Real-Time Electricity-Based Index of Economic Disruption

Updated April 15, 2020

Executive Summary

Electricity use in the Philadelphia region is about 7% below what would be expected over the last seven days, controlling for predictable factors based on long-run demand and weather data. For recent weekdays, demand is down from 7–12%.

Electricity is an important input to most economic activities. Reductions in electricity demand may therefore be useful as an indicator of disruption to economic activity. Relative to many other economic indicators, electricity data is high frequency and readily available. This means it can potentially be used as a real-time, or close-to-real-time, indicator.

The primary challenge of using electricity load data lies in extracting a useful signal of economic activity. Regular variation in daily, weekly, and yearly activity drives much demand, as does weather. We use econometric techniques to remove explainable variation weather demand, and treat the remainder as a measure of activity.

Data begin in 2000 (weekend updates) or 2006 (weekday updates) and 24 to 48 hours before the model is run (lag time for the grid operator, PJM, to post their data). Data coverage is PECO's service area, consisting of much of the Philadelphia MSA: all of Philadelphia and Delaware Counties, and much of Bucks, Montgomery, and Chester Counties. Their 1.3 million accounts serve an area with 4.1 million residents.

Figure 1 provide daily estimates of the index relative to early March; demand on recent days has been as much as 12% below trend (the effect of COVID-19 related orders are likely smaller on Sundays due to higher residential demand). The vertical lines represent school closures and stay-at-home orders. Figures 2 and 3 reveal that recent daily values of the index have been negative and are trending down. The 7-day and 28-day moving averages are below their lowest values in the last two years. (The two figures represent slightly different weather data sources and models.)

Figure 4 shows the raw load data for the PECO area since mid-February by hour-of-week, with 0 being midnight on Sunday (UTC) and 167 being 11pm on Saturday (UTC). The decline in raw usage is visible, but weather has moderated since February. Figure 5 shows log deviations by hour-of-week since the reference week of February 16–22, 2020. The most recent week is 20-30 log points below the reference week. The index in Figure 1 removes weather and other predictable components.

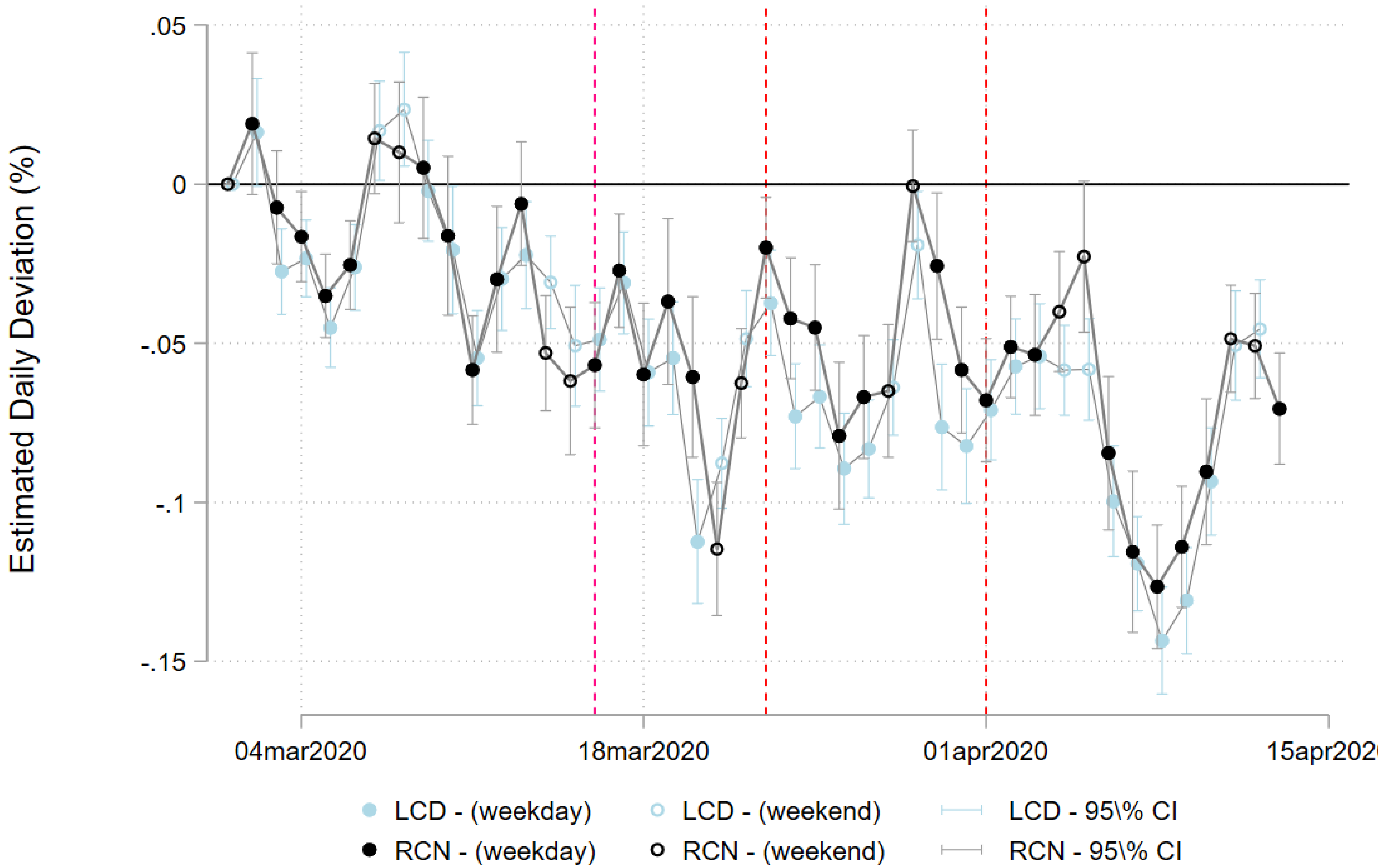

Figure 1: Index of Economic Disruption, Daily Estimates: March 1, 2020 through April 13, 2020.

