Discussion: Seasonal Adjustments

Patrick Bajari

12/9/2022

Disclaimer

- These are the discussant's views alone
- Apologies for abundant abuse of notation
- The private sector tends to approach these questions quite differently and I will share some concerns
- These may not be useful, but there may be something to learn
- More questions than answers
- Seasonal adjustments and measurement of growth particularly tricky around Covid

Why do we care about seasonality?

- There are many reasons, one is that we can’t measure growth without it
- Objectives with seasonality and growth rates:
  - -performance: how did I do this year versus last
  - -trajectory: what is my current rate of acceleration
  - -attribution: how much did some product line contribution to growth
- Many methods (e.g. X13) are model free.
- I think you need a model to get logically coherent answers for all three questions at once.
- The “right” seasonal model depends on what you want to use the model for.
- You should check sensitivity using multiple models

Typical data structure and model

- Consider a panel of time series with multiplicative seasonality:
  $\log(gms_{i,t,a}) = fixed\ effects + time + day\ of\ week + holiday + special\ events + error)$
- *i*-product group (*music*, books, personal computers, clothing, etc….
- *t*-day of week (365 days of calendar year)
- *a-area (DMA, city, etc…)*
- There is also usually an implicit hierarchy (e.g. geo sales sum up to national sales, and auto sales are a component of consumer expenditure)

Issues- Data Frequency

- Daily data makes modeling of retail seasonality much easier.
- Thanksgiving can move by up to 10 days and shopping intensity between Nov and Dec changes.
- Number of Saturdays in a month may change
- With daily data, we can model the seasonals with less coarse assumptions
- This allows for more accurate and timely measurement of growth rates

Multi-colinearity and identification

- In a decades worth of data, you still have only 10 observations for Thanksgiving.
- This can move by 10 days and change retail seasonality.
- Even worse, the seasonality around Thanksgiving may evolve.
- Variable selection guided by shrinkage or out of sample model performance (in time and space)
- Pooling across “seasonal clusters” can improve power (use ML to cluster series with common seasonal factors)
- Hierarchical models can also improve power
- Tests for multi-collinearity are important

Multi-collinearity and identification.

- Pooling across “seasonal clusters” can improve power (use ML to cluster series with common seasonal factors)
- Hierarchical models can also improve power
- Tests for multi-collinearity are important
- Using the model to estimate hundreds or thousands of “lifts” forces you to specify a more sensible model of seasonality (e.g. out of stock adjustments)
- This is inherently tedious and requires some degree of judgement.

Trend breaks

- Trend breaks are one of the most difficult parts of examining series.
- Covid hit differently across time and space.
- If you want to measure your current rate of acceleration, you need to have a view on the prior trends.
- You can do this by trying to flexibly specifying the trend.
- Problematic when you are just exiting the trend break.
- You can also model the trend break using case rates, activity rates or other auxiliary data.
- Macro shocks show up as common factors (e.g. UK World Cup game will show up as a negative effect on all series in the UK).

Hierarchies and pooling

- You have too few observations of many few seasonal factors (e.g. 10 Thanksgiving in a decade)
- Machine learning methods allow for clustering of relate series.
- Hierarchical modeling allow you to borrow strength within hierarchies.
- Borrow strength across geographies.
- ML methods allow for novel forms of nonlinearity in high dimensions.

Tough Comps

- This can occur because of a large residual or cycle.
- It can be useful to remove the residual from a YoY
- If there is a trend break (e.g. Covid) you may need to model that break explicitly across geography and time using case rates, activity or other measures.
- 2 or 3 year CAGRs useful

Conclusion

- In my experience, the modeling of seasonality depends on your objective.
- Past performance, attribution or current rate of acceleration.
- If you want to do all three at once you need a model (I think)
- Multicolinearity requires careful variable selection
- Fewness of observations suggests borrowing strength across regions, clusters or within hierarchies.
- Trend breaks make modeling seasonals and growth particularly challenging.
- Multiple methods, testing the models out of sample and using the models for practical purposes make your seasonal modeling better