse, and this of course is much improved if there is a distinct holiday pattern per EQ03.

Table 7 Dec/Jan Residuals from Three Equations

Residual Standard Deviations During Dec/Jan Transitions
Last Week of December and First Week of January
14 Observations per year

NEWYEAR	Weekday	EQ00	EQ02	v. EQ00	EQ03	v. EQ00
2001	Mon	0.047	0.043	-7%	0.043	-7%
2007	Mon	0.056	0.036	-36%	0.036	-36%
2002	Tue	0.072	0.036	-50%	0.037	-49%
2008	Tue	0.088	0.037	-58%	0.038	-56%
2003	Wed	0.049	0.007	-85%	0.007	-85%
2004	Thu	0.087	0.054	-38%	0.054	-38%
2009	Thu	0.076	0.035	-54%	0.037	-51%
2010	Fri	0.075	0.064	-14%	0.064	-15%
2005	Sat	0.067	0.046	-31%	0.048	-28%
2011	Sat	0.090	0.063	-29%	0.058	-35%
2006	Sun	0.104	0.083	-20%	0.074	-29%
2000	Sat	0.065	0.082	26%	**0.002**	**-98%**
All		0.077	0.052	-32%	0.045	-41%

Note: EQ00 has the same holiday effect regardless of weekday of New Years
Eq02 has varying holiday effects per the borders in the table
EQ03 in addition allows the 2000 transition to have it's own pattern.

Standard deviations of residuals during each of the June/July sequences are reported in Table 8 which includes both the EQ02 and EQ03 results even though allowing for a distinct Dec/Jan effect in 1999/2000 has no evident impact on the June/July values. Overall, allowing the July 4th effect to vary by weekday of the holiday reduces the residual standard error by 62%, a much larger reduction than the Dec/Jan treatment.

Table 8 June/July Outliers from Three Equations

Residual Standard Deviations During June/July Transitions
Last Week of June and First Week of July
14 Observations per year

YEAR	4-Jul	EQ00	EQ02	v. EQ00	EQ03	v. EQ00
2005	Mon	0.080	0.025	-69%	0.024	-70%
2000	Tue	0.071	0.026	-64%	0.025	-65%
2006	Tue	0.066	0.020	-69%	0.020	-70%
2001	Wed	0.032	0.010	-69%	0.011	-67%
2007	Wed	0.035	0.011	-69%	0.011	-69%
2002	Thu	0.063	0.042	-33%	0.042	-33%
2003	Fri	0.041	0.025	-40%	0.025	-39%
2008	Fri	0.039	0.023	-42%	0.022	-43%
2009	Sat	0.050	0.012	-75%	0.012	-75%
1999	Sun	0.059	0.024	-59%	0.024	-60%
2004	Sun	0.069	0.009	-87%	0.009	-87%
2010	Sun	0.076	0.019	-75%	0.020	-74%
All		0.058	0.022	-62%	0.022	-62%

Note: EQ00 has the same holiday effect regardless of weekday of Holidays
Eq02 has varying holiday effects per the borders in the table
EQ03 in addition allows the 2000 New Years transition to have it's own pattern.

Outliers from EQ03

Though the workday and holiday adjustments remove many of the anomalous observations in the daily data, many remain. Table 9 and Table 10 report the EQ03 residuals in excess of 5%, 112 in all. For all of our efforts to remove the holiday effects, Christmas and New Years and Thanksgiving continue to have large numbers of anomalous observations per the results in Table 10. Table 9 reports large residuals that are not associated with the holiday leakage problem. Many of the isolated cases are depressed sales on January and February Sundays, pointing to the need to have a month correction for our weekday adjustment. The offsetting sequences of abnormalities generally have a negative followed by a positive, suggesting sales delayed a single day. The terrorsit attack on 9/11 has the opposite pattern – increased sales on 9/11 and decreased sales on 9/12, as if truckers rushed to fill their tanks in response to the news. In addition to the 9/11 effect, three other patterns in Table 9 are explainable by extreme Northeast snowstorms, but with your help, I think we can track down the explanations of many of the others, for example, the number one residual on October 28, 2002, -23.8%, followed by the number three residual the next day, +22.0%.

Table 9 EQ03 Residual Outliers Larger than 5% excluding Monthly Transition Problems

Isolated Abnormalities

obs	Weekday	Resid	Rank	
January 19, 1999	Tue	0.060	80	
January 28, 2001	Sun	-0.076	47	
January 10, 2005	Mon	-0.053	99	
January 14, 2007	Sun	-0.056	87	
January 18, 2009	Sun	-0.060	75	
January 31, 2010	Sun	0.075	49	
February 23, 2001	Fri	0.052	101	
February 1, 2004	Sun	-0.052	103	
February 6, 2005	Sun	-0.054	98	
February 4, 2007	Sun	-0.054	95	
February 13, 2007	Tue	-0.052	100	1
February 3, 2008	Sun	-0.063	71	
February 1, 2009	Sun	-0.054	93	
March 17, 2000	Fri	0.077	46	
March 3, 2002	Sun	0.071	52	
March 17, 2002	Sun	-0.056	85	
April 22, 2000	Sat	-0.051	110	
April 4, 2001	Wed	-0.054	97	
April 14, 2001	Sat	-0.052	104	
May 29, 2000	Mon	0.054	92	
May 26, 2002	Sun	-0.056	88	
May 26, 2008	Mon	-0.059	82	
May 30, 2010	Sun	0.060	77	
July 27, 2002	Sat	0.057	83	
September 30, 2005	Fri	-0.055	89	
September 1, 2008	Mon	-0.056	84	
December 13, 2000	Wed	-0.097	26	
December 12, 2003	Fri	0.055	90	

Offseting Sequences

obs	Weekday	Resid	Rank	
May 29, 1999	Sat	-0.051	109	
May 31, 1999	Mon	0.081	38	
September 5, 1999	Sun	-0.051	111	
September 6, 1999	Mon	0.067	57	
August 22, 2000	Tue	-0.161	7	
August 23, 2000	Wed	0.148	8	
August 28, 2000	Mon	-0.064	69	
August 29, 2000	Tue	0.101	23	
September 11, 2001	Tue	0.078	42	2
September 12, 2001	Wed	-0.188	4	
October 28, 2002	Mon	-0.238	1	
October 29, 2002	Tue	0.220	3	
November 8, 2002	Fri	0.087	34	
November 9, 2002	Sat	-0.224	2	
November 10, 2002	Sun	0.093	28	
January 28, 2009	Wed	-0.068	55	
January 30, 2009	Fri	0.051	107	

Sequences of Abnormalities

obs	Weekday	Resid	Rank	
February 14, 2003	Fri	-0.067	59	3
February 16, 2003	Sun	-0.084	36	
February 17, 2003	Mon	-0.113	17	
December 2, 2006	Sat	0.072	51	
December 3, 2006	Sun	0.052	105	
January 9, 2011	Sun	-0.051	106	4
January 10, 2011	Mon	-0.116	16	