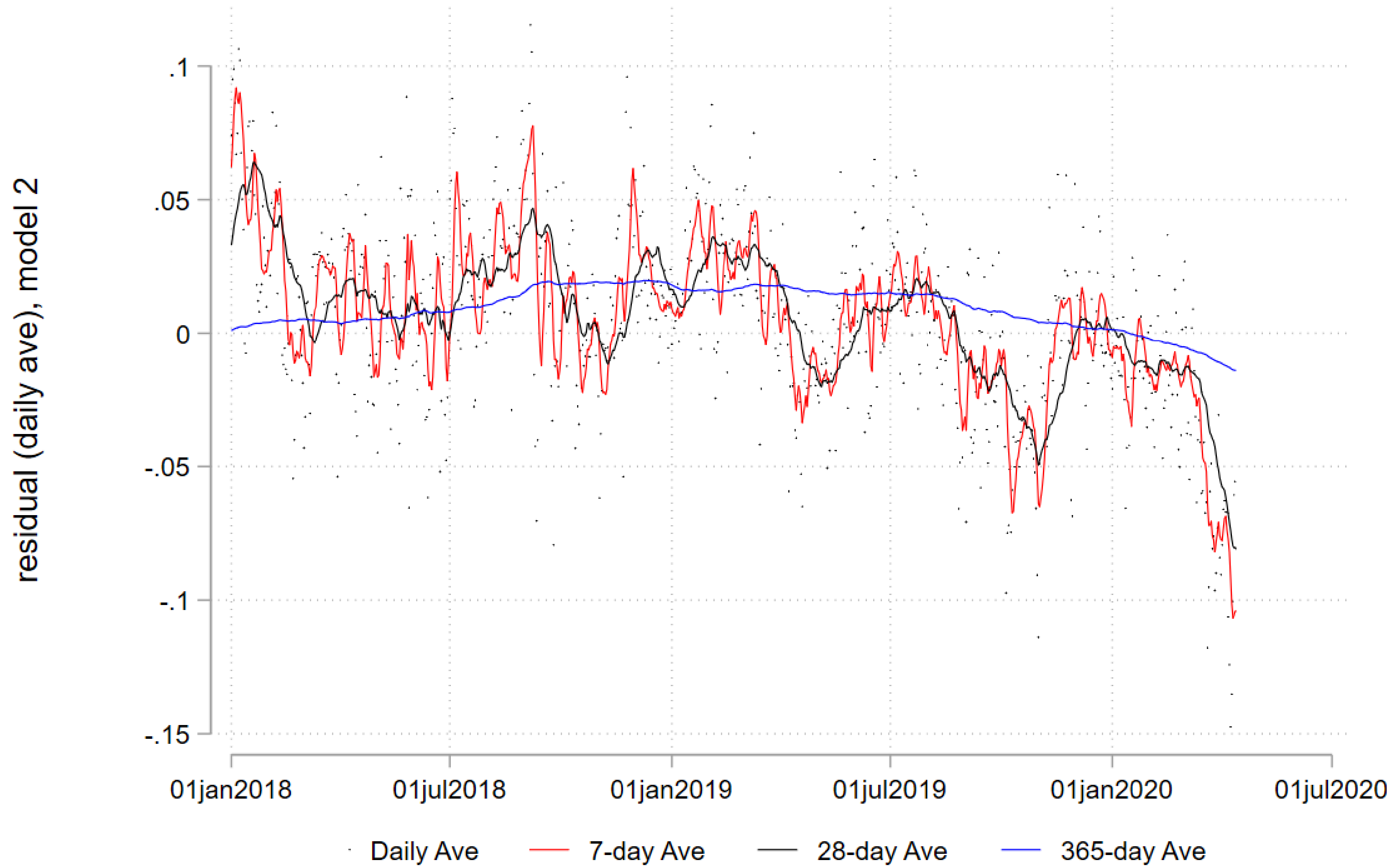

Figure 2: Index of Economic Disruption, LCD Weather Data through April 12, 2020.

