Journal Title
XX(X):1–18

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Community Vibrancy and its Relationship with Safety in Philadelphia

Wichinpong Park Sinchaisri[1] and Shane T. Jensen[2]

Abstract

To what extent can the strength of a local urban community impact neighborhood safety? We constructed measures of community vibrancy based on a unique dataset of block party permit approvals from the City of Philadelphia. We use both regression modeling and propensity score matching to control for the economic, demographic and land use characteristics of the surrounding neighborhood when examining the relationship between crime and our measures of community vibrancy. We conduct our analysis on aggregate levels of crime and community vibrancy from 2006 to 2015 as well as the trends in community vibrancy and crime over this time period. We find that neighborhoods with a higher number of block parties have a significantly higher crime rate, while those holding a greater proportion of spontaneous community-focused events have a significantly lower crime rate. We also find that neighborhoods with an increase in the ratio of spontaneous block parties over time are significantly more likely to have a decreasing trend in total crime incidence over that same time period.

Keywords

urban analytics, community vibrancy, block parties, crime rates, matched pairs experiment

1 Introduction

Why does the crime rate vary so strikingly between neighborhoods in large cities? Common factors associated with high crime rates include poverty levels, job availability, police policy, and the average age of the population. Shaw and McKay (1942) proposes a theory of community social disorganization:

[1]Department of Operations, Information, and Decisions, The Wharton School, University of Pennsylvania
[2]Department of Statistics, The Wharton School, University of Pennsylvania

Corresponding author:
Shane T. Jensen, Department of Statistics The Wharton School, University of Pennsylvania, 463 Huntsman Hall 3730 Walnut Street, Philadelphia, PA, USA 19102
Email: stjensen@wharton.upenn.edu

low economic status, ethnic heterogeneity, residential mobility, and family disruption lead to community social disorganization, which, in turn, increases crime and delinquency rates. Since then, this model has been tested empirically by several researchers who identified potential factors. Crutchfield et al. (1982) finds that high rates of mobility negatively affects social integration, lowering the effectiveness of community informal control mechanisms. Sampson and Groves (1989) shows that between-community variations in social disorganization significantly affects crime rates. Bellair (1997) explores the consequences of frequent and infrequent interaction among neighbors and finds that type of interaction matters. Specifically, getting together once a year or more with neighbors has the most consistent and generally strongest effect on burglary, motor vehicle theft, and robbery.

These findings seem to suggest that the nature of social interaction within a neighborhood is correlated with local safety. More recently, Humphrey et al. (2017) constructed measures of urban vibrancy based upon the built environment of a neighborhood in order to investigate the relationship between safety and the built environment in Philadelphia. The authors observe that neighborhoods with more vacancy have higher crime, but within neighborhoods, more crimes occur near business locations, and that businesses that are active (open) for longer periods are associated with fewer crimes.

In this paper, rather than investigating the role of the built environment and crime of a neighborhood, we focus on the community vibrancy of a neighborhood. We create quantitative measures of *community vibrancy* in local areas of the city of Philadelphia using a unique dataset of block party permit approvals from 2006 to 2015. Since 75% of a street's residents need to agree to hold a block party, this data provides an interesting measure of the cohesion of the community on a particular block.

We construct two quantitative measures of community vibrancy from this data: the number of block party events and the ratio of *spontaneous* to *non-spontaneous* block party events in a neighborhood. Spontaneous events are those organized by a local community on a non-regular basis, while non-spontaneous are related to "public" holidays (e.g., Mother's Day) or federal (e.g., July 4th, Memorial Day) or religious holidays (e.g., Christmas). We then investigate whether there is an association between these measures of community vibrancy and crime incidence at the neighborhood level in Philadelphia. We also examine the relationship between changes in community vibrancy over time and trends in crime over time.

However, these relationships are potentially confounded by many other neighborhood factors that are also related to either our created measures of community vibrancy or crime incidence. To address this possibility, we incorporate data on the economic, demographic, and land use characteristics of Philadelphia neighborhoods into our analyses. We use two statistical techniques, regression modeling and propensity score matching, to estimate the association between crime and community vibrancy while controlling for these other neighborhood factors.

