UHERO
THE ECONOMIC RESEARCH ORGANIZATION
AT THE UNIVERSITY OF HAWAI'I

Combinatorial Optimization for Urban Planning: Strategic Demolition of Abandoned Houses in Baltimore, MD

BY

Philip ME Garboden, Lenny Fan, Tamás Budavári, Amitabh Basu, and John David Evans

Combinatorial Optimization for Urban Planning: Strategic Demolition of Abandoned Houses in Baltimore, MD

Garboden, Philip ME[*1], Fan, Lenny[2], Budavári, Tamás[2], Basu, Amitabh[2] and Evans, John David[4]

[1]Department of Urban and Regional Planning, University of Hawai'i Economic Research Organization, University of Hawai'i at Mãnoa
[2]Department of Applied Math and Statistics, Johns Hopkins University
[4]Department of Housing and Community Development, Baltimore City

July 9, 2019

Acknowledgements

[*]pgarbod@hawaii.edu, 2424 Maile Way, Honolulu, HI 96822

Abstract

In 2017, Baltimore City was awarded $75 million dollars earmarked for the targeted demolition of a portion of its 16,000 vacant and abandoned buildings. Selecting an optimal set of demolition targets is difficult given that the cost per demolition is not independent of the overall demolition pattern; like many older cities, Baltimore's abandoned housing includes a large number of attached rowhouses, which require the construction of retaining walls when a demolished home abuts a non-demolished home. In this paper, we present a method by which planners can use integer linear programming to identify optimal demolition targets for a number of potential objectives. The simplest objective, demolishing the maximum number of houses for a specific budget, is compared to more complex functions that attempt to proxy improved quality of life resulting from the demolitions. The results of different objective functions are then assessed in terms of equity and efficiency using the spatial distribution of proposed targets as a point of comparison.

1 Introduction

For decades, American cities' efforts against housing abandonment have been overwhelmed by the scale and complexity of the problem (Mallach, 2006; Sternlieb and Burchell, 1973). The combined forces of suburbanization and deindustrialization have resulted in significant population loss throughout the rust belt; cities such as Chicago, New York, Pittsburgh, St. Louis, Detroit and Baltimore were left with a housing stock that vastly exceeded their populations. While some of these losses were mitigated by a decrease in average household size, the inevitable result was thousands of homes left vacant and abandoned, gradually decaying until they became uninhabitable (Bier and Post, 2003; Cohen, 2001; Wilson and Margulis, 1994).

The size of blight remediation need relative to available resources has often made comprehensive responses impossible. The majority of interventions have thus focused on accurately identifying (Community Research Partners, 2008), cataloging (Ballard and Kingsley, 2007; Brophy and Vey, 2002), banking (Alexander, 2005; Dewar, 2009), and demolishing a subset of a city's vacant and abandoned properties (Mallach, 2006). When city officials have been successful in securing demolition funding, technical and political challenges related to developing a comprehensive process have often resulted in

2

sub-optimal targeting (Accordino and Johnson, 2000; Cohen, 2001).

Baltimore has followed this post-industrial pattern. The city's population dropped from just under 1 million residents in 1950 to 620,000 today (Cohen, 2001). Forty years after Michael Stegman predicted that, unless Baltimore reversed its nascent trend of housing abandonment, efforts "to renew the inner city will be overwhelmed" (Stegman, 1972, p. 66), Baltimore contains over 16,000 vacant and abandoned houses and Stegman's dire warning has largely been realized. Abandoned properties have become the city's most pernicious symbol of blight. They are simultaneously the symptom of economic decay and its cause, locking neighborhoods into a market position for which no incremental remedy is effective.

They cost money to maintain (Goetz, Cooper, Thiele, and Lam, 1998), increase crime (National Vacant Properties Campaign, 2005; Skogan, 1992; Spelman, 1993), encourage arson (Ahrens, 2009), harm the environment (Farfel, Orlova, Lees, Rohde, Ashley, and Chisolm Jr, 2003) and decrease the value of surrounding properties (Han, 2014, 2017; Mikelbank, 2008; Shlay and Whitman, 2006), further reducing the city's tax base. In Baltimore, city officials estimate that just maintaining each vacant property costs the city $500 per year, totaling over eight million dollars.

The city has tried numerous interventions over the years, ranging from large-scale urban renewal to a variety of rehabilitation incentives (Ballard and Kingsley, 2007). While these programs have made a difference in some areas, none have successfully reversed the trend (Mallach, 2017). Recently, in the wake of the upheaval following the death of Freddie Gray, the Maryland Governor's Office committed an unprecedented $75 million dollars to assist Baltimore City in its fight against blight. But because of the expense involved in the demolition of brick row houses, this funding could only remediate a modest portion of the overall problem.

The question that formed the motivation for our research, then, was how to identify the pattern of abandoned house demolition that maximally benefits city residents, within the budget constraint of an allocation such as this. As a team consisting of staff at Baltimore's Department of Housing and Community Development (DHCD) and [University Name Blinded], we sought to answer the issue of target selection using combinatorial optimization.

Our goal was not to replace traditional planning methods, but to develop an analytic technique to serve as decision support to urban policymakers attempting to balance the needs and interests of multiple stakeholders. It would be neither appropriate nor desirable to use our approach instead of any

3

traditional methods; the final selection of demolition targets will necessarily involve preservationist, political, planning, and strategic considerations.