Notes

1: Category 3 Snowstorm, Feb 12-15, 2017

3 :Category 4 Snowstorm, Feb 15-18, 2003

4: Category 3 Snowstorm, Jan 9-13, 2011

2: 9/11 Attack

Table 10 EQ03 Residual Outliers Larger than 5%, Monthly Transition Problems

Christmas and New Years Abnormalities

	Weekday	Resid	Rank
December 25, 2000	Mon	-0.090	31
December 26, 2000	Tue	-0.106	19
December 30, 2000	Sat	0.067	61
December 23, 2001	Sun	-0.146	9
December 24, 2001	Mon	-0.113	18
December 25, 2001	Tue	0.055	91
December 26, 2001	Wed	0.069	54
December 30, 2001	Sun	-0.060	78
December 24, 2003	Wed	-0.100	25
December 25, 2003	Thu	-0.143	10
December 26, 2003	Fri	0.064	68
December 27, 2003	Sat	-0.101	22
December 23, 2004	Thu	-0.082	37
December 24, 2004	Fri	-0.085	35
December 26, 2004	Sun	0.080	40
December 27, 2004	Mon	0.065	65
December 31, 2004	Fri	-0.066	62
January 2, 2005	Sun	0.073	50
December 25, 2005	Sun	-0.105	21
December 26, 2005	Mon	-0.121	14
December 31, 2005	Sat	0.106	20
January 2, 2006	Mon	-0.176	6
December 23, 2006	Sat	0.124	13
December 25, 2006	Mon	0.077	44
December 26, 2006	Tue	-0.062	73
December 29, 2006	Fri	-0.068	56
December 24, 2007	Mon	0.087	33
December 25, 2007	Tue	-0.077	45
December 26, 2007	Wed	0.064	66
December 30, 2007	Sun	-0.067	60
January 7, 2008	Mon	-0.050	112
December 24, 2008	Wed	0.092	30
December 26, 2008	Fri	-0.076	48
December 27, 2008	Sat	-0.087	32
January 3, 2009	Sat	-0.052	102
December 25, 2009	Fri	0.079	41
December 27, 2009	Sun	0.176	5
December 28, 2009	Mon	-0.078	43
January 3, 2010	Sun	0.060	74
December 23, 2010	Thu	0.065	64
December 25, 2010	Sat	0.125	12
December 27, 2010	Mon	-0.121	15
December 31, 2010	Fri	-0.063	70
January 2, 2011	Sun	0.070	53

July 4 Problems

	Weekday	Resid	Rank
July 4, 1999	Sun	-0.067	58
July 2, 2000	Sun	-0.059	81
July 6, 2002	Sat	-0.142	11
July 6, 2003	Sun	0.064	67
July 3, 2005	Sun	-0.063	72
July 3, 2006	Mon	0.054	94
July 6, 2008	Sun	0.056	86

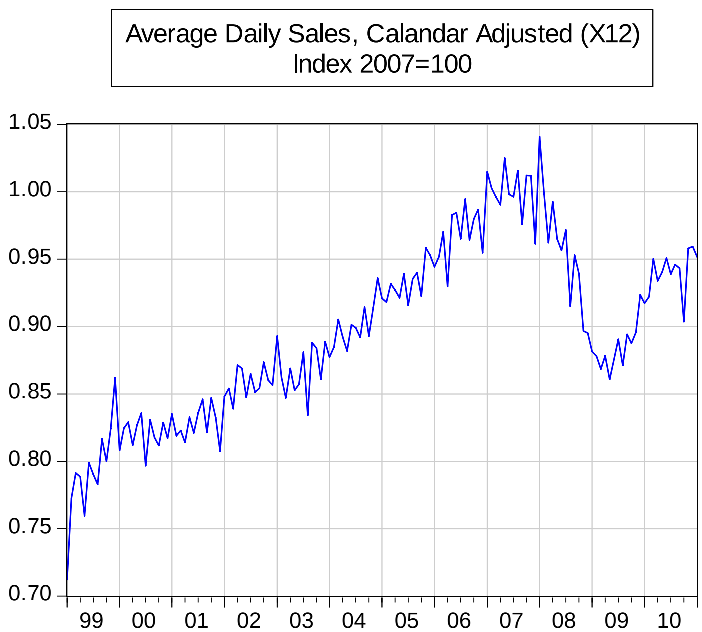
Thanksgiving Problems

November 25, 1999	Thu	-0.060	79
November 23, 2003	Sun	0.081	39
November 28, 2003	Fri	0.060	76
November 26, 2004	Fri	0.066	63
November 16, 2005	Wed	0.054	96
November 24, 2005	Thu	-0.051	108
November 17, 2009	Tue	0.100	24
November 26, 2009	Thu	0.094	27
November 25, 2010	Thu	0.092	29

4. Monthly Data

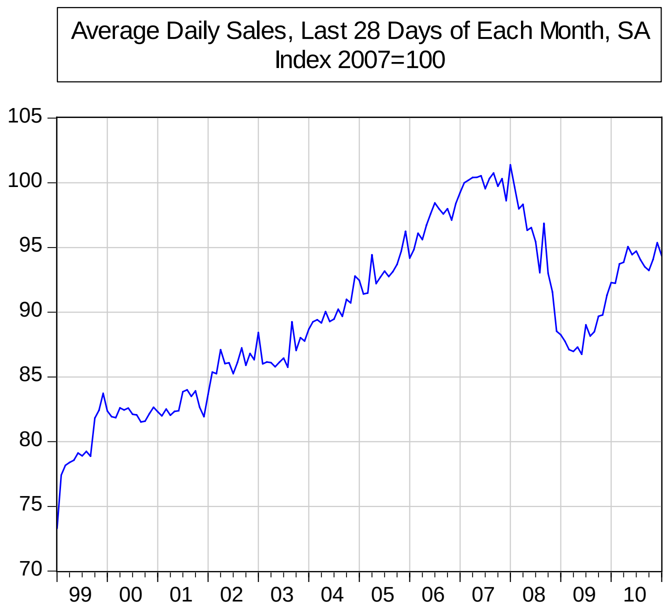
Though the daily data raise interesting issues on their own, here the goal is only to use the daily data to help form monthly aggregates corrected for variable workdays and wandering holidays as well as calendar effects. To set the stage, Figure 24 illustrates calendar adjusted daily average diesel fuel sales for each month, adjusted with X12, without workday or holiday corrections. The sawtooth pattern with up and downs following one after another is very suspicious. It is possible that truckers all over the country near the end of each month decide collectively whether to purchase diesel in the current or subsequent month, or, more likely, we are making an error in allocating sales between adjacent months.

Figure 24 **Straw Man: Calendar Adjusted Index (X12)**
No Workday or Holiday Adjustment



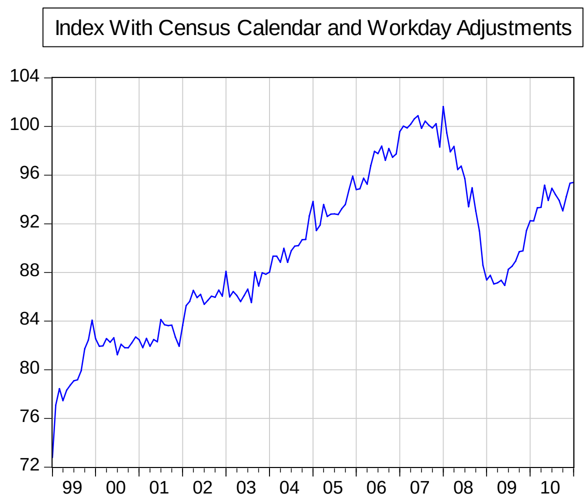
One thing that could give rise to this bi-monthly pattern is the weekday effect since a month that benefits from an extra Wednesday does so at the expense of adjacent months. It is therefore anticipated that the apparent "nervousness" of the index illustrated in Figure 24 will be relieved via the removal of weekday effects. One way to remove the weekday effects is to use aggregates that have a constant composition of weekdays. The last 28 days of each month is an option, and an index based on these last-of-the-month 28-day averages is illustrated in Figure 25, which is much smoother than the traditional index that does not account for workdays.

