Introduction¶

GDP growth is a key metric in considering the economic status of the UK. The ONS produces quarterly growth figures as preliminary and second estimates [1]. The preliminary estimate is typically released approximately one month after the end of the quarter and is based upon data for the first two months of that quarter. The second estimate is released three to four weeks after the preliminary estimate where the figures are based upon data for the majority of the quarter. As these estimates are released after the quarter ends this present an issue of timeliness; for example, the definition of a recession is two consecutive quarters of negative GDP growth, therefore the economy is not officially in recession until seven or eight months after it starts experiencing negative growth. This delay presents a difficulty for policy makers, investors and businesses as they may need to make informed decisions based upon these figures, where even the latest can be considered 'out of date'. Therefore, it would be advantageous to produce early indicators of GDP growth that can reliably predict changes in GDP growth in advance of the GDP figure publication. Here we investigate the potential of using traffic flow data to provide early indication of GDP change.

Previous work

The UK's Department for Transport have developed a National Transport Model (NTM) [2]. The purpose of this model is to estimate future demand for transport over time periods of five years based on several factors including fuel, income, population and GDP per capita. While this model is doing the opposite of what we are attempting to show here; estimating transport demand based, in part, on a measure of GDP, the direction of the relationship between traffic and GDP is not a concern if some conditions of timeliness are met. Even if change in GDP drives change in traffic, where there is little time lag for change in GDP to have an effect and an accurate data source is available for traffic flow that precedes the data used for the GDP estimates, then it is possible to use traffic data as an early indicator of GDP.

Stats Netherlands have produced an analysis with a similar aim of investigating traffic data as an early indicator for numerous economic measures, including GDP [5]. They use traffic counts derived from road sensor data that is seasonally adjusted and aggregated up to give quarterly traffic flow counts which are then compared to the Quarterly GDP estimates. They then identify that change in traffic flow leads change in GDP by 3 months, so applying a lag of 3 months to the traffic flow figure gave the strongest correlation. This means that not only does traffic flow correlate with the GDP estimate, it leads the GDP estimate in time providing proof of concept that real-time traffic data can be used as an early indicator of GDP.

The analysis carried out here is similar to the methods used by Stats Netherlands, we identify a lag and use this to find the strongest correlations. The main differences are that the timeseries of the traffic dataset used here is longer, covering the economic downturn following the 2008 financial crisis but it only contains yearly aggregates and is therefore compared to yearly rather than quarterly estimates as in the Stats Netherlands analysis.

Data

Traffic Flow

Annual average daily flow (AADF) for Major (Roads types) and minor (road types) roads is used as a measure of traffic flow for this analysis [3]. Daily flow is the number of vehicles passing a point on a road on a day. This is averaged across the year to produce the average daily flow. This measure is

based upon approximate 10,000 manual counts per year, taken place between March and October on non-school and public holidays as well as automatic counts. These counts are used to estimate AADF figures for major roads. As it is not possible to gain direct measures of traffic flow from all minor roads, a representative sample of minor road sites are selected as observations points for counts each year, these figures are combined with the change on the previous year to estimate counts for all minor roads that are used to calculate AADF. The mean AADF figure for each vehicle type category for all roads was used as the national AADF figure for all vehicles across all roads.

GDP

The measure used here is the National GDP growth figure as contained in the UK National Accounts Blue Book [4].

Results

Comparing GDP growth and Traffic Flow

This comparison shows the standardised values (mean removed and scaled to variance) for both yearly national GDP growth (Bold, red) and national mean AADF (coloured by vehicle type) with the all vehicle mean AADF shown separately (green). The first thing to notice is the clear dip in GDP growth during 2009, equating to the 4.3% drop as a result of the financial crisis with the rest of the time series remaining approximately level.

We can also see a drop in mean AADF for all motor vehicles, Buses and Coaches, Cars and Taxis, HGVs during 2008 with a bounce back in 2009, but one that does not quite return to previous flows for Cars and Taxis, HGVs and All Motor vehicles. Pedal cycles and LGVs both show troughs for 2008, but it is not clear if this is the same trend as the others from this time series as they show a sharp peak in AADF in 2006 and 2007 respectively so it is not clear if the 2008 trough is return to a previous level. Motorbikes and Scooters also see a drop in 2008, but after an uptick in 2009, there is a decline to its 2008 level indicating the drop in AADF in 2008 may be the start of a longer-term trend.