The mathematical complexity involved in this research rests on the fact that the cost of a row-house demolition is not independent of the demolition pattern. Specifically, the cost to demolish a row house depends on the number of adjacent non-demolished properties because of the cost associated with building retaining walls. Thus the determination of an optimal set of demolition targets is non-trivial and must be structured as a combinatorial optimization problem.

In this paper, we employ integer linear programming to identify optimal demolition targets for a number of potential objectives. The simplest objective, demolishing the maximum number of houses, is compared to more complex functions that attempt to proxy improved quality of life resulting from the demolitions, e.g., the number of abandoned properties within the immediate vicinity of occupied homes. The results of different objective functions are then assessed in terms of equity and efficiency using the spatial distribution of proposed targets as a point of comparison. Of course, neither of these objectives represents the full range of issues that demolition planners may be looking to optimize. Many researchers, for example, have used some form of hedonic modeling to estimate the impact of blight on property values (Rosen, 1974) or crime (Skogan, 1992; Spelman, 1993). Where accurate parameter estimates exists, these factors could be incorporated into a optimization procedure customized for a particular locality.

The actual target sets provided in this paper's figures are for illustrative purposes only and intentionally do not represent any actual set of demolition targets considered by the city. This is for two reasons. First, the planned demolition of abandoned properties has real-world implications and we would not wish for any resident of Baltimore to believe they have received news of the city's plans through an academic publication. Second, the target sets most useful to the city involve substantial on-going city-specific model tuning, which would be of little interest to a broader audience.

This paper presents the tool with sufficient flexibility to be useful to urban planners in a variety of contexts, specifically those in cities that struggle with housing abandonment. For the basic modeling, the tool requires very little data beyond building footprints with an abandoned building flag, making it implementable in a variety of contexts.

4

2 Defining An Optimal Demolition Strategy

Our quantitative approach to strategic planning is motivated by optimizing the impact of budget-constrained interventions at both the neighborhood and citywide level. In our framework, every house is considered for demolition and any possible combinations of demolition targets can be evaluated in terms of its impact and cost. Based on these assumptions, we can reject unfeasible plans that are too expensive, and compare competing feasible proposals. To do so, we develop a computationally efficient strategy by which to assess an exponentially large number of demolition target sets.

As noted above, we focus on abandoned houses as the most plausible demolition targets. The definitions of vacant and/or abandoned and the distinction from habitable houses are not always consistent in the literature (Cohen, 2001; Sternlieb, Burchell, and Paulus, 1972). For our purposes, a vacant or abandoned property is one on which the city has placed a Vacant Building Notice (VBN hereafter). Such a property is not legally occupied and is deemed unfit for human habitation. A habitable property is either occupied currently or unoccupied but part of the active housing market, i.e., it is either being renovated, marketed for sale or rent, or being held with no obvious signs of disrepair.

Our goal is to not treat these two habitability states fundamentally differently, especially because the status of each house can change. Instead, we formulate the mathematical model to allow for interventions on any house to make improvements in the whole city. The questions we will address are the following:

1. How can competing planning strategies be represented and compared? Is there a way to articulate an approach to housing demolition as a quantifiable objective?

2. Can we identify the best set of intervention targets such that it optimizes some objective notion of social good, subject to budgetary constraints?

3. To what degree does the optimal set compare to the status quo or a random selection of vacant targets? How does it compare to the naive strategy of maximizing total demolitions?

4. How do the assumptions, both explicit and implicit, in our various objective functions effect the distribution of demolition targets throughout

the city?

2.1 Representation and Formalism

We formulate the above problem in the realm of combinatorial optimization where binary variables assigned to all houses represent the proposed interventions. For example, when planning demolitions, one can introduce a set of $\{x_i\}$ variables for all houses H, ($i \in H$), to determine for every house i if it is recommended for demolition ($x_i = 1$) or not ($x_i = 0$). The objective and the budget can now be expressed as functions of these variables that together will be represented by $\boldsymbol{x}$ for short.

The cost of demolition c_i will be different for each house i depending on a number of factors such as the number of stories or the potential relocation cost of residents, but the total budget required for a given plan will also contain line items that come from a more complicated function of these variables. A typical example is the cost of support walls for row-houses. If a house i is proposed for demolition but the attached property j next door is not, an extra w_{ij} cost is incurred to build a wall for structural integrity. The total budget can be written as a sum of two terms

$$B(\boldsymbol{x}) = \sum_{i \in H} c_i \, x_i \; + \sum_{(i,j) \in E} w_{ij} \, (x_i \oplus x_j) \tag{1}$$

where E represents the set of all neighbors in row houses. The first term simply means that the cost for house i is only incurred if the intervention is planned, $x_i = 1$, which one just has to add up for all houses. The second term, however, needs more explanation. The $\oplus$ operator represents the Boolean XOR, which yields 1 if only one of the two arguments is 1, and 0 otherwise. The meaning of these terms is that the w_{ij} cost is only incurred if only one of the neighboring houses (sharing a wall) is demolished. If both are kept or both are demolished, no extra cost is involved. This non-linear aspect makes the optimization problem much more complicated. With the budget $B(\boldsymbol{x})$, we can formulate the constraint that each of our solutions will have to obey, which is that the total cost is less than a given threshold B_T, i.e.,