Figure 25 **28 Day Averages, Seasonally Adjusted Data**



But, Figure 24 is labeled a straw man since Census statisticians have thought a lot about how to remove trading day (weekday) effects from the monthly data. After all, the weekday composition of every month is known, and can be removed from the monthly totals if the number of months is enough to estimate the weekday effects. I wonder if the 145 months of data are enough to do this with accuracy. Surprise: When the Census X11 seasonal adjustment with trading day effects is applied to the monthly sums, the result is Figure 26, which is very similar to the 28-day averages. That seems disappointing since it suggests that the daily data may not be all that useful for improving on Census "trading day" adjustment.

Figure 26 **Census "Trading Day Adjustment" and Calendar Adjustment**



Monthly Aggregates from Daily Data

Finally, we arrive near the end of this long journey. We use the daily data adjusted for workday and holiday per EQ01, EQ02 and EQ03 to form monthly averages, and then adjust these monthly averages for calendar effects using X12. This produces three monthly series displayed in Figure 27, each transformed into an index, 2007=100. Here we can see that index based on EQ02 which allows two of the holiday effects to depend on the holiday weekday is much smoother than the index based on EQ01 which has fixed holiday effects. There is little difference between EQ03 which allows a distinct New Years pattern for 1999 and EQ02 which does not. Below these in Figure 28 is an index based on the monthly sums corrected for workday, holiday and calendar effects per Census methods.[4]

[4] CENSUS holiday adjustments:

> "The basic model used by X-12-ARIMA for Easter and Labor Day effects assumes that the level of activity changes on the wth day before the holiday for a specified w, and remains at the new level until the day before the holiday. For Thanksgiving the model used assumes that the level of activity changes on the day that is a specified number of days before or after Thanksgiving and remains at

It should be visually apparent that both the EQ02 and EQ03 Indexes aresmoother than the Census index.   This visual impression is confirmed by the means and standard errors of the change in the log of the Census-based index and the EQ03 based index reported in Table 11, overall and month by month. Overall, the EQ03 approach reduces the standard error by 8.3%, but the improvement is concentrated in the months influenced by holidays not adequately dealt with by Census methods: Memorial Day, the 4th of July, Labor Day, Thanksgiving and New Years.

---

the new level until December 24. The regression variable constructed for the holiday effect is, for a given month t, the proportion of the affected time period that falls in month t. (Actually, as noted in Table 4.1, these regressors are deseasonalized by subtracting off their long-run monthly means.) Essentially the same Easter effect variable applies also to quarterly flow time series, but Labor Day and Thanksgiving effects are not present in quarterly series. X-12-ARIMA does not provide built-in variables for possible holiday effects in stock series.

Figure 27 EQ00, EQ02 amd EQ03 Based Indexes

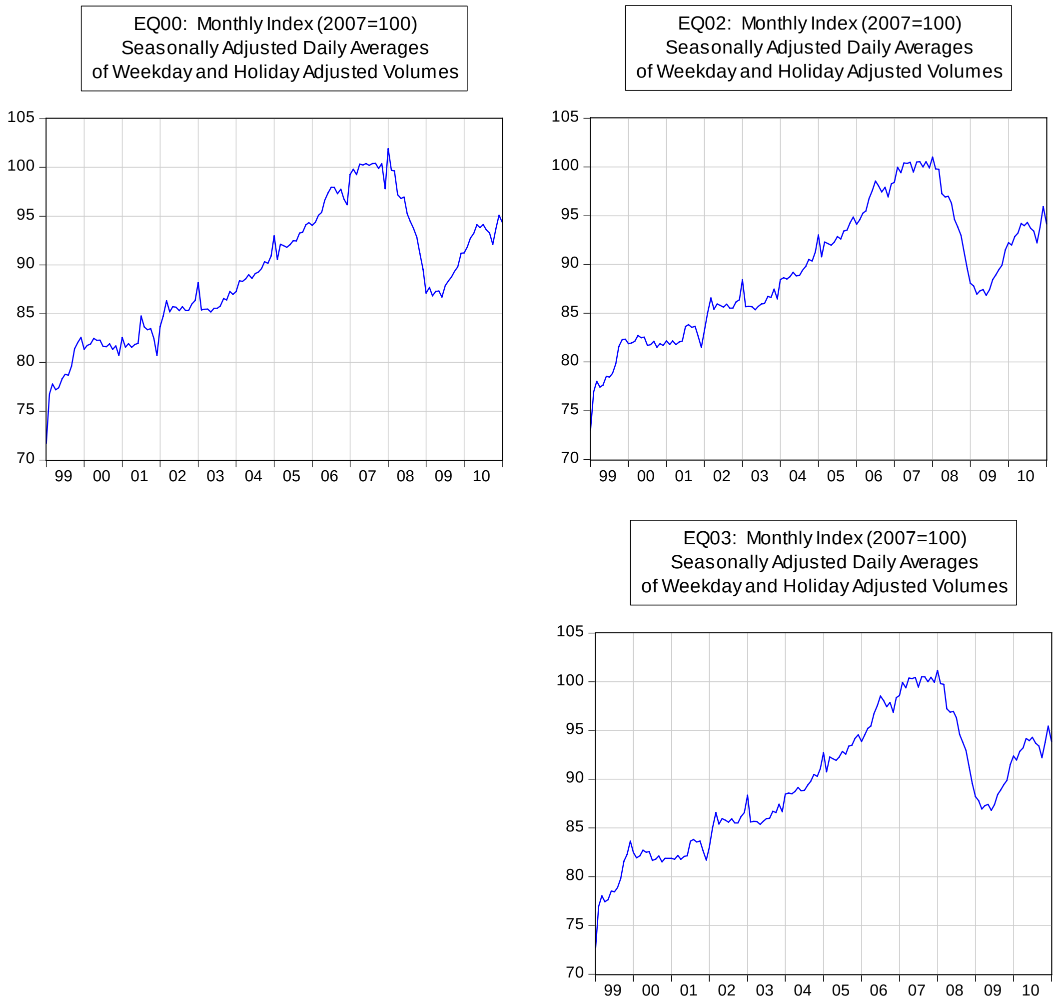


Figure 28 Index Based on Census Methods Applied to Monthly Sum

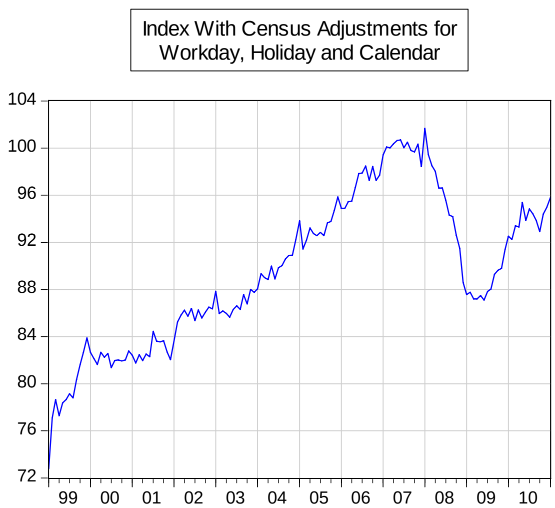


Table 11 Comparison of Census-based Indexe and EQ03-based index

Measures of the Improvement In Monthly Data from Daily Data
Volatility of Indexes Based on Census Methods applied to Monthly data
Versus Methods That First Adjust the Daily Data for Weekday and Holiday Effects