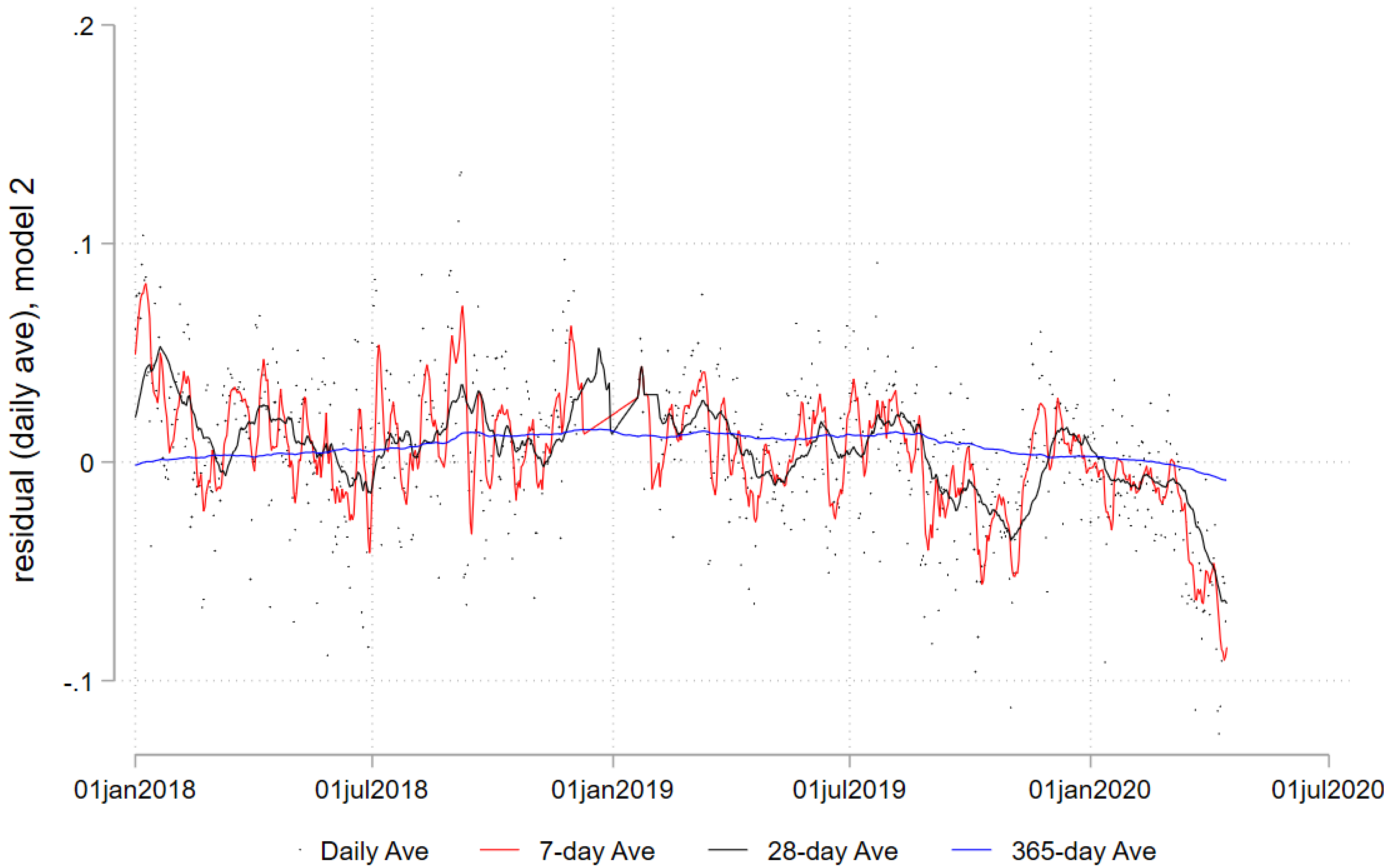

Figure 3: Index of Economic Disruption, RCN Weather Data through April 13, 2020.