$$B(\boldsymbol{x}) < B_T \,. \tag{2}$$

In this paper, we will employ several objective functions that are relevant for urban planning. We will introduce increasingly complex and more meaningful measures and study their solutions. The simplest and perhaps most

intuitive objective is as follows: Given a budget, which are the houses one should target to maximize the number of demolitions. This can be written as

$$F_H(\boldsymbol{x}) = \sum_{i \in H} x_i \qquad \text{or} \qquad F_V(\boldsymbol{x}) = \sum_{i \in V} x_i \tag{3}$$

if focusing on only the vacant subset $V \subset H$. The above combinatorial optimization is non-trivial not only because of the hundreds of thousands of variables but due to the non-linear and combinatorial constraints.

2.2   Integer Linear Programming

To solve the problem at hand, we transform the above budget function that simplify the equations at the expense of the introduction of extra variables and constraints. Let us substitute z_{ij} variables in place of the non-linear $x_i \oplus x_j$ terms. To achieve the same effect we also have to require the following constraints

$$
\begin{align}
z_{ij} &\geq x_i - x_j \tag{4}\\
z_{ij} &\geq x_j - x_i \tag{5}\\
z_{ij} &\leq x_i + x_j \tag{6}\\
z_{ij} &\leq 2 - x_i - x_j \tag{7}
\end{align}
$$

which can be verified by testing the possible variable combinations. Considering that the above are linear inequalities, our original optimization problem of an F objective function

$$\max_{\boldsymbol{x}} F(\boldsymbol{x}) \quad \text{such that} \quad B(\boldsymbol{x}) < B_T \tag{8}$$

is reduced to a linear problem over $\boldsymbol{x}$ and the new dependent $\boldsymbol{z}$ variables, which now can be efficiently solved using methods of Integer Linear Programming (hereafter ILP). In addition to the above simple objectives in Equation (3) that illustrate the concept, we will also study variants of a function that intends to capture the sentiment that home owners prefer to live far away from vacant properties (Han, 2017). In the next sections we describe these models.

2.3 Modeling a more realistic objective

Based on the assumption that each vacant house v generates negative externalities born by each occupied house o within a particular radius, we formate an "impact score" $h_{ov} = f(o, v)$ for each mixed pair in the city. There are many options for the concrete formula by which f is calculated. A sophisticated approach could, in theory, consider myriad externalities such as crime, population density, traffic, property values, and so forth (Han, 2014, 2017). But to illustrate our approach, we employ simpler functions related to the spatial distance of each (o, v) pair, making only the basic assumption that externalities decrease at greater distances.

Our first approach is to let the impact score be proportional to $(1/d)^p$ where d is distance and p represents the rate of decay. The impact score decays faster with large p. However, the impact approaches infinity when two houses are very close to each other, meaning that the model largely ignores the impact on non-adjacent parcels. To resolve this characteristic issue, we use the function below to normalize the impact. d_0 is a parameter representing an "distance." We now consider the following formula for

$$h_{ij} = e^{-(d_{ij}/d_0)^p} \, . \tag{9}$$

For example, when p is sufficiently large, the impact score is approximately equal to 0 if the distance is greater than d_0.

In terms of our notation from Section 2, we can now formulate the new objective. Let $V \subset H$ be the vacant set of houses and $O = H \setminus V$ denote the set of occupied houses. The objective would look like

$$F_I(\boldsymbol{x}) = \sum_{i \in O} \sum_{j \in V} h_{ij} (1 - x_i)(1 - x_j) \tag{10}$$

which needs to be minimized.

While Equation (9) is suitable for estimation, it assumes that the impact of an abandoned property declines with Euclidean distance. This assumption is not necessarily the case; an abandoned property "on the same block" as an occupied property may have the same impact regardless of whether that "block" is 500 or 1,000 feet long. Indeed, as we will see in Section 4, the use of impact functions based on Euclidean distances naturally prefers demolition targets in high-density parts of the city.

As an alternative, we formulate impact functions that consider the impact on each occupied property based on its k-nearest neighbors for an appropriate

choice of k. This would allow to account for difference in density in different parts of the city. More precisely, given a value of k, for each occupied house, only the nearest k vacant houses will have impact on it. The value of k was selected by calculating the average number of row houses for a block and increasing that value exponentially to generate a set of results. This assumes that each of the k-nearest vacant homes will have the same impact on each occupied house. However, some areas of Baltimore are lower density. To address this, we proposed to set d_0 equal to the average/median distance

$$d_0 = \frac{1}{k\,|O|} \sum_{i \in O} \sum_{j \in V_i} d_{ij} \qquad \text{or} \qquad d_0 = \underset{i \in O,\ j \in V_i}{\text{median}} \big\{ d_{ij} \big\} \qquad (11)$$

where $V_i \subseteq V$ is the subset of k-nearest vacant houses to the occupied house $i \in O$.

3   Sources of Data

The project uses data from multiple administrative data sources supplied by the Baltimore City Department of Housing and Community Development linked by the city's "block lot" identifier.