Our results demonstrate that neighborhoods with a larger number of block parties per year have a significantly higher annual crime rate. However, those holding a greater proportion of spontaneous block party events have a significantly lower crime rate. We also find that neighborhoods with an increase in the ratio of spontaneous block parties over time, potentially indicating a stronger sense of community, are significantly more likely to have a decreasing trend in total crime incidence over that same time period.

This paper is organized as follows. We use Philadelphia block party permit data to define measures of community vibrancy in Section 2. We explore the relationship between our community vibrancy measures and other neighborhood characteristics (economic, demographic, and land use) in Section 3. We examine crime incidence at the neighborhood level in Philadelphia in Section 4 and then investigate the association between overall crime incidence and our community vibrancy measures in Section 5. We also investigate the relationship between changes over time in crime incidence and community vibrancy in Section 6. We summarize our findings in Section 7.

2 Measuring Community Vibrancy through Block Parties

2.1 Data: Block Parties in Philadelphia

Our dataset contains 68,553 permit approvals for a block party across 10,347 unique locations (by street address) in the city of Philadelphia from January 2006 to May 2016. This data was made available to us by the author of Geeting (2016). All permits in this data are for one-day events, although we do observe that some blocks organize events on consecutive days. Since we do not observe the full details of the event nor its planner, we consider events on consecutive days as separate events.

In this paper, we study community vibrancy at the neighborhood level of resolution. We will define our neighborhood units as the "block group" geographical units established by the US Census Bureau. There are 1,336 US Census block groups in the city of Philadelphia. These US census block groups consist of 10-20 city blocks which generally matches our concept of a "neighborhood," and the block-group level is the highest resolution at which the US Census Bureau publicly releases economic data. We aggregate the 68,553 block party permits within these 1,336 neighborhoods in Philadelphia.

There are 30 unique event types for these block party permits, which we group into two main categories: regular events such as national or religious holidays versus *spontaneous* community events that are not tied to a regular holiday. The breakdown of the event types within these two categories is:

- **Regular events** (7.45%)
  - *Public holiday*: 4th of July, Labor Day, Memorial Day, New Year's Day, New Year's Eve, May Day, Christmas Party, Father's Day, Mother's Day (7.30%)
  - *Religious*: Church Service, Communion, Easter Egg Hunt, Halloween Party, Other Religious Event (0.15%)
- **Spontaneous events** (92.54%)
  - *Community*: Community Fun Day, National Night Out, Prom, Spring Festival, Arts & Crafts Show, Health Fair, Stop The Violence Crusade, Dedication, Serenade (92.13%)
  - *Personal*: Baby Shower, Birthday Party, Graduation Party, Repass, Wedding Reception, Wedding (0.41%)

One question we will consider in this paper is whether this distinction between regular versus spontaneous events is helpful when constructing measures of community vibrancy for a neighborhood. In the subsections below, we propose our two primary measures of vibrancy that reflect either regular or spontaneous events or both.

2.2 Community Measure 1: Total Number of Block Party Events

We first consider the total number of block party events (regular or spontaneous) held within each neighborhood. The total number of block party events held in a particular neighborhood is a simple and intuitive measure of the community vibrancy of that neighborhood. In Figure 1 (left), we show the total number of block party events within each neighborhood of Philadelphia, aggregated across the entire time span of our data (2006-2016).

We see in Figure 1 (left) that neighborhoods that have the largest total number of block party events are in the North Philadelphia area. West Philadelphia and South Philadelphia also have several neighborhoods with a large total number of block party events, whereas the outlying suburban communities in the Northwest and Northeast parts of the city have relatively few total number of block party events. We will examine the trend over time in the total number of block party events aggregated by year in Section 2.4 below.

2.3 Community Measure 2: Spontaneous Ratio

In addition to the total number of block party events held in each neighborhood, we are also interested in the distinction between spontaneous versus regular block party events, as outlined in Section 2.1 above. For each neighborhood in Philadelphia, we compute the ratio of the number of spontaneous events to the total number of events (spontaneous or regular).

This *spontaneous ratio* is generally quite high since over 90% of block party events are categorized as spontaneous. Almost all neighborhoods (97.5%) have a spontaneous ratio above 0.8, but there is still considerable variation in spontaneous ratio between neighborhoods. In Figure 1 (right), we show the spontaneous within each neighborhood of Philadelphia, based on the total number of spontaneous and regular events across the entire time span of our data (2006-2016).