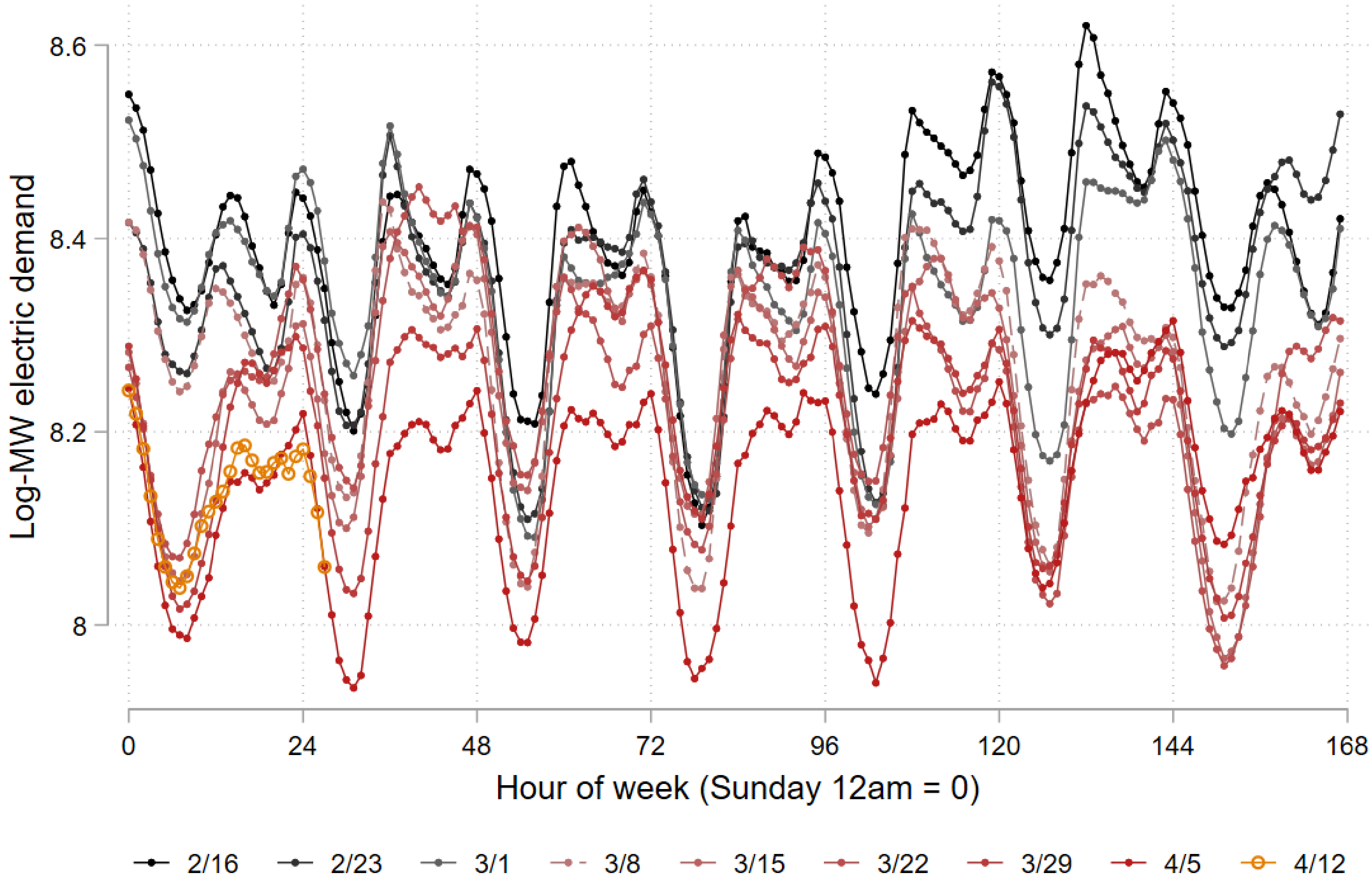

Figure 4: ln(MW) load in PECO area by hour-of-week, 2/16/2020 – 4/13/2020, times UTC.

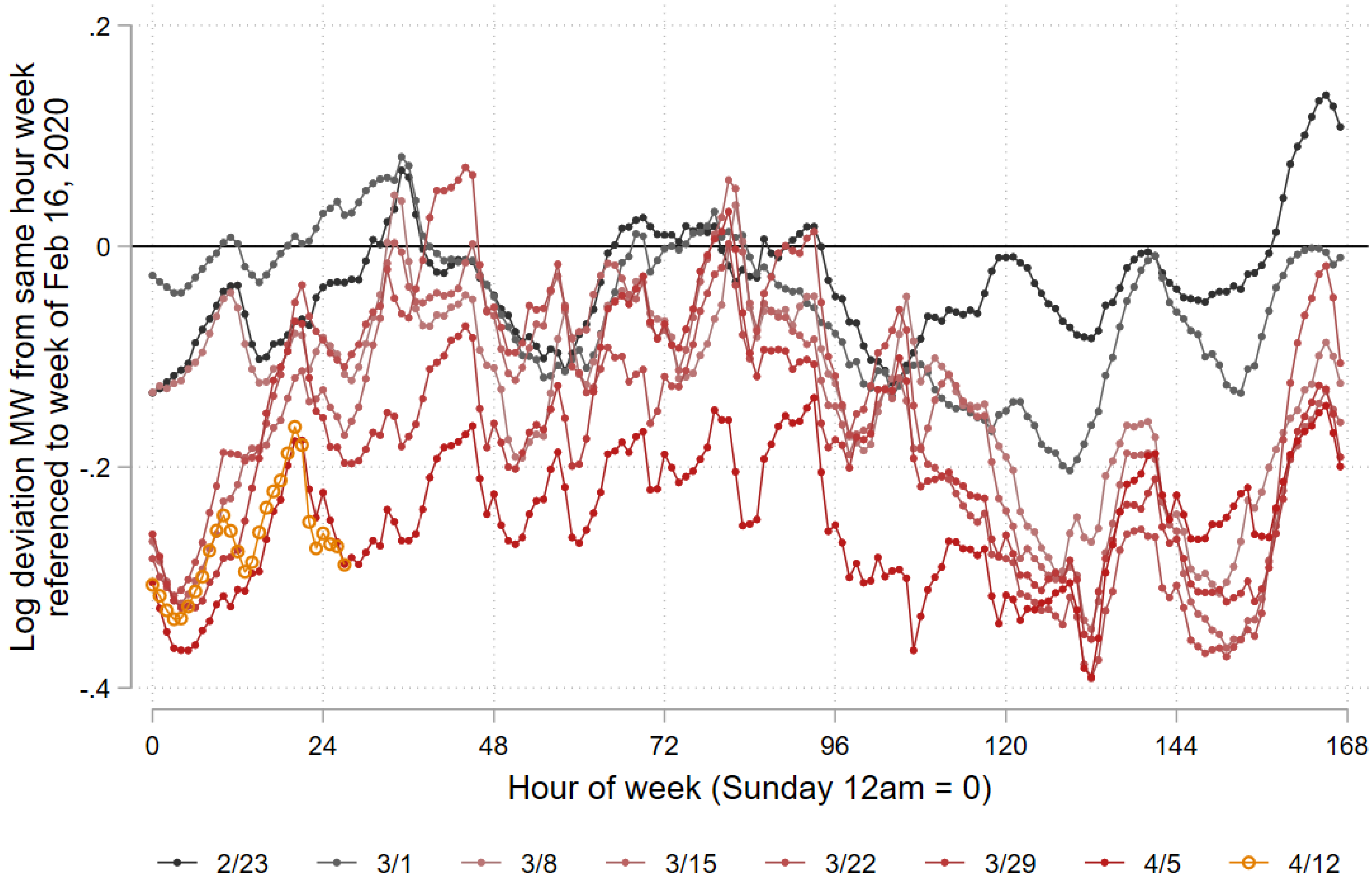

Figure 5: Deviations in ln(MW) load in PECO area by hour-of-week (using week of 2/16/2020 as reference), 2/23/2020 – 4/13/2020, times UTC.

REAL-TIME ELECTRICITY-BASED INDEX OF ECONOMIC DISRUPTION

UPDATED APRIL 15, 2020

Christopher Severen*

Goal

To develop a (close to) real-time index in the change in economic activity (economic disruption) using electricity data.

Reasoning, Strategy, and Concerns

Electricity is an important input to most economic activity. Electricity consumption can be thought of a summary measure of the current state of economic production and activity. Furthermore, Electricity data is high frequency (grid load data is easily available in hourly increments) and is updated daily by grid operators.