Mean and Standard Error of DLOG of Monthly Data

		Census Methods		EQ03 Methods			
	Obs.	Mean	Std. Dev.	Mean	Std. Dev.	Difference	Reason
Jan	12	0.0072	0.0147	0.0038	0.0143	-2.4%	New Years
Feb	12	0.0011	0.0227	0.0016	0.0228	0.6%	
Mar	12	0.0031	0.0087	0.0048	0.0088	0.3%	
Apr	12	-0.0001	0.0079	-0.0020	0.0098	23.9%	
May	12	0.0033	0.0106	0.0026	0.0055	-48.3%	Memorial Day
Jun	12	-0.0001	0.0082	0.0013	0.0051	-38.1%	Mem Day, 4th of July
Jul	12	0.0020	0.0117	0.0014	0.0084	-28.2%	4th of July
Aug	12	-0.0004	0.0072	0.0008	0.0081	13.2%	
Sep	12	0.0025	0.0104	0.0014	0.0066	-36.3%	Labor Day
Oct	12	0.0004	0.0091	0.0005	0.0092	1.3%	
Nov	12	0.0027	0.0102	0.0017	0.0105	3.4%	
Dec	12	0.0013	0.0150	0.0034	0.0124	-17.2%	Thanksgiving
All	144	0.0019	0.0118	0.0018	0.0108	-8.3%	

5. A Radical Alternative: A Calendar That Eliminates the Leakages

The current Gregorian calendar reflects a compromise between the earth's cycle around the sun, the moon's cycle around the earth and earthly religions that require observances every seven days. Given the daily data we have the freedom to divide the days of the year in a way best suited to describing the trucking data.

One idea already explored with some considerable success is the use of only the 28 days at the end of each month, thereby adjusting for the strong weekly cycle in diesel purchases. This works pretty well but inappropriately omits data from the beginning of all months except the 28-day Februaries,[5] but more importantly for our purposes it doesn't deal with the holiday effects that wander between months from year to year. The extraction of the holiday effects from the daily data described above has been difficult and not necessarily entirely successful, which leads us to explore another idea: minor revisions to the current calendar that prevents the wandering of the holidays by appropriate expansions/contractions of months. Here are the proposed calendar revisions that are meant to deal with wandering New Years, wandering Memorial Days, wandering 4ths of July and wandering Labor Days, leaving untouched the wandering Easters:

Calendar Revisions

	Start	End	Effect
Jan	8-Jan		New Years in December
Feb			
Mar			
Apr			
May		3-Jun	Memorial Day in May
Jun	4-Jun	27-Jun	4th of July in July
Jul	28-Jun		4th of July in July
Aug		28-Aug	Labor Day in September
Sep	29-Aug		Labor Day in September
Oct			
Nov			
Dec		7-Jan	New Years in December

[5] An early 13-month proposal was the 1849 Positivist calendar, created by Auguste Comte. It was based on a 364-day year which included one or two "blank" days. Each of the 13 months had 28 days and exactly four weeks, and each started on a Monday.
http://en.wikipedia.org/wiki/Calendar_reform#Perpetual_calendars

Figure 29 illustrates the effect of this proposed calendar change on the trucking index with workday and Easter corrections per EQ03, but no other holiday corrections. This doesn't seem much smoother than the results already discussed.

Figure 29 **Effect of Calendar Change on Trucking Index**

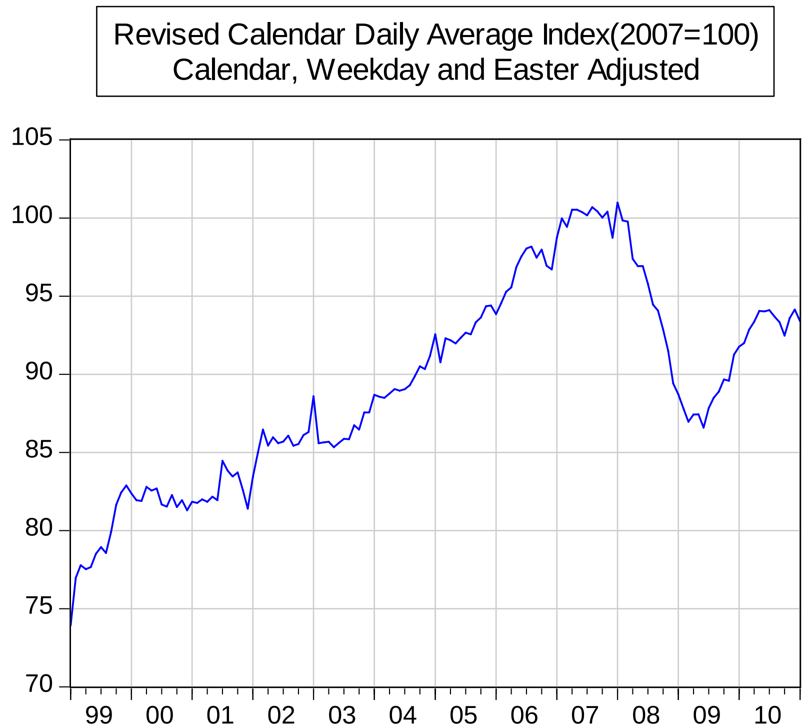


Table 12 reports a numerical comparison of the monthly variability of the index based on the revised calendar with the best adjusted data using the traditional calendar. Overall the revised calendar reduces the standard deviation of the month-to-month variability by 3.4%, but the month by month comparison is a very mixed bag, with no clear winner. Call this horse race a tie. .

Table 12 **Normal Calendar vs. Revised Calendar**

Two Approaches to Holiday Adjustment
Traditional and Revised Calendar
Sample (adjusted): 1999M02 2011M01
12 Observations Per Month

	DLOG(EQ03 Index)		DLOG(Revised Calendar Index)			Change
	Mean	Std. Dev.	Revision	Mean	Std. Dev.	Std. Dev.
Jan	0.004	0.014	Shortened	0.009	0.013	-7.1%
Feb	0.002	0.023		0.000	0.019	-16.1%
Mar	0.005	0.009		0.004	0.008	-3.5%
Apr	-0.002	0.010		0.000	0.010	0.3%
May	0.003	0.006	Lengthened	0.002	0.005	-2.1%
Jun	0.001	0.005	Shortened	0.001	0.005	5.7%
Jul	0.001	0.008	Lengthened	0.003	0.011	34.0%
Aug	0.001	0.008	Shortened	-0.001	0.006	-25.5%
Sep	0.001	0.007	Lengthened	0.002	0.008	22.7%
Oct	0.001	0.009		0.001	0.009	2.7%
Nov	0.002	0.011		0.001	0.010	-7.0%
Dec	0.003	0.012	Lengthened	-0.002	0.012	-5.0%
All	0.002	0.011		0.002	0.010	-3.4%

Timely and Accurate

Among the most important reports that make use of survey-based data sets are the quarterly National Income and Product Accounts provided by the US Bureau of Economic Analysis and the monthly Employment Situation Summary provided by the US Bureau of Labor Statistics. For those who are interested in an early and accurate read on the health of the economy, these data have two disconcerting features – they are delayed by at least a month and they are subject to very substantial revisions. The middle of 2008 was a particularly telling example because the original data provided by the BEA and the BLS was not suggestive of a recession but the revised data revealed a very troubled economy. Figure 30 illustrates data on US Real GDP commencing the first quarter of 2006 and ending the second quarter of 2008.[6] The question mark in this figure is asking for your forecast for the second half of 2008. The initial release by the BEA on August 28, 2008 had GDP growing substantially in the second quarter, clearly indicating that the US was not in recession. Based on these data I imagine you might have been forecasting continued growth through the rest of the year. The first sign of trouble came with the release almost a year later on July 31, 2009, when the revised data indicated a clear GDP peak in the last quarter of 2007. Based on these data I imagine you would have been very worried about the second half of 2008.