It is interesting to observe that while North Philadelphia contains the neighborhoods with the largest total number of block party events, these North Philadelphia neighborhoods also have a lower spontaneous ratio than the areas of the city that have a smaller number of block party events. Center City and the Northwest and Northeast suburban communities contain the neighborhoods with the highest spontaneous ratios in Philadelphia, despite having the fewest total number of block party events. We will also examine the trend over time in the spontaneous ratio in Section 2.4 below.

2.4 Trends over Time in Community Measures

Figure 1 shows the spatial distribution of our two community measures that have been aggregated over the entire time span of our data (2006- 2016). However, we can also examine how each of these community measures varies over that same period, both in terms of monthly and yearly trends.

In Figure 2, we show the variation by year and variation by month of the total number of block party events and the spontaneous ratio of block party events that we introduced in Section 2.2 and 2.3. Note that 2016 is not included in Figure 2 since we only have data for part of that year.

We observe in the top left of Figure 2 that the total number of events has been changing over time, with the number of events increasing in 2006-2008 and then decreasing from 2009 onwards. This range of total events across Philadelphia corresponds to around 5-7 events per block group per year.

In the top right of Figure 2, the spontaneous ratio is around 0.86-0.9 in the earlier years of our data but increases to 0.94 in 2010 and then moves very close to 1 from 2011 onwards. In these more recent years, it seems that almost all block party permits were issued for spontaneous parties rather than non-spontaneous events.

In terms of monthly or seasonal variation, we see a clear trend in the lower left of Figure 2 for a greater number of block party events during the warmer months from May (5) to September (9). The spontaneous

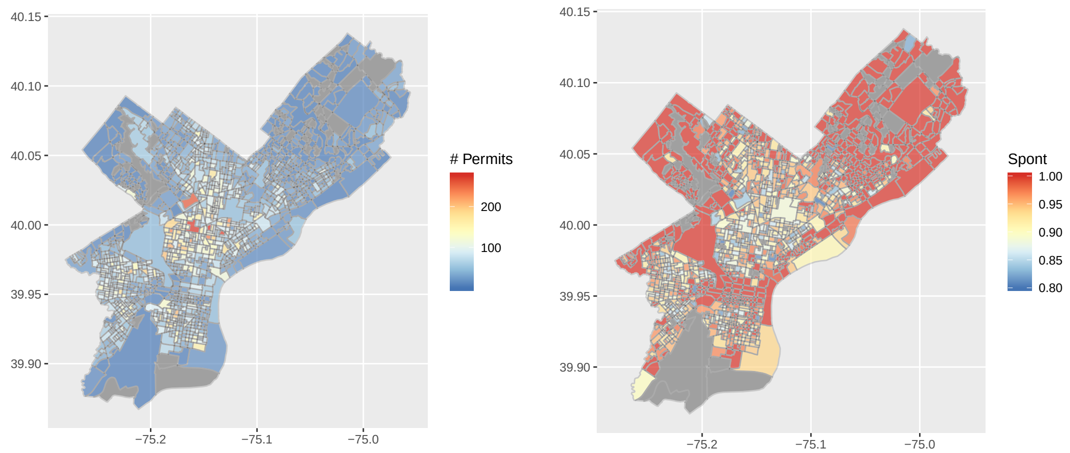


Figure 1. **Left:** Map of Philadelphia showing the total number of permits per neighborhood. **Right:** Map of Philadelphia showing the ratio of spontaneous to regular events per neighborhood

ratio is lower in May (5), July (7), and September (9), which is due to the prominence of regular holidays (Memorial Day, 4th of July, and Labor Day) during those months.

3 Community Vibrancy and other Neighborhood Characteristics

Before we investigate the association between our two measures of community vibrancy and neighborhood safety, we first incorporate into our analysis other neighborhood characteristics that could be associated with either community vibrancy or neighborhood safety. For example, Wu et al. (2018) define neighborhood vibrancy using a GPS-based activity survey in suburban Beijing and found that high density and mixed land use were positively correlated with neighborhood vibrancy. In this section, we explore the relationship between our measures of community vibrancy and various economic, demographic, and built environment characteristics of Philadelphia neighborhoods.