However, there are some concerns with the use of electricity data to model disruptions in economic activity:

1. Electricity demand closely follows weather.
2. There are predictable, high-frequency drivers of electricity demand (e.g., demand is always lower on Sunday at 3am than Thursday at 3pm).
3. Demand has become much more efficient in the last two decades (greater production per unit input of electricity; think of more efficient light bulbs, air conditioners, and machines).
4. Different electricity use across sectors reflect different changes in economic activity.
5. Some important industries and companies do not use grid-provided electricity.
6. There may be other, confounding drivers of electricity consumption.

Issues (1) and (2) can be addressed through statistical modeling techniques. Fortunately, data weather are readily available at high frequency, and predictable cyclical components easily controlled for with fixed effects. Issue (3) likely reflects a relatively slow moving change, and can be safely ignored. A related issue is the switch to distributed renewable energy. As long as this is slow, this should be ignorable. A more pernicious effect would be to alter the relationship between grid load and activity, which we cannot capture.

Issue (4) is important and cannot be addressed at this point. Because of the nature of the COVID-19 disruption, there may be load displacement from commercial (and manufacturing) locations to

*Jonás Arias, Thorsten Drautzburg, Kyle Mangum, and Arthur van Benthem have provided valuable input. Nathan Schor has provided excellent research assistance.

residential locations. Similarly, Issue (5) is challenging to address. Large manufacturers sometimes produce their own power. We do not address this issue here, which is akin to assuming it constant across our period of study. In principle, both of these issues could be solved with load or demand data broken out be sector. However, the allocation of demand to sectors does not take place at the level of the grid operator; individual utilities typically make these accounting distinctions. We will pursue this data with the local utility.

Issue (6) is challenging. We will try to deal with this through time averaging (to remove the most volatile components), but this may not be totally sufficient.

Data

Hourly electricity data are from the 'PE' (PECO) load area of the PJM interconnection. This corresponds to PECO's service area, which consists of Philadelphia County and its PA suburbs (all of Delaware County, and most of Bucks, Montgomery, and Chester Counties). In total, this area covers about 1.3 million accounts and 4.1 million people. Hourly data allow careful modeling of weather and predictable (cyclical) drivers of demand.

Hourly weather data are drawn from one of two sources: 1.) NOAA's Local Climatological Data (LCD), or 2.) the US Climate Reference Network (CRN). LCD data are from the observation station at the Philadelphia Airport; CRN data are from a station in Avondale, PA. The LCD data provide a longer time series and contain more weather variables, but are updated less frequently (no updates during the week) and are more susceptible to equipment malfunctions resulting in suspicious readings. CRN data begin more recently and contain fewer variables, but are update more frequently and are less likely to have problematic observations. Unfortunately, CRN data contain several periods with no data. Primary variables of interest are temperature, humidity, precipitation, and wind.

The sample is restricted to begin on January 1, 2000.[2] Figure 6 shows the raw hourly time series of log megawatt electricity demand. Note the low values in 2012; this corresponds to Hurricane Sandy. Figure 7 shows the most recent 15 months with a local linear moving average. This provides initial motivation for this analysis; recent readings are low. There does appear to be a downward trend in March 2020, though this may predate COVID-19 or reflect weather or other seasonal factors.

[2]Prior research found evidence a structural break in the path of electricity stochasticity around 2000 in Delmarva electricity demand (Arias & Fernández-Villaverde 2018).

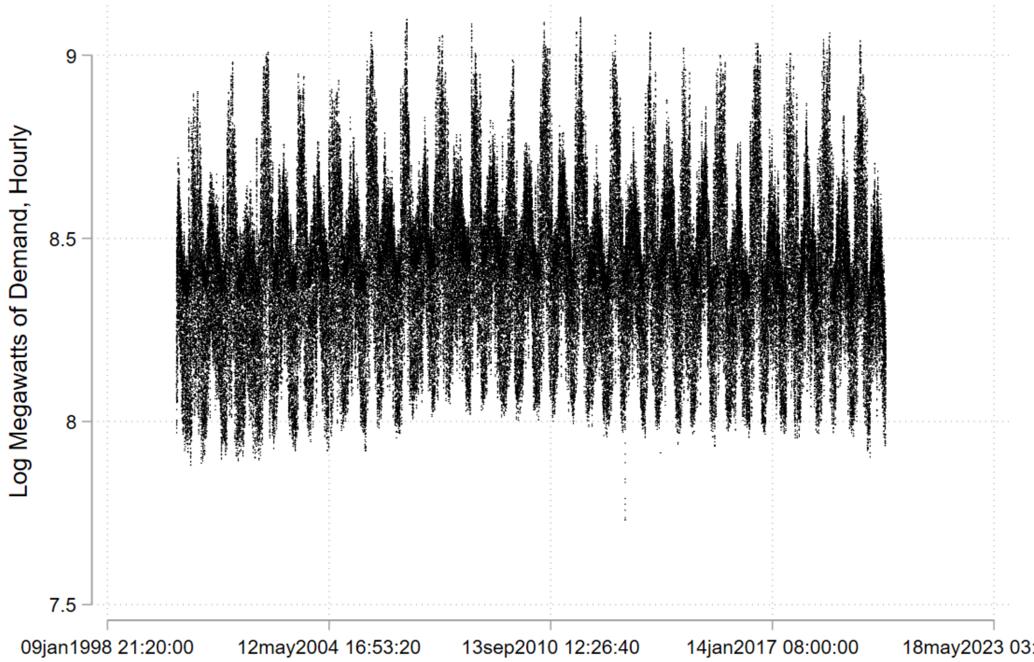

Figure 6: Hourly Electricity Load since 2000.