Parcel & Building Polygons In order to determine property adjacency and pairwise distances, we utilized building polygons maintained by Baltimore City. Corrections to this data were made to account for recent property demolitions.

Abandoned Properties Baltimore City tracks abandoned properties through its Code Enforcement Division, which places Vacant Building Notices (VBNs) on all un-occupied properties deemed not suitable for habitation, generally due to broken windows and doors (although many have fallen into far worse condition over time). Data used for this project consists of all VBN notices in the city active in January 2018.

Census Data Neighborhood characteristics are operationalized at the census tract level and use data from the 2015 American Community Survey, a probability sample of all households in the United States estimated in five-year ranges (the 2015 ACS is thus the average across 2013-2017).

4   Results

In this section, we report the answers to our research questions. First we report the computational efficiency of the optimization results for the city of Baltimore, which includes 176,474 properties and 15,726 vacant homes.

We implement with different combinations of budget B_T, the characteristic distance d_0, and the shape parameter of the impact curve p. We employ effective distances ranging from 10 to 700 meters and budgets ranging from 1 to 40 million dollars with different impact shapes $p = 1, 2$, and ∞. In addition, we compare the result to the model only considers the top k nearest houses ($k \in \{25, 100, 225\}$) and the exponent $p = 1$.

For each model, we consider how the optimal objective value compares to the same value for a model demolishing the maximum number of properties, a model that minimizes the objective value while holding the number of demolitions constant, and a set of models that select demolition targets randomly (while remaining under the budget constraint).

Finally, we examine how these various models compare in terms of city-wide equity, comparing how the demolition targets are distributed in terms of neighborhood concentration and neighborhood characteristics.

4.1   Computational Efficiency

The general problem is NP-complete (proof available upon request) and therefore very hard in the worst case, from a theoretical complexity perspective. However, we find that the models in our experiments are always solved in a reasonable amount of time. We used 3.1 GHz Intel Core i7 processor with 16GB of RAM and 4 threads. The experiments completed under 5 minutes in average. In the worst case, a model took up to 20 minutes; on the other hand, a few models were solved under 30 seconds.

4.2   Descriptive Results

In order to minimize the non-trvial objective functions, an ideal demolition target is one that is 1) inexpensive to tear down and 2) close to a large number of habitable properties (with "close to" being defined in multiple ways). This encourages the selection of demolition targets which are contiguous to other demolition targets (thus avoiding the costs associated with building retaining walls), while prioritizing areas with lower levels of property abandonment.

As shown in Figure 1, the optimization model rarely selects demolition targets with occupied properties on either side. Instead, it priorizes whole sections of blockfaces with contiguous abandonment, greatly reducing the cost per demolition.

The models also appear to succeed in minimizing the distance from habitable structures to the demolition targets. As show in Figure 2, the demolition targets are not distributed, in a spatial sense, similar to all vacant structures in the city (pink on the map). Instead, the area where few habitable houses remain (such as the area due west of center city) have disproportionately few demolition targets (red) compared to areas with high levels of abandonment that are spatially proximate to healthier neighborhoods.

4.2.1 Relative Improvement Compared to Random Demolition Targets and Worst Case Selection

The true comparison for our results in not the status quo of zero demolitions, but alternative targeting approaches that might more accurately reflect the planning process in the absence of our tool.

Figure 3 illustrates these comparisons. For two models (other models produce similar patterns, plots available upon request), we estimate the relative change of the objective depending on the number of houses selected for demolition, which we allow to vary between the optimal number of demolitions and the maximum number of demolition possible under the budget constraint. We also consider the minimum improvement in the objective throughout this range.

As clearly indicated by the orange lines, it is possible have almost no relative impact if an inappropriate set of targets is selected. We also consider the impact if the demolition targets are selected randomly (represented by the blue dots), with again the suggestion that our model dramatically improves on random targeting as well.

4.3 Distribution of the Targets

We note that the while we allowed for our models to select inhabited properties for demolition — a politically and ethically complex intervention — no non-vacant property was ever selected for demolition by any of our models. We expect that more advanced models that realize the benefits of site assemblage will propose to utilize such interventions, where entire blocks could be

11

turned into valuable urban assets, such as parks.

In addition to the question of objective maximization, we can evaluate each model in terms of the distribution of its demolition targets across the city (which can in turn impact the project's political legitimacy).

While distributional equity could be incorporated as a tertiary objective in the optimization model, the goal here is to see how various alternative models "score" on this dynamic, allowing policymakers to anticipate the consequences of their modeling decisions.

As shown in Figure 4, none of the models select a set of demolition targets with a distribution identical to the overall distribution of vacant properties in the city. Not only is the overall distribution more concentrated within high poverty neighborhoods (with poverty rates around 30 percent), but the distribution is somewhat bimodal, with a substantial number of vacant properties in extremely high poverty neighborhoods (around 25 percent poverty). Because all the models endeavor to select target in areas adjacent to healthier housing markets, all models produce target sets more evenly distributed throughout the city.

The differences between models that choose either Euclidean or k-nearest buffers suggest that the choice to correct for changes in neighborhood density has significant impacts on the distribution of demolition targets. By allowing the model to incorporate less dense portions of the city, it not surprisingly selects targets that are in substantially lower poverty neighborhoods, while remaining roughly similar with regards to the median.