Much the same problem afflicted the payroll jobs estimates in the first half of 2008 illustrated in Figure 31. According to the data available on July 3, 2008, payrolls were declining at a rate well under 100 thousand per month. Keeping in mind that recession-level declines are 200 thousand per month or more, this together with the early GDP estimates revealed a troubled economy that had not yet fallen into a full-blown recession. At the time it seemed possible that the second half of 2008 would be more of the same – a troubled job market and weak GDP growth, but not recession-level declines. The subsequent revisions increased the job losses to 150 thousand per month, and more, and make the first half of 2008 look a lot weaker, but we didn't know this until Feb 6, 2009.

[6] ALFRED data from the St.Louis Fed: http://alfred.stlouisfed.org/category?cid=18

Figure 30 **Three Estimates of Real GDP Through 2008Q2**

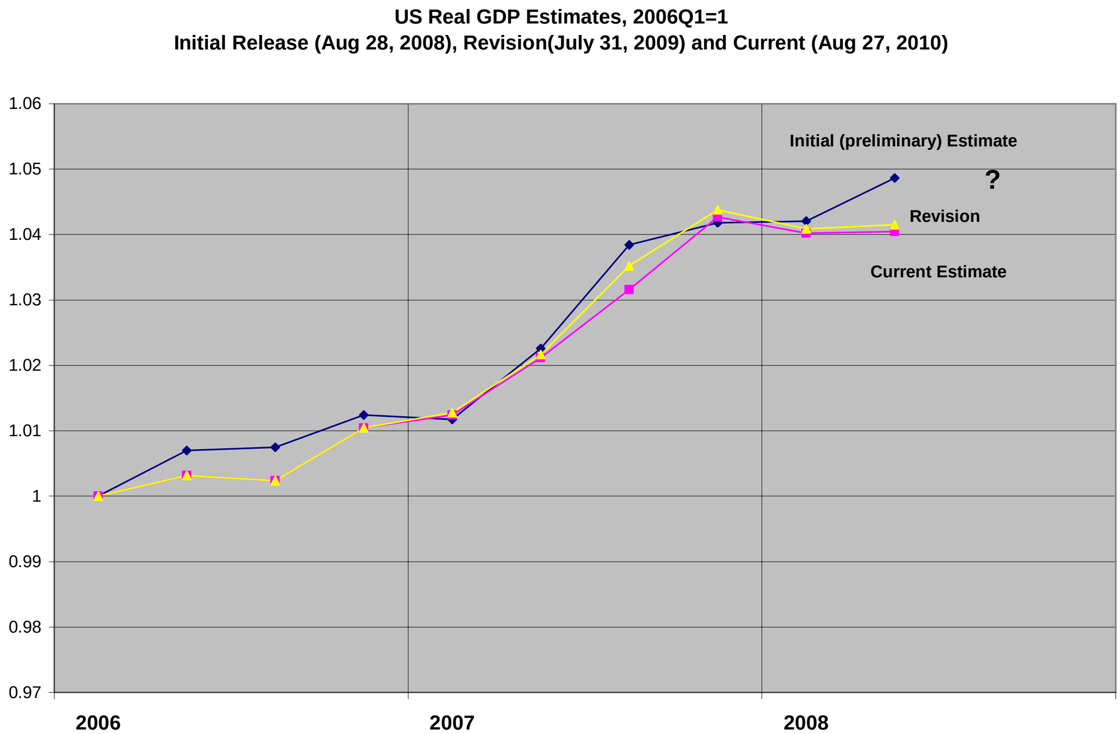
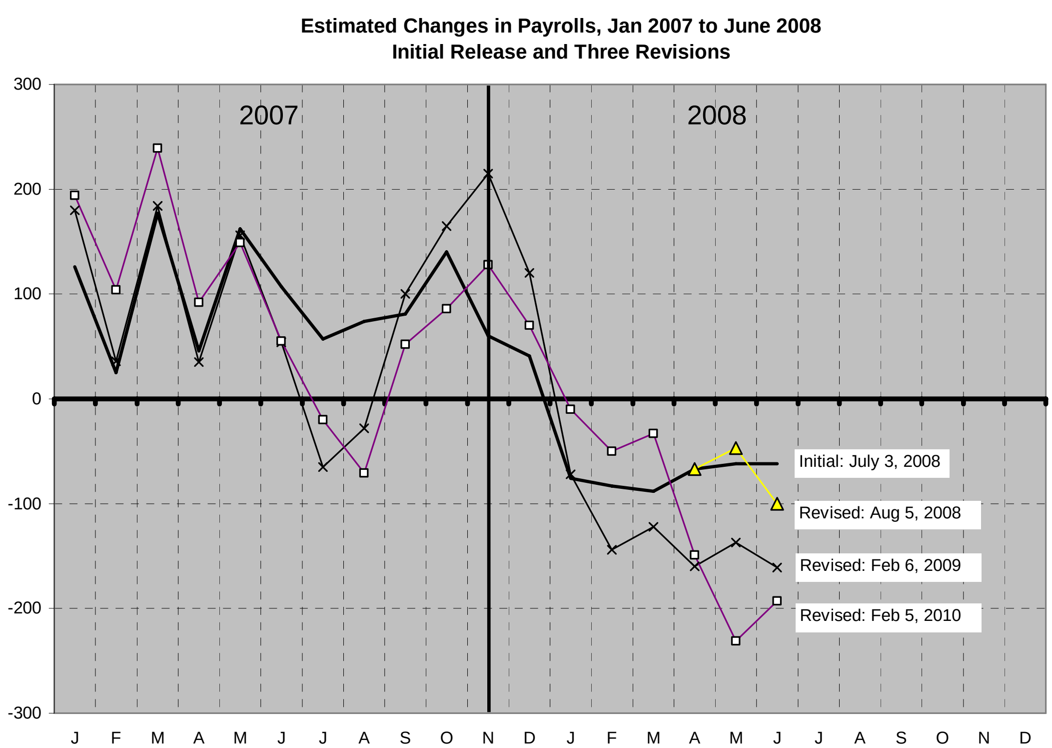


Figure 31 **Four Estimates of Payroll Employment through June 2008**



Conclusion

The most salient feature of the daily diesel data is a very large weekly cycle with weekend purchases only half of midweek purchases. This matters much for forming a monthly index because the weekday composition of any given month varies from year to year. Census seasonal adjustment methods X11 and X12 include an option for weekday adjustment inferred from monthly data, but with daily data we have direct evidence of the workday effect. One surprise is how well Census methods using monthly data do in comparison with direct purging of the weekday effect using the daily data. The next step after removing the weekday effect from the daily data is holiday adjustment using traditional dummy-variable regression estimates of the holiday impacts. Though Census X11 and X12 allow holiday adjustments inferred from monthly data, Census has missed the way that some of the holiday effects drift back and forth, which is evident in the daily diesel data.

The bottom line here is that these daily data afford an understanding of the health of the economy that is a substantial improvement over monthly data. Moreover, these are actual transactions – real data, in real time.

References

Bell, William R. Bell and Donald E. K. Martin, Modelling Time-Varying Trading-Day Effects in Monthly Time Series. http://www.census.gov/srd/www/sapaper/jsm2004dem_abs.html

Bell, W. R. and S. C. Hillmer (1983). Modeling time series with calendar variation. *Journal of the American Statistical Association* 78, 526-534.

Bell, W. R. and S. C. Hillmer (1984). Issues involved with the seasonal adjustment of economic time series. Journal of Business and Economic Statistics 2, 299-320. (invited paper with discussion).

Bell, W. R. and S. C. Hillmer (1985). A reply (to comments by Sims in Bell and Hillmer(1984)). *Journal of Business and Economic Statistics* 3, 95-97.