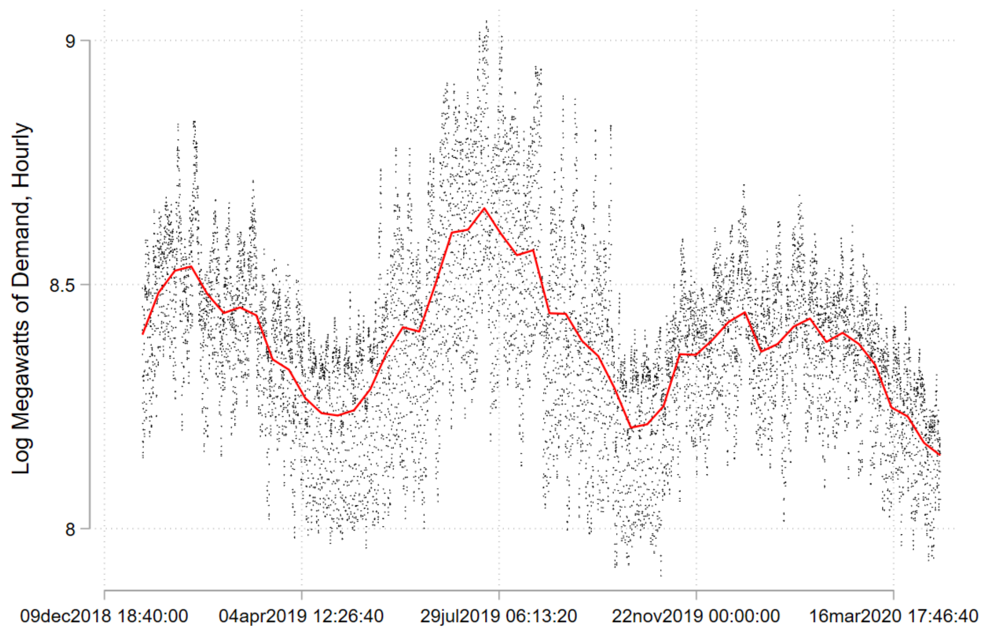

Figure 7: Hourly Electricity Load since Jan 1, 2019, with local moving average.

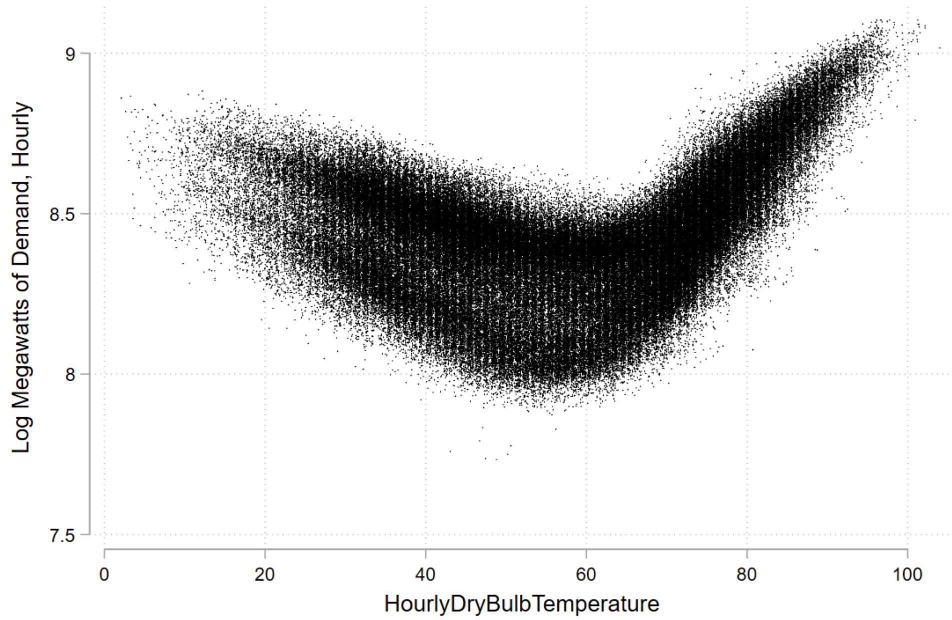

Figure 8: Temperature and Load.

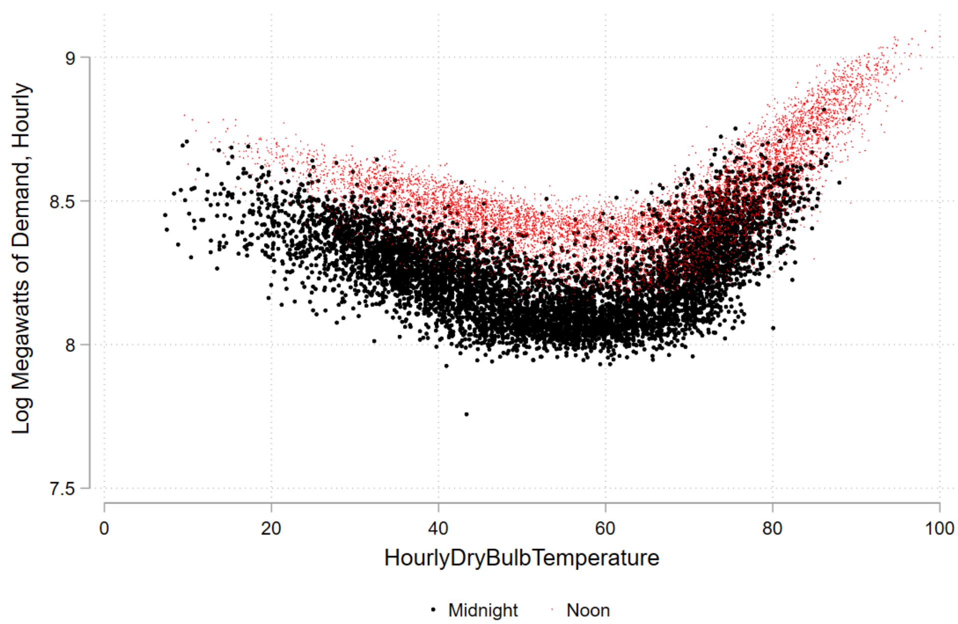

Figure 9: Temperature and Load, by Time of Day.