5   Discussion

Our results show that not only is the proposed optimization problem solvable, but it is sufficiently efficient computationally as to be a practical tool for urban planners. The sets of demolition targets selected by the models follow expected patterns: to reduce costs, the process selects targets that are adjacent to one another and thus can be demolished without the construction of retaining walls. At the large scale, the model tends to privilege areas of the city with a mix of abandoned and occupied housing, ignoring the trivial number of vacant properties in low-poverty neighborhoods as well as the largely unoccupied blocks of Baltimore's most blighted communities. In all cases, the process significantly outperforms a random selection of targets.

Tuning decisions related to the objective function can have substantial im-

pact of the sets of demolition targets. The most substantial differences were found between approaches that modeled distance in different ways, specifically those that considered all properties within a particular buffer radius and those that considered the k-nearest properties. Demolition target sets for the former tended to ignore portions of the city that consisted only of detached single family homes (as compared to row- and town-homes). The former distributed targets throughout the city and thus incorporated lower poverty areas. While beyond the scope of this paper, the decision between those two metrics rests ultimately on an empirical question regarding which properties are impacted by blight externalities. Other changes in the impact function such as the rate of distance decay and the maximum buffer size appeared to have less of an impact on the results. While each objective function described above produced subtly different demolition sets, they were distributed similarly across neighborhoods within the Euclidean and k-nearest groups.

6 Conclusions

This paper presents a novel method for blight planning that uses integer linear programming produce a quantifiably optimal set of demolition targets to achieve a particular objective. We presented two plausible objectives: 1) the demolition of the maximum number of vacant and abandoned properties; 2) the maximal reduction in exposure to abandoned properties by city residents. For this second objective, we have considered a number of ways this exposure can be measured, adjusting the rate of decay and the distance metric. We then compared the set of demolition targets generated by optimization based on each of these objectives in terms of their distribution through Baltimore city's neighborhoods.

The process described in this article neither should nor could replace standard blight planning processes. The decision to demolish a house, no matter how dilapidated, must be made with myriad practical, social, and political considerations. Planning professionals, political representatives, and citizens themselves must all be brought to the table to determine how best to utilize a demolition budget in a way that is both fair and effective. Nevertheless, a tool that is able to efficiently consider the entire city and highlight demolition targets that would maximize particular public goals can prove invaluable to the planning process.

For example, simply knowing the maximum number of potential demolitions would allow planners to benchmark their set of identified targets in terms of budgetary efficiency. Moreover, the combinatorial optimization process makes it possible for the team to identify specific clusters of vacant properties that are inexpensive to demolish (per structure) and would likely benefit legacy residents in the area. These objectively identified targets, then, could be subsequently evaluated by neighborhood planners and historical preservationists partnering with the city.

Our goal in this project was not to develop a single-use tool for a particular city but to establish a model decision support tool that can be replicated and implemented by planners in cities across the country. The data needed to run this analysis consist of little more than accurate building footprint polygons, a list of vacant and abandoned properties, and a set of rough budget estimates for demolition. Given the increasing availability of such data, we believe the tool described herein could be used to assist planners with blight remediation across the nation.

6.0.1 Future Work

As an extension of this research, we are working to move beyond the identification of specific demolition sets – a level of specificity that masks the uncertainty related to myriad tuning decisions. When multiple models appear theoretically valid, it is possible to create a score for each abandoned property based on the number of models which recommend that property for demolition. Similarly, the absolutely optimal target set may not be practically different from the top 5 percent of target sets for a particular objective function. The goal of the probabilistic approach, would be to help urban planners see which properties represent "low hanging fruit" insofar as they appear in targets sets produced by multiple models.

In addition, neither of our exemplary objective functions reflects the myriad of factors considered during blight remediation planning. We have considered, for example, only the distance between occupied homes and abandoned buildings, ignoring the fact that residents would also wish to shop and attend school in areas free of blight. Nor have we incorporated any consideration of which abandoned buildings would likely be renovated without government intervention and which are harbingers of long-term neighborhood decline.

The good news is that each of these considerations, and dozens of others, could be incorporated into the objective function with only minimal changes

14

to process presented here. Many researchers, for example, have used some form of hedonic modeling to estimate the impact of blight on property values (Rosen, 1974). These models, which incorporate dozens of variables and allow for numerous non-linarites of effects, could be used to select a set of demolition targets that would maximize the total assessed value of properties throughout the city targeting neighborhoods in which blight has the most deleterious effect on property values. Similarly, econometric models have shown that blight is associated with crime, generally by providing a largely un-surveilled location in which to participate in illegal activities (Skogan, 1992; Spelman, 1993). These models too could be incorporated into the optimization procedure and a set of targets could be constructed that would maximally reduce crime.

It is likely that these two objectives would produce very different demolition sets: a focus on property values would push targets into areas with middling housing markets, where a handful of abandoned properties can have large negative effects on value. A focus on crime would likely push targets to higher poverty areas, where open-air drug dealing and vagrancy are more common. The decision to prioritize one objective over another cannot be determined objectively, but emerges from a political process. Our tool is designed to assist that process once objectives have been identified, resulting in smarter more equitable decisions.