Cleveland, W. P. and M. R. Grupe (1983). Modeling time series when calendar effects are present. In Proceedings of the Conference on Applied Time Series Analysis of Economic Data, pp. 57-73. Washington, DC: U. S. Census Bureau.

Cleveland, W. S. and S. J. Devlin (1980). Calendar effects in monthly time series: Detection by spectrum analysis and graphical methods. *Journal of the American Statistical Association* 75, 487-496.

Findley, D. F. and R. J. Soukup (2000). Modeling and model selection for moving holidays. Proceedings of the American Statistical Association, Business and Economic Statistics Section, 102{107. http://www.census.gov/ts/papers/asa00_eas.pdf.

US Census Bureau, X-12 Arima Reference Manual, June 2002

U.S. Census Bureau, Statistical Research Division (2009), “X-12-ARIMA Reference Manual, Version 0.3, December 17, 2009 http://www.census.gov/srd/www/x12a/

Appendix: Estimated Regressions

Table 13 **EQ00: Regression Estimate for Holiday Adjustment**

Dependent Variable: LOG(TRACTORGALLONS/WEEKDAY_ADJ)
Sample (adjusted): 1/08/1999 1/31/2011

Variable	Coeff	t-Stat	Effect
XMAS=-5	0.003	0.3	0.3%
XMAS=-4	-0.036	**-3.9**	-3.5%
XMAS=-3	-0.145	**-15.5**	-13.5%
XMAS=-2	-0.426	**-45.7**	-34.7%
XMAS=-1	-1.026	**-110.5**	-64.2%
XMAS=0	-1.557	**-168.3**	-78.9%
XMAS=1	-0.687	**-74.4**	-49.7%
XMAS=2	-0.331	**-35.3**	-28.2%
XMAS=3	-0.209	**-22.2**	-18.9%
JAN1=-3	-0.193	**-20.4**	-17.5%
JAN1=-2	-0.281	**-30.0**	-24.5%
JAN1=-1	-0.542	**-58.9**	-41.8%
JAN1=0	-0.810	**-88.2**	-55.5%
JAN1=1	-0.382	**-41.6**	-31.8%
JAN1=2	-0.178	**-19.4**	-16.3%
JAN1=3	-0.077	**-8.4**	-7.4%
JAN1=4	-0.029	**-3.3**	-2.9%
JAN1=5	-0.014	-1.7	-1.4%
EASTER=-5	0.006	0.8	0.6%
EASTER=-4	0.000	0.0	0.0%
EASTER=-3	-0.044	**-4.8**	-4.3%
EASTER=-2	-0.148	**-16.2**	-13.7%
EASTER=-1	-0.192	**-21.5**	-17.5%
EASTER=0	-0.203	**-22.7**	-18.4%
EASTER=1	-0.065	**-7.2**	-6.3%
EASTER=2	-0.012	-1.3	-1.2%
EASTER=3	0.001	0.1	0.1%
EASTER=4	0.001	0.1	0.1%
EASTER=5	0.001	0.1	0.1%
MEM_DAY=-5	0.006	0.8	0.6%
MEM_DAY=-4	-0.002	-0.2	-0.2%
MEM_DAY=-3	-0.042	**-4.6**	-4.1%
MEM_DAY=-2	-0.131	**-14.3**	-12.2%
MEM_DAY=-1	-0.408	**-45.5**	-33.5%
MEM_DAY=0	-0.551	**-61.5**	-42.4%
MEM_DAY=1	-0.145	**-16.1**	-13.5%
MEM_DAY=2	-0.029	**-3.2**	-2.8%
MEM_DAY=3	0.010	1.1	1.0%
MEM_DAY=4	0.013	1.4	1.3%
MEM_DAY=5	0.021	**2.6**	2.1%
JULY4=-5	0.008	1.0	0.8%
JULY4=-4	0.002	0.3	0.2%
JULY4=-3	-0.026	**-2.9**	-2.6%
JULY4=-2	-0.107	**-11.7**	-10.2%
JULY4=-1	-0.280	**-31.1**	-24.4%
JULY4=0	-0.586	**-65.1**	-44.3%
JULY4=1	-0.375	**-41.7**	-31.2%
JULY4=2	-0.169	**-18.6**	-15.6%
JULY4=3	-0.072	**-7.9**	-7.0%
JULY4=4	-0.031	**-3.5**	-3.1%
JULY4=5	-0.014	-1.7	-1.3%

Variable	Coeff	t-Stat	Effect
LB_DY=-5	0.003	0.4	0%
LB_DY=-4	-0.005	-0.6	-1%
LB_DY=-3	-0.036	**-3.9**	-4%
LB_DY=-2	-0.127	**-13.9**	-12%
LB_DY=-1	-0.392	**-43.7**	-32%
LB_DY=0	-0.553	**-61.6**	-42%
LB_DY=1	-0.159	**-17.7**	-15%
LB_DY=2	-0.032	**-3.5**	-3%
LB_DY=3	0.014	1.5	1%
LB_DY=4	0.021	**2.4**	2%
LB_DY=5	0.023	**2.9**	2%
THNKSGIV=-5	0.034	**4.2**	3%
THNKSGIV=-4	0.058	**6.5**	6%
THNKSGIV=-3	0.002	0.2	0%
THNKSGIV=-2	-0.054	**-6.0**	-5%
THNKSGIV=-1	-0.284	**-31.6**	-25%
THNKSGIV=0	-0.991	**-110.5**	-63%
THNKSGIV=1	-0.805	**-89.8**	-55%
THNKSGIV=2	-0.501	**-55.1**	-39%
THNKSGIV=3	-0.155	**-16.9**	-14%
THNKSGIV=4	-0.048	**-5.3**	-5%
THNKSGIV=5	0.000	0.0	0%

C	16.761	648.2
AR(1)	-0.172	-1.3
AR(2)	-0.001	0.0
AR(3)	0.278	4.4
AR(4)	0.054	0.9
AR(5)	0.381	6.3
AR(6)	0.137	2.0
AR(7)	0.308	5.7
MA(1)	0.717	5.3
MA(2)	0.380	4.9
MA(3)	-0.105	-1.7
MA(4)	-0.111	-2.0
MA(5)	-0.443	-7.9
MA(6)	-0.317	-4.5
MA(7)	-0.234	-4.0

R-squared	0.97	Mean dependent var	16.70
Adjusted R-squared	0.97	S.D. dependent var	0.17
S.E. of regression	0.03	Akaike info criterion	-4.28
Sum squared resid	3.42	Schwarz criterion	-4.16
Log likelihood	9529.87	Hannan-Quinn criter.	-4.24
F-statistic	1779.07	Durbin-Watson stat	2.00
Prob(F-statistic)	0.00		

Table 14 EQ03: Regression Estimate for Holiday Adjustment

Dependent Variable: LOG(TRACTORGALLONS/WEEKDAY_ADJ)

Z1=(@YEAR=2000 OR @YEAR=2005 OR @YEAR=2006)
Z2=(@YEAR=2002 OR @YEAR=2003 OR @YEAR=2008)
Z3=(@YEAR=1999 OR @YEAR=2004 OR @YEAR=2010)
Z4=(@YEAR=2000 OR @YEAR=2006 OR @YEAR=2001 OR @YEAR=2007)
Z5=(@YEAR=2001 OR @YEAR=2007 OR @YEAR=2002 OR @YEAR=2008)
Z6=(@YEAR=2003 OR @YEAR=2008 OR @YEAR=1998 OR @YEAR=2009)
Z7=(@YEAR=2004 OR @YEAR=2009 OR @YEAR=1999 OR @YEAR=2010)
Z8=(@YEAR=1999 OR @YEAR=2004 OR @YEAR=2005 OR @YEAR=2010)
Z9=(@YEAR=2000 OR @YEAR=2005 OR @YEAR=2006 OR @YEAR=2011)