To highlight the importance of controlling for weather variation, Figure 8 plots the relationship between (dry bulb) temperature and load. This relationship is not linear. Furthermore, there appears to be two lobes defining the relationship between lower temperatures and demand. This represents differences in demand by time of day (see Figure 9), and underscores why controlling for seasonality will be important. Buildings are typically kept cool all night during summer (because cooling air requires reducing humidity, which is costly). In contract, it is not very difficult to let a building cool a bit in the winter, and then warm it back up when people arrive again.

Methods

I use a linear model supplemented with high-dimensional fixed effects to purge the electricity data of regular cyclical variation and weather driven components. Specifically, I estimate:

$$\ln(MW_{hdy}) = \mathrm{L}[\ln(MW_{dy})] + f(\text{weather}_{dy}) + \delta_{h\times\text{doy}} + \mu_{h\times\text{dow}} + \text{events} + g(\text{trends}) + u_{hdy}$$

for hour h on day d of year y. L[] represents (autoregressive) lags of *daily* electricity load. In most models, I include the 364, 365, and 371 lags, primarily to capture changes in seasonality.[3] I use lagged values of daily average demand rather than hourly values for model simplicity. Shorter-frequency lags are avoided so as to not contaminate estimates of the COVID-19 event (using a 7-day lag might misattribute low load this week to low load last week, rather than to the event).

The precise specification varies depending on which weather-data source is used. In both specifications, I use:

- One-degree temperature bins interacted with hour of day. The binning approach flexibly captures non-linearities in electricity demand related to weather, and the interaction with hour of day captures the issues presented in Figure 9. Degrees are F for LCD data, and C for CRN data.
- A quadratic in precipitation, interacted with whether it is warm or cold outside (above or below roughly 60F).
- Relative humidity, interacted with whether it is warm or cold (above or below roughly 60F), and with temperature. This is captures the much greater load placed on cooling systems on humid, hot days than dry, hot days.

Additionally, the LCD data contain windspeed and direction information, which I also include in this model.

The fixed effects capture repeating, predictable variation in demand:

[3]The 364-day and 371-day lags are divisible by 7 and so capture excess weekly variation, and the 365-day lag (366-day lag in leap years) captures excess annual variation.

- Hour-by-day of week fixed effects. This captures the average daily cycle of electricity demand across different days of the week. This flexibly controls, for example, for average differences in load at 3am on Sunday, at 3pm on Sunday, and at 4pm on Thursday.
- Hour-by-day of year fixed effects. This captures the average daily cycle of electricity demand across different days of the year. This flexibly controls, for example, for average differences in load at 1am on December 25th, at 2pm on December 25th, and at 4pm on June 28th. This also captures a considerable amount of seasonality in demand; with the rich weather data, seasonality is well controlled for.

The trend terms are a shorthand way to try to fit a moving average model, and remove aggregate, slow-moving trends (up to a quadratic or cubic in days).[4] Events capture special power grid events; for now, this is primarily an indicator for Hurricane Sandy.[5]

One variant of the models (used to produce Figure 1) includes an indicator for a baseline period from February 1, 2020 – March 1, 2020 and indicators for each day since March 1, 2020, to provide estimates of the impacts so far of COVID-19. This model is other identical to the others, but is less likely to lead to bias in the estimated fixed effects and weather coefficients due to the use of current period data (that potentially reflects reductions induced by COVID-19). Standard errors are calculated allowing for up to 28 days (672 hours) of autocorrelation in the error terms.

Results

This model is highly explanatory. The hour-by-day of week and hour-by-day of year fixed effects explain about 75% of the variation in electricity demand. Including the weather variables allows the model to explain about 94.6% of the variation in electricity demand.

Hourly residuals, $\hat{u}_{hdy}$ are produced after estimation. These are a bit too volatile and numerous to be studied individually, so daily averages are taken and used as the index of disruption. Even then, the series is volatile, so additional smoothing techniques will be employed.

Estimates of the residual over the length of the sample, along with 7-day and 365-day moving averages, are shown in Figures 10 and 11. Versions of these graphs, focused in on the period following 2018, are shown in Figures 2 and 3 in the Executive Summary.

[4] I have not found a straightforward way to estimate an ARIMA-type model with many fixed effects.

[5] Estimates show that Hurricane Sandy was accompanied by a 31%-35% reduction in electricity demand; the grid operator and utilities intentionally cut power along many transmission lines to ensure minimal damage and fire risk.