References

John Accordino and Gary T. Johnson. Addressing the vacant and abandoned property problem. 22(3):301–315, 2000. ISSN 1467-9906. doi: 10.1111/0735-2166.00058. URL http://onlinelibrary.wiley.com/doi/10.1111/0735-2166.00058/abstract.

Marty Ahrens. Vacant building fires, 2009.

Frank Alexander. Land bank authorities: A guide for the creation and operation of local land banks, 2005.

William Ballard and G. Thomas Kingsley. Systems to improve the management of city-owned land in baltimore, 2007.

Thomas Bier and Charlie Post. Vacating the city: An analysis of new homes vs. household growth, 2003.

Paul C. Brophy and Jennifer S. Vey. Seizing city assets: Ten steps to urban land reform, 2002.

James R. Cohen. Abandoned housing: Exploring lessons from baltimore. 12(3), 2001.

Community Research Partners. $60 million and counting: The cost of vacant and abandoned properties to eight ohio cities, 2008.

Margaret Dewar. What helps or hinders nonprofit developers in reusing vacant, abandoned, and contaminated property? findings from detroit and cleveland. 2009.

Mark R Farfel, Anna O Orlova, Peter SJ Lees, Charles Rohde, Peter J Ashley, and J Julian Chisolm Jr. A study of urban housing demolitions as sources of lead in ambient dust: demolition practices and exterior dust fall. *Environmental Health Perspectives*, 111(9): 1228, 2003.

Edward G Goetz, Kristin Cooper, Bret Thiele, and Hin Kin Lam. Pay now or pay more later: st. paul's experience in rehabilitating vacant housing. *CURA Reporter*, 14, 1998.

Hye-Sung Han. The impact of abandoned properties on nearby property values. *Housing Policy Debate*, 24(2):311–334, 2014.

Hye-Sung Han. Exploring threshold effects in the impact of housing abandonment on nearby property values. *Urban Affairs Review*, page 1078087417720303, 2017.

Alan Mallach. *Bringing buildings back: From abandoned properties to community assets: A guidebook for policymakers and practitioners*. Rutgers University Press, 2006.

Alan Mallach. Tackling the challenge of blight in baltimore: An evaluation of baltimore's vacants to value program, 2017.

Brian A. Mikelbank. Spatial analysis of the impact of vacant, abandoned and foreclosed properties, 2008.

National Vacant Properties Campaign. Vacant properties: The true costs to communities, 2005.

Sherwin Rosen. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1):34–55, 1974.

Anne B Shlay and Gordon Whitman. Research for democracy: Linking community organizing and research to leverage blight policy. 5(2):153–171, 2006. ISSN 1540-6040. doi: 10.1111/j.1540-6040.2006.00167.x.

Wesley G Skogan. *Disorder and decline: Crime and the spiral of decay in American neighborhoods.* Univ of California Press, 1992.

William Spelman. Abandoned buildings: Magnets for crime? *Journal of Criminal Justice*, 21(5):481–495, 1993.

Michael A Stegman. *Housing investment in the inner city: the dynamics of decline; a study of Baltimore, Maryland, 1968-1970.* The MIT Press, 1972.

George Sternlieb and Robert W Burchell. *Residental abandonment: The tenement landlord revisited.* 1973.

George Sternlieb, Robert W Burchell, and Virginia Paulus. Residential abandonment: The environment of decay, 1972.

David Wilson and Harry Margulis. Spatial aspects of housing abandonment in the 1990s: The cleveland experience. 9(4):493, 1994. ISSN 02673037.

Figure 1: Small-scale illustration of Demolition Targets (red), Vacant Houses (pink), and Habitable buildings (grey) for model ($20mil budget, 125 buffer, p=inf).

