Class -

Whether or not this is your first time reading an economic journal article, this is likely the first time reading a journal article with a critical eye towards understanding the econometric methods. I've written up a little bit here to help guide you as to what to expect and what to look for. I've added notes in line as well, and have tried to point out the key concepts that we have talked about in class.

This paper follows a standard pattern for journal articles: intro, data, methods, results, discussion, conclusion. Here's what to look for in each.

Introduction
This is where the research question is motivated, and the setting is introduced. It gives background on the question, reviews the existing literature on the topic, summarizes the method, and teases the result.

Here, the question is about the impact of a unique solar panel financing program called PACE on the rate of residential solar panel installations installations in a city with this policy. It's an interesting policy because it's not a subsidy - it charges a market interest rate on a loan used to install solar, but the loan effectively stays with the house when someone moves. Pages 359-361 make the case for why this is an interesting thing to study by grounding the question in theory. Admittedly, the theory is not very rich, which we make clear.

The unit of observation, *i,* is "city" and time, *t* , is "quarter", so we have panel data.

Sample and data
This section introduces all of the data sources we will be using, and the covariates (x's) we'll be including as controls - city-specific data on demographics (median age, average education, etc.). Table 2 on page 366 shows all of the variables (the treatment / variable of interest is PACE but isn't on the table).

This section also discusses whether or not we dropped or eliminated any cities from the sample. It is very common to have two or more different slices of the data and report regression results for all of them (hopefully with similar results!). We do that here mainly because some cities don't have demographic data.

Methods and results
Difference in differences
You'll recognize a lot of the terminology in the Difference-in-differences method section. Parallel trends assumption is front and center, and we discuss things about the data that suggest it is (or is not) a safe assumption. Note that we are concerned that there seems to be a "jump" just before treatment starts. We'll talk in class why this might be of concern to our identification strategy.

DID results are in Table 3. The "main result" from our "preferred specification" is the first coefficient in the first column, 3.82. A PACE program has a causal effect of an increase in watts installed per household per quarter of 3.82, ceteris paribus. The specification of the regression is just below Table 3.

Synthetic Counterfactual Method
We might be a little worried about the parallel trends assumption, so we use another method called the Synthetic Control Method.

The gist of the method is this: we are trying to find a counterfactual for each treated city in the data. If we had a perfectly matched (but untreated) version of the two treated cities, we could just compare the two and say that the effect is the difference - essentially saying that Y_{i0} is the same as Y_{j0} for some city *j* that really really really looks and acts like our treatment city, *i* .

The question is, how do we determine which city is the best match? There are lots of methods that match on observable things - find a city in the same state with the same size population and same median age and same average income, and maybe it'd be similar. But there are lots of unobserved things that would get left out - what if one city was dependent on the auto industry while the other has a burgeoning but unobserved tech industry?

What the synthetic method does is look at both observable covariates *and* pre-treatment outcomes and tries to pick a control that matches them closely. This main idea is that those unobserved things can't be matched directly because they're unobserved, but you can match on the ways the pretreatment outcome changes due to these unobserved things.

The synthetic method goes a little further than finding the one closest match to each treatment city. It comes up with weights for each city that, combined, make the synthetic. For instance, one of the earlier uses of this method was to estimate the effect of cigarette taxes on smoking. California had taxed cigarettes, so the authors needed a synthetic California. Using this method, it turned out 35% Utah, 20% Colorado, 12% Connecticut, and a couple other states, once weighted and summed, made a pretty good California!

There will be some matrix algebra in the methods section here, do your best with it, but just keep in mind we're trying to make a synthetic city for each of the treatment cities (Palm Desert, Yucaipa, and Santa Rosa) where the pretreatment outcomes look pretty similar. Figures 5, 6, and 7 show the synthetic (the gray boxes) and the actual outcomes (the gray lines). Especially in 7, in the pretreatment periods to the left of the first line, the actual outcome is really similar to the gray boxes (parallel trends!), but after treatment the actual is much higher than he gray boxes - that's the treatment effect!