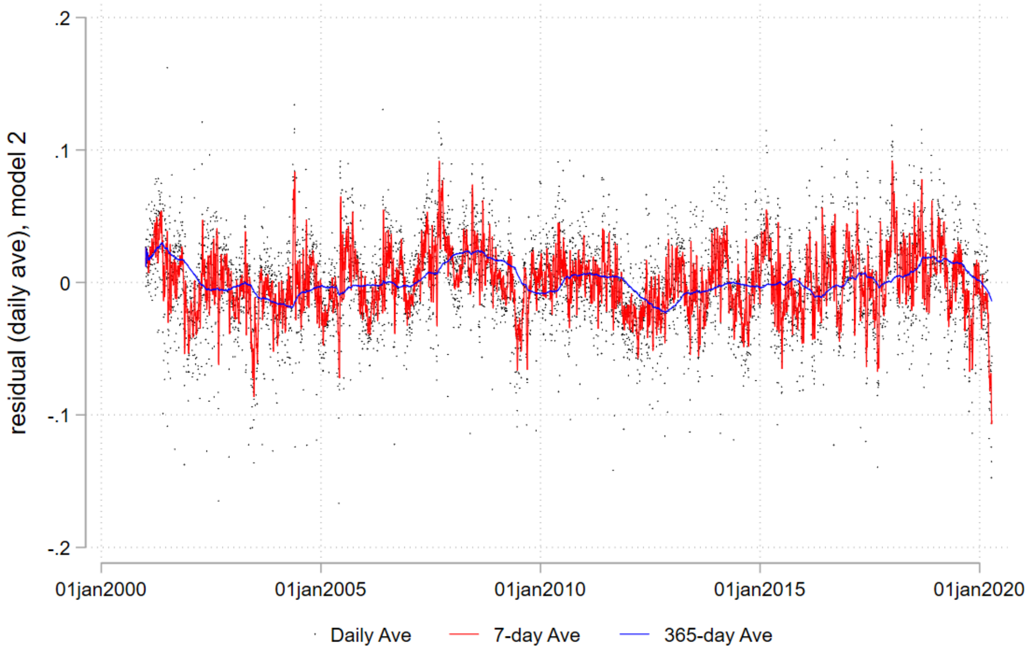

Figure 10: Index of Economic Disruption, LCD Weather Data. Full Series.

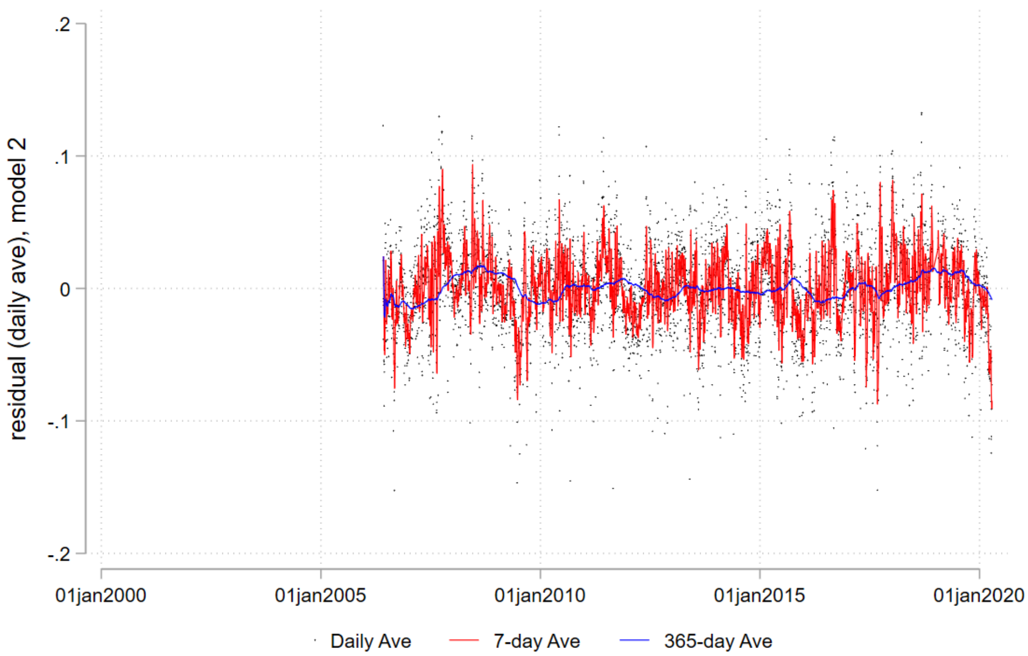

Figure 11: Index of Economic Disruption, CRN Weather Data. Full Series.

Interpretation, Benchmarking, Scalability, and Next Steps

We know from a variety of sources (e.g., initial unemployment claims) that the economy is responding quickly and strongly to COVID-19 and consequent policy actions. This electricity-based index may not capture the full severity of these impacts, largely because there will likely be substitute activity that smooths demand (Netflix takes a lot of energy to run). The magnitude may be somewhat information, but perhaps a better interpretation will be understand when the "bottom" has been hit, and how recovery is going (when and if the economy is "back to normal").

To help understand how useful this exercise is, we can attempt to benchmark this data to other events and recessions. The use of Hurricane Sandy captures one effect: a short (1-day) drop of greater than 30% of load demand (though in this case, the storm directly reduced the ability of the grid to supply load). I hope to provide more careful benchmarks to local unemployment data in following iterations.

The approach here could be improved in several dimensions. However, all are relatively costly:

- Modeling different geographic areas might reduce some of the residual volatility in the index, and allow extraction of a cleaner signal. The challenge is that data come from disparate providers, and take significant investments in cleaning. Steve Cicala (Chicago Harris) is working on a such a project.
- Pursuit of sectoral demand data to report separate series for residential, commercial, and industrial loads. I do not believe such data exist in real time (rather, they are created only after metering is completed monthly). I hope to get this data from PECO.
- Formal moving average or ARIMA modeling. The challenge here is that such models with many fixed effects are difficult to fit.

One advantage of the relatively similar approach outlined above is that it is scalable. This model could be easily implemented for every area with good data coverage, and results aggregated. For analysis that has occurred at similar large scales, see Auffhammer, Baylis, & Hausman (2017).

References

Arias, J. & Fernández-Villaverde, J. (2018). Tracking Business Conditions in Delaware. *Economic Insights* (Federal Reserve Bank of Philadelphia).

Auffhammer, M., Baylis, P., & Hausman, C. H. (2017). Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States. *Proceedings of the National Academy of Sciences*, 114(8), 1886–1891.