Figure 2: Large-scale illustration of Demolition Targets (red), Vacant Houses (pink), and Habitable Properties (grey) for model ($20mil budget, 125 buffer, p=inf).
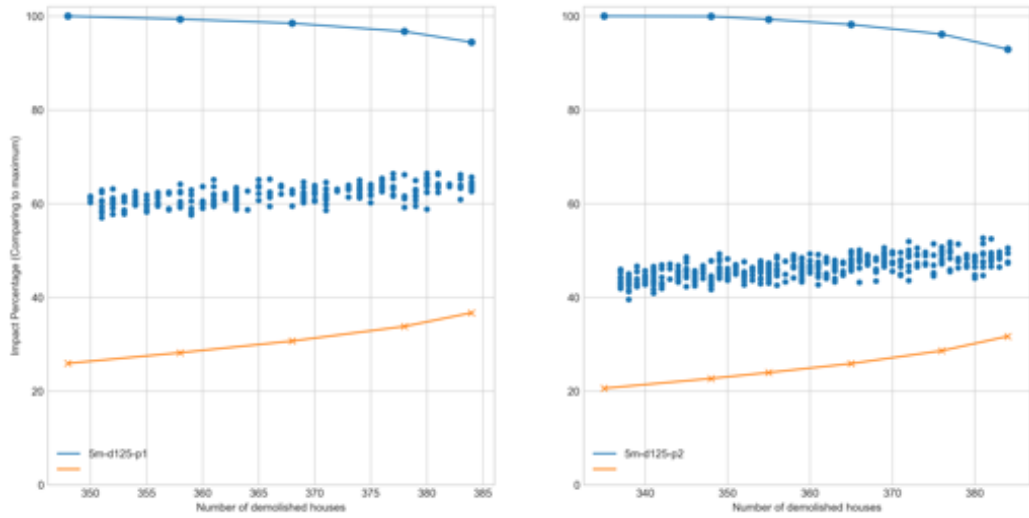
Figure 3: Relative Improvement of Objective Function for Optimal Solution, Random Solutions, and Minimum Solution with 5 million budget.
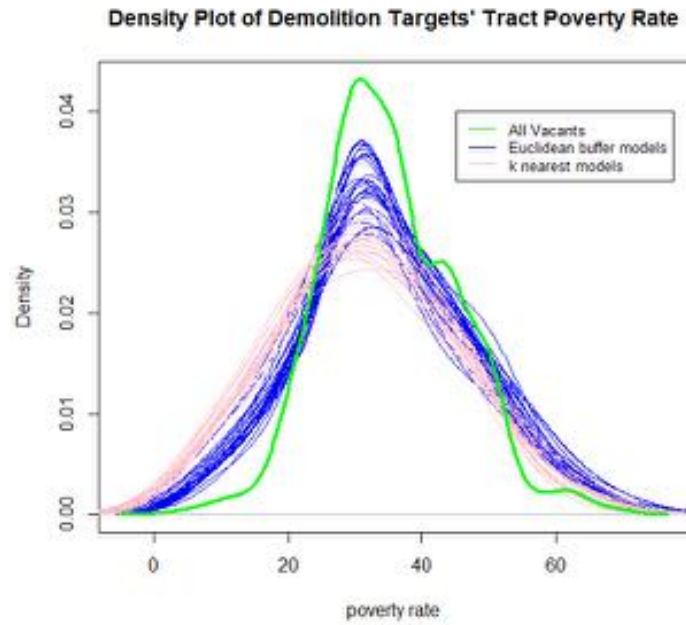
Figure 4: Curves represent distribution of demolition targets in terms of neighborhood poverty rates. Blue curves are from euclidean models, pink from k-nearest models, and green represents the distribution of all abandoned properties in the city.

A   Technical Appendix

A.1   Notation

We use the following notation for the rest of this Appendix.

$$
\begin{aligned}
B_T &= \text{ The total budget} \\
H &= \text{ The collection of houses ID} \\
O &= \text{ Subset of H, the collection of occupied houses ID} \\
V &= \text{ Subset of H, the collection of vacant houses ID} \\
E &= \text{ The collection of adjacent houses pairs} \\
C &= \{(i,j) \mid i \in O \text{ and } j \in V\}
\end{aligned}
\tag{12}
$$

A.2   The integer programming model for maximizing the numbers of houses to be demolished

The original problem is to tear down as many vacant houses as possible. We can formulate this as an integer programming model. For each vacant house, we assign a binary variable x_i, for every vacant house $i \in V$, such that

$$
x_i = \begin{cases} 1 \text{ , if the house i is demolished} \\ 0 \text{ , otherwise} \end{cases}
$$

and our objective function is

$$
\max \sum_{i \in V} x_i
$$

The budget constraint can be written as

$$
\underbrace{\sum_{i \in V} c_i x_i}_{\text{cost of demolishing the houses}} + \underbrace{\sum_{(i,j) \in E} c'_i x_j (1 - x_i)}_{\text{cost of adjacent wall for house i}} +
$$

$$
\underbrace{\sum_{(i,j) \in E} c'_j x_i (1 - x_j)}_{\text{cost of wall for house j}} + \underbrace{\sum_{(i,j) \in E} p_{ij} x_i x_j}_{\text{time/cost benefit for demolishing houses i and j}} \leq B_T,
\tag{13}
$$

where c_i is the cost of demolishing the the vacant house i, c_i' is the cost to build the wall for house i if an adjacent house j (i.e., $(i,j) \in E$) is demolished and p_{ij} represents the scale economies related to demolishing adjacent houses. To make the constraint linear, we replaced all product terms $x_i x_j$ with y_{ij}, and imposed the following additional constraints:

$$
\begin{aligned}
x_i + x_j - y_{ij} &\leq 1 \\
-x_i + y_{ij} &\leq 0 \\
-x_j + y_{ij} &\leq 0
\end{aligned}
\tag{14}
$$

$$
y_{ij} = \{0, 1\}
$$

These constraints encode $y_{ij} = x_i x_j$. Therefore, we can convert the nonlinear constraint (13) into the linear constraint

$$
\sum_{i \in V} c_i x_i + \sum_{(i,j) \in E} (c_i' x_j + c_j' x_i) - \sum_{(i,j) \in E} (p_{ij} + c_i' + c_j') y_{ij} \leq B_T
\tag{15}
$$

A.3 The integer linear programming model for minimizing total impact

Let h_{ij} be a function that measures the impact score between the house i and house j, where $i \in O$ and $j \in V$; see the discussion in Section 2.3. The vacant house j has impact h_{ij} on the occupied house i if and only if $x_j = x_i = 0$, which can be written as

$$
h_{ij}(1 - x_i)(1 - x_j) = h_{ij}(1 - x_i - x_j + y_{ij})
\tag{16}
$$

Therefore, this objective function can be modeled by the following integer program.