A clear finding from this is that it is possible to observe a downturn in AADF for most, if not all the vehicle types in the year preceding the largest drop of annual GDP growth. If we are looking at early indicators of change in GDP, then traffic flow at the national level, for all roads looks a promising candidate.

Cross correlating GDP and traffic

A more robust way of showing that changes in traffic flow tend to lead changes in GDP is by using a cross correlation. This shifts the GDP time series in time, then computes the similarity (intuitively, this is the area of overlap) of the shifted GDP growth timeseries and the AADF time series across a range of different time shifts giving the cross-correlation function. Here, we look for the time shift (or lag) of the GDP time series that is most similar to the AADF time series, indicated by the lag with the largest peak.

We compute the cross-correlation function for each vehicle type. This shows that these follow a similar shape, peaking at a lag of -1 years, indicating we need to shift the GDP data to one year in the past to see the largest degree of similarity meaning that our AADF measure could be used to indicate change in GDP in advance. This is unsurprising given the previous figure, however it provides a clear unbiased indication of our AADF measure leading GDP growth supporting our initial finding.

Correlations of traffic flow and time-lagged GDP¶

Here we apply a time-lag to the GDP data so we can shift it forwards and backwards in time to compute the Pearson's correlation coefficient with traffic flow for different types of vehicles as shown by the scatter plot. We should make sure with this method that the data is normally distributed, hence we include a log transform of the traffic flow data as well as probability density functions for the GDP (right) and traffic flow (top).

With a time lag of -1 applied to the GDP time series, buses and coaches, cars and taxis and all motor vehicles both show the similar coefficients ($R_{(10)} = 0.76$, $p < 0.01$ $R_{(10)} = 0.77$, $p < 0.01$ and $R_{(10)} = 0.76$, $p < 0.01$ respectively) No other vehicle type shows correlation with GDP significant to $p < 0.05$. This finding is promising, showing that AADF for all vehicles is a potential early indicator for GDP. The fact that cars and taxis also correlate with GDP is an indicator that they make up the majority of the all vehicles AADF figure [figure here]. No other time-lag produces significant correlations.

There is still reason to be cautious over this finding; the correlation is driven largely by a single outlier at (~14600, -4.3%). If that were removed, the pattern may well indicate a negative correlation, the opposite to what we might expect. Since this point represents the dip in AADF in 2008 and the drop in GDP growth in 2009 it is perhaps the most important point in the dataset as if we were to use traffic flow as an earlier indicator for GDP, it is situations where GDP drops that we would like to get early warning for.

The other issue is one of multiple comparisons; however, since we have clear reason to look at each of the different vehicle types' correlation with GDP growth as discussed previously, we have a priori reason for performing this analysis.

Fundamentally the limitation of this dataset is that it is based upon relatively few data points (11 with the lagged data as we lose the most recent and farthest data points). In order to provide more robust datasets, we will need to use datasets with a higher temporal resolution. GDP releases are provided quarterly, that could essentially quadruple the data points. Highways England also produce a dataset based upon automatic counting of vehicles that gives traffic counts at 15-minute intervals. The issues with this dataset are that it only covers counting points along motorways and major A-roads and gives an indication of vehicle size rather than a clear categorisation. This still means there is the potential to give an almost real-time economic indicator based upon traffic flow data.

Summary

This shows that the use of traffic flow does have potential as an earlier indicator of GDP the time-series' have the greatest similarity when GDP growth has a time lag of minus 1 year, meaning change in traffic flow appears to precede GDP. The measures that best correlated with annual GDP growth that were examined here are annual average daily flow (AADF) for cars and taxis, all vehicles, and buses and coaches. There are limitations with the datasets here, as there are just eleven data points that make up this analysis. Correlations are driven by a single data outlier data point, that also happens to be the post financial crisis dip and is therefore the most important, therefore it is crucial to obtain a richer dataset with more data points with negative values of GDP to confirm if this is an outlier. Further work should be undertaken with datasets that provide a higher temporal resolution that should provide a more robust analysis and allow more potential for nowcasting/earlier indicator of GDP.