**PhD Progress Self-Review**

**Leon Stirk-Wang**

**1 year report**

**1. Overview of Project**

The thesis as a whole will comprise three chapters. The common thread across the three chapters will be the application of matching as a sample pre-processing method for causal inference in observational datasets in the field of real estate economics. Each of the three chapters will be a different application of matching techniques to important areas in the current literature. To illustrate the project and it’s aims I will use one of my chapters as an example.

An important area of the literature is the estimation of the effect of natural hazards or the risk of natural disasters on properties. This is often used to study and inform policy responses to natural hazards. Flooding is one such hazard that may have an impact on property values. The standard way of estimating the impact of a flood event on house prices is to use a ‘difference-in-differences’ regression. Such a model estimates the effect of the flood on the properties that flooded (the difference in the prices before and after the flood) and compared that to the difference in prices of a comparison group (that did not flood). Thus if there is some component of the change in prices that is not attributable to the flood event then the comparison group should separate that from the true effect of the flood (implicit control of confounders). Such a comparison group provides us with a ‘counterfactual’. An estimate of what would have happened to prices in the absence of a flood. By comparing the effect of the flood on prices to this counterfactual we should be able to determine the true effect of the flood.

The problem is that counterfactuals are, by their nature, only an estimate of what would have happened based on what happened to houses that did in fact not flood as we cannot observe what would have happened to the houses that did flood (in the absence of a flood). As such, a good estimate of the counterfactual relies on the houses that did not flood being comparable to the houses that did flood (a comparison of like-to-like). If there are systematic differences between the two groups then the integrity of the counterfactual is undermined.

This is a property that is often exhibited in observational studies/datasets as the assignment of the treatment (whether a property floods or not) is not assigned in a way that conforms to the framework of good research design i.e. a fully-blocked, randomised controlled trial. As research of this nature is just about impossible to arrange with respect to real-estate it is a field particularly well suited to the application of matching techniques.

Matching provides a systematic and consistent way of correcting for the problem of poor counterfactual estimates by generating a more appropriate comparison group and therefore a more appropriate counterfactual.

My thesis will look at three areas of the real-estate economics literature and examine the effects of improving the counterfactual estimates using matching techniques. At the moment, one chapter will be dedicated to the causal effect of flooding on prices, another will be the causal effect of time on prices (price indexing). The specifics of the third chapter have yet to be determined although I wonder if it might be worthwhile retaining a thread related to natural hazards or climate change. I am currently thinking about using Christchurch residential property sale data.

**2. Overview of activities in the last reporting period**

In the last 6 months (since my last review) I have obtained a relatively complete set of empirical results for what I am calling my second chapter (on flooding). I have drafted substantial portions of this chapter.

Paul (supervisor) purchased some additional data required to complete the analysis before starting the research.

We’re also in the process of obtaining data that will potentially be used for the third chapter.

The chapter on indexing (first chapter) has been on the backburner for the last three or four months while I’ve been working on the second chapter. I have made some progress although I feel like the methodology requires additional refinement prior to reporting the results. I’ve drafted some theory and literature chapters associated with the indexing project.

**3. Issues**

No issues

**4. Changes**

I don’t think there have been substantive changes in the thesis. While my proposal and six-month self-review probably look slightly different, I feel like the research topics have started to solidify quite well between now and my last review. The common thread of matching and real-estate still underpins the thesis as a whole and any changes in direction have been guided in part by gaining a more thorough understanding of the current literature.

**5. Outputs since last report (use any of these headings which are relevant in your case)**

* Department seminar

**6. Activities and other highlights (use any of these headings which are relevant in your case)**

* None

**7. Employment**

Employed by the university/economics department as a tutor.

**8. Work Plan**

* Build additional residual diagnostic mapping tools for flooding chapter
* Complete additional matching diagnostics (flooding chapter)
  + CEM matching comparison
  + Global distance optimisation matching comparison
  + k-to-k and non-greedy weighted matching comparisons
* Submit drafts to supervisors in February
* Finish analysis scripts for indexing chapter
* Build visualisation and diagnostic tools for indexing chapter
* Produce results and submit drafts (indexing) to supervisors maybe August/September 2019
* Third chapter (unsure of specifics)

**9. Expected Submission Date**

* November 2020

**10. Thesis outline and writing schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **Chapter** | **Section** | **Drafted** | **Completed** |
| Matching methods | Intro | 10% | No |
|  | Theory | 50% | No |
| Indexing | Literature | 30% | No |
|  | Data | 20% | No |
|  | Method | 10% | No |
|  | Results | 10% | No |
|  | Conclusions | 0% | No |
| Flooding | Literature | 80% | No |
|  | Data | 80% | No |
|  | Method | 80% | No |
|  | Results | 80% | No |
|  | Conclusions | 80% | No |
| Third chapter | Literature | 0% | No |
|  | Data | 0% | No |
|  | Method | 0% | No |
|  | Results | 0% | No |
|  | Conclusions | 0% | No |

**12. Attached Documents (may not be used for most review reports and/or candidates)**

If there is anything you particularly want to discuss at the review meeting, or report back to the PhD Convener, feel free to attach it/copy it to the of the self-review document. It would be helpful if you had a sentence at the end explaining the reason behind any attachments.