$$\min \quad \sum_{i,j \in C} h_{i,j}(1 - x_i - x_j + y_{ij})$$

$$\text{s.t.} \quad \sum_{i \in H} c_i x_i + \sum_{(i,j) \in E}(c'_i x_j + c'_j x_i) - \sum_{(i,j) \in E}(c'_i + c'_j + p_{ij})y_{ij} \le B_T$$

$$\forall (i,j) \in E \uplus C$$
$$x_i + x_j - y_{ij} \le 1$$
$$-x_i + y_{ij} \quad \le 0$$
$$-x_j + y_{ij} \quad \le 0$$

$$x_i \in \{0,1\} \quad \forall i \in H$$
$$y_{ij} \in \{0,1\} \quad \forall (i,j) \in E \uplus C \tag{17}$$

A.3.1 Formulating the problem using Big M

In the model (17), we have variables y_{ij} for every $(i,j) \in E \uplus C$. The cardinality of C is $|V| \times |O|$ which might be very large. In this section we will introduce Big M method to reduce the dimension.

$$
\begin{aligned}
\min \quad & \sum_{i \in O} \sum_{j \in V} h_{ij}(1 - x_i)(1 - x_j) \\
= \quad -\max \quad & -\sum_{i \in O} \sum_{j \in V} h_{ij}(1 - x_i)(1 - x_j) \\
= \quad -\max \quad & \sum_{i \in O} \sum_{j \in V} h_{ij}(1 - x_i)(x_j - 1)
\end{aligned}
\tag{18}
$$

For any occupied house i, depending on the two cases whether $x_i = 0$ (house is not demolished) and $x_i \neq 0$ (house is demolished), the total impact on the occupied house i can be formulated as

$$
\text{The Impact Score} = \begin{cases} \sum_{j \in V} h_{ij}(1 - x_j) & \text{if } x_i = 1 \\[2ex] 0 & \text{if } x_i = 0 \end{cases}
$$

Using these observations, we can replace the $y_{ij} \in C$ variables introduced earlier with new variables t_i by using so-called Big-M method. In particular, we propose to impose the following constraints on these variables t_i:

24

$$\forall\, i \in O \quad \begin{cases} t_i \leq \quad \sum_{j \in V} h_{ij}(x_j - 1) + Mx_i \\[2ex] t_i \leq \quad 0 \end{cases} \tag{19}$$

where M is a sufficiently large number. In the inequalities (19), the first inequality will always be nonpositive when $x_i = 0$, and thus t_i is constrained by the first inequality. When $x_i = 1$, there is no impact on the occupied house i. By adding a sufficient large value M, the upper bound of t_i is controlled by the second inequality. The final integer linear programming model will be the following.

$$\max \quad \sum_{i \in O \uplus R} t_i$$

$$\text{s.t.} \quad \sum_{i \in H} c_i x_i + \sum_{(i,j) \in E}(c'_i x_j + c'_j x_i) - \sum_{(i,j) \in E}(c'_i + c'_j + p_{ij}) y_{ij} \leq B_T$$

$$\begin{aligned} &\forall (i,j) \in E \\ &\quad x_i + x_j - y_{ij} \leq 1 \\ &\quad -x_i + y_{ij} \quad \leq 0 \\ &\quad -x_j + y_{ij} \quad \leq 0 \\ &\forall\, i \in O \\ &\quad t_i \leq \sum_{j \in V} h_{i,j}(x_j - 1) + Mx_i \end{aligned}$$

$$\begin{aligned} &x_i \in \{0,1\} \quad \forall i \in H \\ &y_{ij} \in \{0,1\} \quad \forall (i,j) \in E \\ &t_i \ \in \{-\infty, 0\} \quad \forall i \in O \end{aligned}$$

(20)

As discussed in (16), the final impact score will be the negative of the optimal value of the above integer linear optimization problem.

A.4   NP-hardness of minimizing total impact

It is clear that the problem is in NP, since given a particular set of houses to be demolished, one can check if they satisfy the budget constraint and whether the impact is above a certain threshold in time that is linear in the number of houses. To show the problem is NP-hard, we reduce the well-known Knapsack Problem to our problem by following steps.

1. Let the knapsack instance be as follows: $v_1, \ldots, v_n$ are the values/profits of the different items, let $w_1, \ldots, w_n$ be the weights of different items, and W be the total capacity of the knapsack. Thus, we want to select a subset of items to maximize the total value, while maintaining total weight below the capacity of the knapsack.

2. For all $i \in \{1, \ldots n\}$, create a vacant house i such that the cost to demolish the vacant house i is w_i.

3. For all $i \in \{1, \ldots n\}$, create an occupied house i' such that the cost to demolish the occupied house i' is some constant $C > W$ (Note that there are no adjacent houses in this setting)

4. For each pair of (i, i'), set the impact between them to be equal to $-v_i$ which is the impact of vacant house i to occupied houses i'.

5. Set the budget to be W .

Note that we can only demolish vacant houses in this instance. The optimal set of vacant houses to be demolished to maximize impact will correspond to an optimal solution to the knapsack problem.