Variable	Coeff	t-Stat	Effect
EASTER=-5	0.007	1.090	1%
EASTER=-4	-0.001	-0.202	0%
EASTER=-3	-0.043	**-6.313**	-4%
EASTER=-2	-0.147	**-21.501**	-14%
EASTER=-1	-0.193	**-28.650**	-18%
EASTER=0	-0.203	**-30.137**	-18%
EASTER=1	-0.064	**-9.412**	-6%
EASTER=2	-0.011	-1.609	-1%
EASTER=3	0.001	0.083	0%
EASTER=4	0.003	0.419	0%
EASTER=5	0.002	0.285	0%
MEM_DAY=-5	0.004	0.650	0%
MEM_DAY=-4	-0.003	-0.460	0%
MEM_DAY=-3	-0.042	**-6.157**	-4%
MEM_DAY=-2	-0.133	**-19.384**	-12%
MEM_DAY=-1	-0.407	**-60.237**	-33%
MEM_DAY=0	-0.551	**-81.649**	-42%
MEM_DAY=1	-0.145	**-21.535**	-14%
MEM_DAY=2	-0.031	**-4.541**	-3%
MEM_DAY=3	0.010	1.399	1%
MEM_DAY=4	0.013	1.919	1%
MEM_DAY=5	0.020	**3.362**	2%
LABOR_DAY=-5	0.003	0.461	0%
LABOR_DAY=-4	-0.007	-1.067	-1%
LABOR_DAY=-3	-0.039	**-5.641**	-4%
LABOR_DAY=-2	-0.131	**-19.074**	-12%
LABOR_DAY=-1	-0.392	**-58.040**	-32%
LABOR_DAY=0	-0.554	**-82.150**	-43%
LABOR_DAY=1	-0.160	**-23.684**	-15%
LABOR_DAY=2	-0.036	**-5.235**	-4%
LABOR_DAY=3	0.013	1.921	1%
LABOR_DAY=4	0.019	**2.821**	2%
LABOR_DAY=5	0.024	**3.944**	2%
THANKSGIVING=-5	0.033	**5.498**	3%
THANKSGIVING=-4	0.053	**8.034**	5%
THANKSGIVING=-3	-0.004	-0.541	0%
THANKSGIVING=-2	-0.059	**-8.671**	-6%
THANKSGIVING=-1	-0.286	**-42.224**	-25%
THANKSGIVING=0	-0.996	**-147.407**	-63%
THANKSGIVING=1	-0.810	**-119.832**	-56%
THANKSGIVING=2	-0.505	**-73.789**	-40%
THANKSGIVING=3	-0.160	**-23.381**	-15%
THANKSGIVING=4	-0.052	**-7.852**	-5%
THANKSGIVING=5	-0.003	-0.528	0%
C	16.759	**656.699**	

AR(1)	-0.962	**-11.1**
AR(2)	0.363	**4.0**
AR(3)	0.437	**4.2**
AR(4)	-0.091	-1.0
AR(5)	0.262	**2.9**
AR(6)	0.641	**11.4**
AR(7)	0.325	**5.5**
MA(1)	1.448	**17.2**
MA(2)	0.451	**3.8**
MA(3)	-0.071	-0.8
MA(4)	0.093	1.1
MA(5)	-0.180	**-2.3**
MA(6)	-0.588	**-11.2**
MA(7)	-0.366	**-10.2**

R-squared	0.984507	Mean dependent var	16.7
Adjusted R-squared	0.983759	S.D. dependent var	0.169
S.E. of regression	0.021534	Akaike info criterion	-4.793
Sum squared resid	1.949023	Schwarz criterion	-4.497
Log likelihood	10765.74	Hannan-Quinn criter.	-4.689
F-statistic	1315.661	Durbin-Watson stat	2.0
Prob(F-statistic)	0		

Variable	Coeff	t-Stat	Effect	
JULY4=-5	0.003	0.211	0%	Wed
JULY4=-4	0.014	0.866	1%	
JULY4=-3	0.021	1.277	2%	
JULY4=-2	-0.060	**-3.609**	-6%	
JULY4=-1	-0.199	**-12.003**	-18%	
JULY4=0	-0.540	**-32.699**	-42%	
JULY4=1	-0.257	**-15.538**	-23%	
JULY4=2	-0.138	**-8.250**	-13%	
JULY4=3	-0.088	**-5.244**	-8%	
JULY4=4	-0.055	**-3.381**	-5%	
JULY4=5	-0.023	-1.521	-2%	
(JULY4=-5)*Z1	0.001	0.033	0%	Mon Tu
(JULY4=-4)*Z1	-0.031	-1.486	-2%	
(JULY4=-3)*Z1	-0.089	**-4.127**	-7%	
(JULY4=-2)*Z1	-0.156	**-7.201**	-19%	
(JULY4=-1)*Z1	-0.240	**-11.231**	-35%	
(JULY4=0)*Z1	-0.155	**-7.256**	-50%	
(JULY4=1)*Z1	0.028	1.308	-20%	
(JULY4=2)*Z1	0.046	**2.136**	-9%	
(JULY4=3)*Z1	0.049	**2.265**	-4%	
(JULY4=4)*Z1	0.046	**2.177**	-1%	
(JULY4=5)*Z1	0.019	0.967	0%	
(JULY4=-5)*Z2	0.016	0.820	2%	Thurs F
(JULY4=-4)*Z2	-0.004	-0.209	1%	
(JULY4=-3)*Z2	-0.054	**-2.486**	-3%	
(JULY4=-2)*Z2	-0.009	-0.400	-7%	
(JULY4=-1)*Z2	-0.031	-1.451	-21%	
(JULY4=0)*Z2	-0.105	**-4.947**	-48%	
(JULY4=1)*Z2	-0.221	**-10.353**	-38%	
(JULY4=2)*Z2	-0.131	**-6.079**	-24%	
(JULY4=3)*Z2	-0.019	-0.878	-10%	
(JULY4=4)*Z2	0.017	0.791	-4%	
(JULY4=5)*Z2	0.012	0.635	-1%	
(JULY4=-5)*(@YEAR=2009)	-0.007	-0.264	0%	Sat
(JULY4=-4)*(@YEAR=2009)	-0.018	-0.656	0%	
(JULY4=-3)*(@YEAR=2009)	-0.060	**-2.079**	-4%	
(JULY4=-2)*(@YEAR=2009)	-0.056	-1.921	-11%	
(JULY4=-1)*(@YEAR=2009)	-0.152	**-5.309**	-30%	
(JULY4=0)*(@YEAR=2009)	0.081	**2.823**	-37%	
(JULY4=1)*(@YEAR=2009)	0.008	0.267	-22%	
(JULY4=2)*(@YEAR=2009)	0.036	1.242	-10%	
(JULY4=3)*(@YEAR=2009)	0.052	1.789	-4%	
(JULY4=4)*(@YEAR=2009)	0.037	1.305	-2%	
(JULY4=5)*(@YEAR=2009)	0.017	0.648	-1%	
(JULY4=-5)*Z3	-0.001	-0.062	0%	sun
(JULY4=-4)*Z3	-0.017	-0.833	0%	
(JULY4=-3)*Z3	-0.040	-1.840	-2%	
(JULY4=-2)*Z3	-0.012	-0.569	-7%	
(JULY4=-1)*Z3	-0.008	-0.381	-19%	
(JULY4=0)*Z3	0.051	**2.374**	-39%	
(JULY4=1)*Z3	-0.275	**-12.872**	-41%	
(JULY4=2)*Z3	-0.046	**-2.124**	-17%	
(JULY4=3)*Z3	0.021	0.986	-6%	
(JULY4=4)*Z3	0.030	1.406	-3%	
(JULY4=5)*Z3	0.007	0.351	-2%	

