**CAPP 30254 - Machine Learning for Public Policy**

**Project Proposal**

Allison Collins, Sinclair Target, Erika Tyagi

1. **Background and Goals**

Eviction is the removal of a tenant from a rental property – typically due to late payment, but it can also be a result of other causes such as property damage, disturbance, or illegal subletting. It remains a key challenge because of the lack of affordable housing, which causes families and individuals to spend a greater amount of income on rent and subsequently puts them at risk of eviction.

Housing instability in turn underpins a number of social problems, including poverty, education, and healthcare access. Children may have to switch schools; families’ belongings are left on the street or can be taken, requiring a fee to get them back; and there are a number of studies which indicate higher levels of depression and other mental health issues among those who have been evicted. Addressing these challenges cannot be successful without first tackling housing security—and many cities have recently made moves to address this challenge.

In particular, our project aims to explore this issue and related interventions within the city of Philadelphia, as they are considering adapting legislation similar to that in NYC to provide universal legal coverage to tenants to reduce evictions. Philadelphia ranks fourth on the list of cities with the highest number of evictions at 10,264 in 2016, according to the dataset from Eviction Lab. Further, a recent task force report from the mayor’s office highlighted a number of other compelling statistics – 20% of people entering homeless shelters in Philadelphia cited eviction as the cause; 1/14 renting households have been subjected to an eviction filing; and there is an extreme dearth of affordable housing, with only 24 units available for every 100 extremely low-income renter households. Particularly relevant to the question of legal interventions, this report further found that 81% of the time in eviction court cases landlords were represented, while tenants were only represented in 8.5% of cases.[[1]](#footnote-1)

Eviction impacts a number of stakeholders—first and foremost, there are those who are at risk of eviction. Many individuals do not know what their options are upon receiving an eviction notice. They lack the resources to pursue how they could stay in their homes, and in many cases, the necessary legal information and support. Another set of key players is the community organizations who advocate for these individuals at risk of eviction. These can include organizations ranging from Philadelphia Tenants Union who do grassroots advocacy to highlight landlords’ illegal practices and fight for resources and legal change, to Community Legal Services of Philadelphia, which provides support across a range of topic areas, including a tenant resources guide for those at risk of eviction. Finally, local government sets forth the legislation related to eviction and supportive/preventative services, including funding such programs, and local courts are responsible for implementation and trying of cases.

We will investigate how to prioritize availability of government-provided legal services to tenants at risk of eviction based on highest need. This could help direct resources to areas where eviction risk is higher, in support of an intervention as outlined above which would drive toward universal legal coverage for tenants at risk of eviction. Specifically, this prioritization would be used by the Philadelphia city council allocating funding during the rollout to full availability of legal resources, and then to prioritize the distribution of presumably limited resources at the end stage. Additionally, accompanying information on these services could be distributed by either the government or advocacy groups to ensure that people will take advantage of them (since as can be seen in the relative take-up across boroughs in NYC, messaging of their availability is key).

In sum, the goal we are driving towards is reducing the number of evictions in the city of Philadelphia, by optimally allocating legal support resources through predicting evictions. This will roll up into the city of Philadelphia’s multi-pronged effort to combat the eviction epidemic.

1. **Current Solutions and Proposed Interventions**

As alluded to above, activists in Philadelphia have the city council and mayor considering emulating New York’s universal access law. A report commissioned by the Philadelphia Bar Association calls for a $3.5M annual investment, which it believes could provide legal services to all tenants unable to afford representation (and it believes that this could avoid $45.2M in other costs, for instance resources for the homeless, annually).[[2]](#footnote-2) In 2017, New York’s city council implemented The Universal Access to Legal Services law to provide free legal services to tenants whose landlords try to evict them, currently available to four zip codes in each of the boroughs (the aim is to offer this to all residents who make up to double the federal poverty level by 2022). It has seen early success: by Q4 FY2018, 30% of tenants appearing in Housing Court for eviction cases were represented by attorneys, as compared to only 1% in 2013—and in neighborhoods targeted for the intervention, the legal representation rate was 56%. Notably, in eviction cases resolved in FY18 citywide, 84% of tenants were allowed to remain in their residences (previously, ~50% remained in their homes).[[3]](#footnote-3)   
  
Such an intervention would also need to be rolled out in phases in Philadelphia, and in fact, in 2017, Philadelphia began piloting the delivery of such free legal services to residents in specific zip codes. The Philadelphia Eviction Prevention Project allocated a $500,000 investment to expand access to legal services and information and anticipated serving 1,800 tenants over the course of 2018 (both through direct legal representation and/or advising). Notably, the aforementioned report by the Philadelphia Bar Association found that only 5% of tenants in Philadelphia who have attorneys are evicted through eviction proceedings – but among those who are not represented, that number increases to 78%.[[4]](#footnote-4) This work to provide legal services to poor tenants is part of a broader effort created through a task force set up by the mayor, which will also include steps such as the creation of a database of offending landlords, providing counseling services and emergency financial support, and educating tenants on rights and responsibilities.  
  
Our proposed intervention is to continue and grow funding universal legal access for tenants at risk of eviction through local legislation, with information-sharing standards in place so that eligible tenants are aware of and able to utilize these services. These legal services would help reduce the eviction rate through eliminating evictions that happen because tenants are unaware of their rights and unable to advocate for themselves. As will be discussed in further depth later on in the proposal, our aim is to use supervised regression models to predict evictions and the gap between filings and evictions at the block group level. We will compare if this suggests the same parts of the city and use this information to inform allocation of resources for legal advising and an accompanying information campaign both in prioritization for first stages of rollout, and eventually to optimize distribution among areas in steady-state.

1. **Data and Analysis**

The primary data that we intend to use is provided by the Eviction Lab at Princeton University.[[5]](#footnote-5) This data combines information on formal eviction records with Census demographic information. Notably, this data only includes court-ordered evictions made publicly-available by municipalities (i.e., it excludes informal evictions, evictions from court-sealed cases, and evictions from localities without archived data).  
  
This data is available annually at the Census block group level. Philadelphia County (conterminous with the city of Philadelphia) spans 1,336 block groups in the dataset, all of which have complete evictions data from 2002 through 2016. Thus, the relevant dataset that we intent to use to build and evaluate our models includes 20,040 records (1,336 block groups across 15 years). Because the Eviction Lab provides data from municipalities across the country (albeit with varying levels of completeness), we can also create models incorporating data from other cities and evaluate how this affects model performance.  
  
For a given block group and year, the following eviction information is available:

* The number of eviction judgements in which renters were ordered to leave
* The number of eviction cases filed

And the following demographic information is available:

* The total population
* The percentage of the population with income in the past 12 months below the poverty level
* The percentage of occupied housing units that are renter-occupied
* The median gross rent
* The median household income
* The median property value
* The median gross rent as a percentage of household income
* The percentage of the population that is White, African American, Hispanic, Asian, etc.

From this data, an eviction rate (i.e., the ratio of evictions to renter-occupied households) for each block in a given year is also provided, along with an eviction filing rate. Each record also includes information indicating if that record's data is imputed, subbed, from non-Eviction Lab sources, and estimated to be lower than in reality.

The demographic data provided in the Eviction Lab data relies on the following:

* 2000 Census for 2000-2004
* 2009 5-year ACS for 2005-2009
* 2010 Census for 2010
* 2015 5-year ACS for 2011-2016

Because the data is provided at a Census block group level and includes the Census FIPS code, additional Census data can easily be incorporated to supplement the existing information at the block group level.

The fundamental goal of this analysis is to identify the block groups that are likely to have the highest eviction rates in the future, which requires predicting eviction rates by block group. In a machine learning context, we're solving a supervised regression problem where our target is the eviction rate for each block group. Separately, we're also interested in predicting the difference between the filing rate and the eviction rate (i.e., the difference between eviction cases filed and judgements in which renters were actually ordered to leave), as this gap appears to be widening over time.

Given that our model must be appropriate for regression where the target is a continuous variable, we intend to compare the following types of methods:

* Simple linear regression
* Polynomial regression
* Support vector regression
* Decision tree regression
* Random forest regression

Because our data spans 15 years, we intend to apply temporal validation using various splits, with the goal of identifying the best model to predict future eviction rates. As previously noted, we intend to consider including data from other cities to train our models and compare their performance against models trained only using data from Philadelphia.

1. **Validation**

Our project aims to reduce the number of evictions across Philadelphia by more efficiently providing legal aid to tenants that otherwise would not be represented in housing court. We would like to find a model that, over the years we have available, reduces the city-wide number of evictions by the largest average amount. In order to estimate how much each model reduces the city-wide eviction count, we have to make a number of assumptions.

We assume that the city will not be able to provide legal aid to everyone. The Philadelphia Bar Association has suggested that $3.5 million in funding would be sufficient to provide legal aid to all tenants earning below 30% of area median income, but as of this writing the city has only allocated $500,000 for legal aid. We therefore conclude that the city is only be able to provide legal aid to roughly 14% of the census blocks within the city limits. Philadelphia’s pilot program now offers legal aid in a select few zip codes chosen for a variety of factors; we would like to see if we can prevent more evictions by instead targeting the top 14% of census blocks as ranked by predicted eviction rate.

We also assume that providing legal aid to a census block prevents 79% of the evictions that would otherwise have occurred within that census block from occurring. This number is based on the Philadelphia Bar Association report, which found that tenants without legal representation are evicted in 95% of cases whereas tenants with legal representation are evicted in only 16% of cases.

Using these two assumptions, we can estimate how much each model reduces the number of evictions city-wide by subtracting from the city-wide total the prevented evictions in the census blocks that were prioritized by the model. The models that more accurately predict eviction rates in each census block should do better according to this metric.

We will also report the RMSE (root mean-squared error) and adjusted R^2 value for each of our models as another metric for model accuracy.