Variable	Coeff	t-Stat	Effect	
XMAS=-5	-0.002	-0.098	0%	Wed
XMAS=-4	-0.027	-1.150	-3%	
XMAS=-3	-0.123	**-5.106**	-12%	
XMAS=-2	-0.345	**-14.202**	-29%	
XMAS=-1	-0.977	**-40.375**	-62%	
XMAS=0	-1.667	**-68.944**	-81%	
XMAS=1	-0.675	**-27.968**	-49%	
XMAS=2	-0.333	**-13.580**	-28%	
XMAS=3	-0.233	**-9.420**	-21%	
JAN1=-3	-0.160	**-6.476**	-15%	
JAN1=-2	-0.285	**-11.612**	-25%	
JAN1=-1	-0.511	**-21.204**	-40%	
JAN1=0	-0.790	**-32.695**	-55%	
JAN1=1	-0.342	**-14.162**	-29%	
JAN1=2	-0.162	**-6.663**	-15%	
JAN1=3	-0.091	**-3.786**	-9%	
JAN1=4	-0.057	**-2.459**	-6%	
JAN1=5	-0.051	**-2.417**	-5%	
(XMAS=-5)*Z4	-0.009	-0.400	-1%	Mon, Tues
(XMAS=-4)*Z4	-0.042	-1.597	-7%	
(XMAS=-3)*Z4	-0.068	**-2.541**	-17%	
(XMAS=-2)*Z4	-0.211	**-7.783**	-43%	
(XMAS=-1)*Z4	-0.231	**-8.530**	-70%	
(XMAS=0)*Z4	0.019	0.689	-81%	
(XMAS=1)*Z4	0.115	**4.282**	-43%	
(XMAS=2)*Z4	0.044	1.605	-25%	
(XMAS=3)*Z4	0.046	1.681	-17%	
(JAN1=-3)*Z4	0.002	0.059	-15%	
(JAN1=-2)*Z4	-0.010	-0.368	-26%	
(JAN1=-1)*Z4	-0.115	**-4.282**	-47%	
(JAN1=0)*Z5	-0.052	-1.931	-57%	
(JAN1=1)*Z5	0.085	**3.136**	-23%	
(JAN1=2)*Z5	0.071	**2.612**	-9%	
(JAN1=3)*Z5	0.081	**3.001**	-1%	
(JAN1=4)*Z5	0.075	**2.880**	2%	
(JAN1=5)*Z5	0.053	**2.245**	0%	
(XMAS=-5)*Z6	0.041	1.661	4%	Thurs Fri
(XMAS=-4)*Z6	0.038	1.422	1%	
(XMAS=-3)*Z6	0.032	1.164	-9%	
(XMAS=-2)*Z6	0.066	**2.361**	-24%	
(XMAS=-1)*Z6	0.157	**5.620**	-56%	
(XMAS=0)*Z6	0.096	**3.452**	-79%	
(XMAS=1)*Z6	-0.207	**-7.429**	-59%	
(XMAS=2)*Z6	-0.129	**-4.542**	-37%	
(XMAS=3)*Z6	0.013	0.455	-20%	
(JAN1=-3)*Z6	-0.044	-1.538	-18%	
(JAN1=-2)*Z6	0.053	1.875	-21%	
(JAN1=-1)*Z6	0.060	**2.162**	-36%	
(JAN1=0)*Z7	-0.075	**-2.704**	-58%	
(JAN1=1)*Z7	-0.166	**-5.967**	-40%	
(JAN1=2)*Z7	-0.123	**-4.425**	-25%	
(JAN1=3)*Z7	-0.047	-1.717	-13%	
(JAN1=4)*Z7	0.009	0.357	-5%	
(JAN1=5)*Z7	0.048	1.969	0%	
(XMAS=-5)*Z8	-0.024	-0.981	-3%	Sat. Sun
(XMAS=-4)*Z8	-0.031	-1.142	-6%	
(XMAS=-3)*Z8	-0.031	-1.099	-14%	
(XMAS=-2)*Z8	-0.081	**-2.892**	-35%	
(XMAS=-1)*Z8	0.009	0.328	-62%	
(XMAS=0)*Z8	0.286	**10.250**	-75%	
(XMAS=1)*Z8	-0.005	-0.162	-49%	
(XMAS=2)*Z8	0.049	1.744	-25%	
(XMAS=3)*Z8	0.010	0.341	-20%	
(JAN1=-3)*Z8	-0.056	-1.966	-19%	
(JAN1=-2)*Z8	-0.003	-0.105	-25%	
(JAN1=-1)*Z8	0.010	0.356	-39%	
(JAN1=0)*Z9	0.155	**5.544**	-47%	
(JAN1=1)*Z9	-0.054	-1.929	-33%	
(JAN1=2)*Z9	-0.006	-0.220	-15%	
(JAN1=3)*Z9	0.011	0.393	-8%	
(JAN1=4)*Z9	0.014	0.521	-4%	
(JAN1=5)*Z9	0.031	1.287	-2%	
(XMAS=-5)*(@YEAR=1999)	0.021	0.873	2%	Outlier
(XMAS=-4)*(@YEAR=1999)	0.014	0.524	-1%	
(XMAS=-3)*(@YEAR=1999)	-0.007	-0.234	-12%	
(XMAS=-2)*(@YEAR=1999)	-0.036	-1.281	-32%	
(XMAS=-1)*(@YEAR=1999)	-0.214	**-7.678**	-70%	
(XMAS=0)*(@YEAR=1999)	-0.205	**-7.364**	-85%	
(XMAS=1)*(@YEAR=1999)	0.021	0.737	-48%	
(XMAS=2)*(@YEAR=1999)	0.029	1.027	-26%	
(XMAS=3)*(@YEAR=1999)	-0.001	-0.026	-21%	
(JAN1=-3)*(@YEAR=1999)	-0.029	-1.029	-17%	
(JAN1=-2)*(@YEAR=1999)	-0.068	**-2.414**	-30%	
(JAN1=-1)*(@YEAR=1999)	-0.132	**-4.741**	-47%	
(JAN1=0)*(@YEAR=2000)	-0.416	**-14.894**	-70%	
(JAN1=1)*(@YEAR=2000)	-0.052	-1.868	-33%	
(JAN1=2)*(@YEAR=2000)	-0.040	-1.417	-18%	
(JAN1=3)*(@YEAR=2000)	0.000	0.005	-9%	
(JAN1=4)*(@YEAR=2000)	0.012	0.446	-4%	
(JAN1=5)*(@YEAR=2000)	0.005	0.205	-5%	

Appendix:
Monthly Totals Versus Monthly Daily Averages

For those who are choosing seasonally adjustment options in Eviews and in other software I suspect, I warn you about a mistake I found myself making: opting for "trading day" adjustments of daily averages. The CENSUS trading day adjustment includes a leap year correction and assumes that the data are monthly sums. It therefore reduces the February data in leap years by 1/29. An error in the opposite direction is to seasonally adjust monthly sums without a Leap Year adjustment. These are both illustrated in Figure 32 for the three leap years: 2000, 2004 and 2008. See the big discrepancy in February, with the monthly sums too high and weekday adjusted daily averages too low.

Figure 32 **Incorrect Handling of Leap Year**