1. **Caveats and Policy Recommendations**

Our analysis will ultimately produce a list of the census blocks that the city should provide legal aid to first. Prioritizing these census blocks should be the most efficient way to use the city’s limited resources until enough funding is made available to help everyone.

We cannot guarantee that our list will always be correct. Because we are making simplifying assumptions about the intervention resources available to the city and the effectiveness of providing legal aid within each census block, our key evaluation metric is only an estimate. If the model is used in practice, each year the city should compare the number of evictions prevented by the model with the number of evictions that could have been prevented by targeting the top 17% of census blocks as ordered by their actual eviction rates for the year.

There is also at least one methodological and one ethical issue to consider. The methodological issue is that our data only includes formal evictions that have been recorded at a municipal housing court. Many tenants may be forced out of their homes without ever appearing in housing court. Some courts also do not report eviction cases. So, strictly speaking, our analysis only efficiently reduces the kind of eviction that will appear in our data. The ethical issue is that we are presuming that reducing the absolute number of evictions city-wide is the goal; it may be more equitable to focus on neighborhoods that historically have been targeted by unscrupulous landlords or perhaps on a collection of neighborhoods evenly distributed across the city.

While these are important caveats, our analysis should still help to further understanding of how legal aid to tenants can be efficiently allocated across the city. This will be important even when more funding is made available, because efficiently allocating the available resources might free up money for other important efforts such as providing legal aid to city residents facing deportation.

1. **Appendix: Descriptive Statistics**

The following tables and figures summarize the Eviction Lab data at the block group level for Philadelphia.

Table 1: Summary of Evictions, Filings, Eviction Rate, and Filing Rate

A screenshot of a cell phone

Description automatically generated

Table 2: Total Evictions, Total Filings, Average Eviction Rate, and Average Filing Rate Over Time

A screenshot of a cell phone

Description automatically generated

Figure 1: Total Evictions and Total Filings Over Time

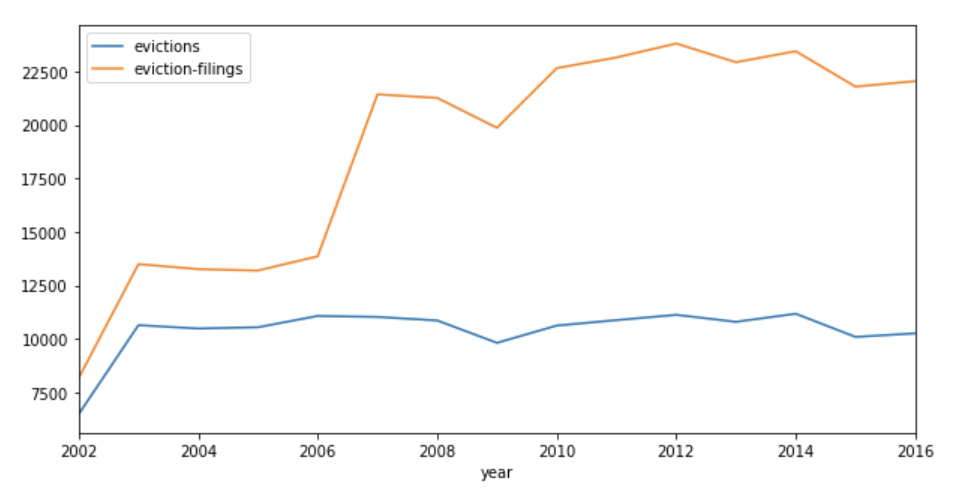


Figure 2: Average Evictions, Filings, Eviction Rate, and Filing Rate Over Time

A close up of a map

Description automatically generated

Figure 3: Comparison of Evictions to Filings and Eviction Rate to Filing Rate

A screenshot of a social media post

Description automatically generated

1. <https://www.phila.gov/hhs/PDF/Mayors%20Task%20Force%20on%20Eviction%20Prevention%20and%20Response-Report.pdf>

   [↑](#footnote-ref-1)
2. <https://www.philadelphiabar.org/WebObjects/PBA.woa/Contents/WebServerResources/CMSResources/PhiladelphiaEvictionsReport.pdf> [↑](#footnote-ref-2)
3. <https://www1.nyc.gov/assets/hra/downloads/pdf/services/civiljustice/OCJ_Annual_Report_2018.pdf> [↑](#footnote-ref-3)
4. <https://www.citylab.com/equity/2018/12/philadelphia-could-be-next-provide-lawyers-low-income-tenants/577519/> [↑](#footnote-ref-4)
5. This research uses data from The Eviction Lab at Princeton University, a project directed by Matthew Desmond and designed by Ashley Gromis, Lavar Edmonds, James Hendrickson, Katie Krywokulski, Lillian Leung, and Adam Porton. The Eviction Lab is funded by the JPB, Gates, and Ford Foundations as well as the Chan Zuckerberg Initiative. More information is found at evictionlab.org. [↑](#footnote-ref-5)