Our demographic data come from the 2010 Decennial Census whereas our economic data come from the 2015 American Community Survey. Land use data is provided by the City of Philadelphia through the `opendataphilly.org` data portal. This data gives the area and land use zoning designation for every single lot in Philadelphia. From these additional datasets, we construct the following measures for each neighborhood (i.e., census block group) in Philadelphia:

- **Demographic measures:** total population and the proportion of residents that identify as White, Black, Asian, Hispanic, or other
- **Economic measures:** mean household income, poverty index (0 = poorest, 1 = wealthiest)
- **Built environment measures:** total area and the proportion of that area designated as commercial vs. residential vs. vacant land

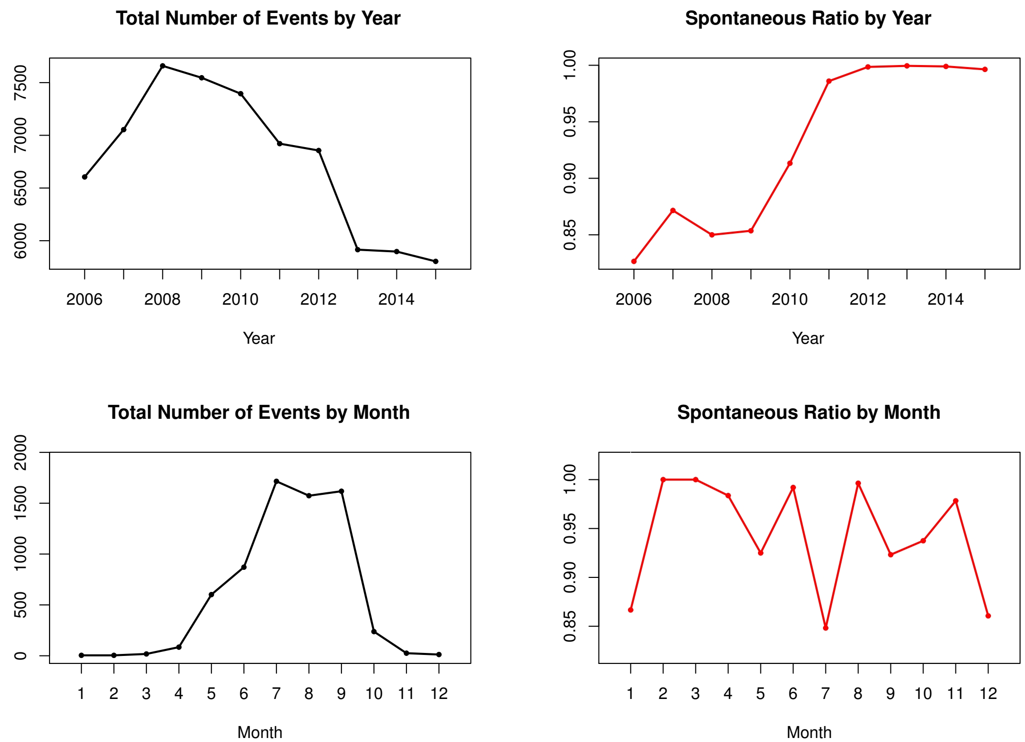


Figure 2. Yearly and monthly trends in the total number of block party events and the spontaneous ratio of block party events. Blue lines are the 95% confidence intervals for those yearly or monthly averages. Updated plots.

Figure 3 provides correlations between our measures of community vibrancy and these demo- graphic, economic, and built environment measures across all neighborhoods in Philadelphia. This correlation matrix also includes several crime measures that we will introduce in Section 4.

We observe that the spontaneous ratio of block parties is not strongly correlated with any of these other neighborhood characteristics. However, the total number of block party permits is correlated with both economic measures (median income and poverty index) as well as the proportion of Black residents in a neighborhood. We also see that the total number of block party permits is correlated with two measures of crime incidence that we will introduce in Section 4, and that those crime measures are strongly correlated with several other neighborhood characteristics.

The association between community vibrancy, crime incidence, and these other neighborhood characteristics means that any comparison of crime incidence that we make between high vibrancy and low vibrancy neighborhoods could be confounded by an imbalance on these other neighborhood characteristics. This imbalance is apparent in Figure 4 where we see significant differences in median

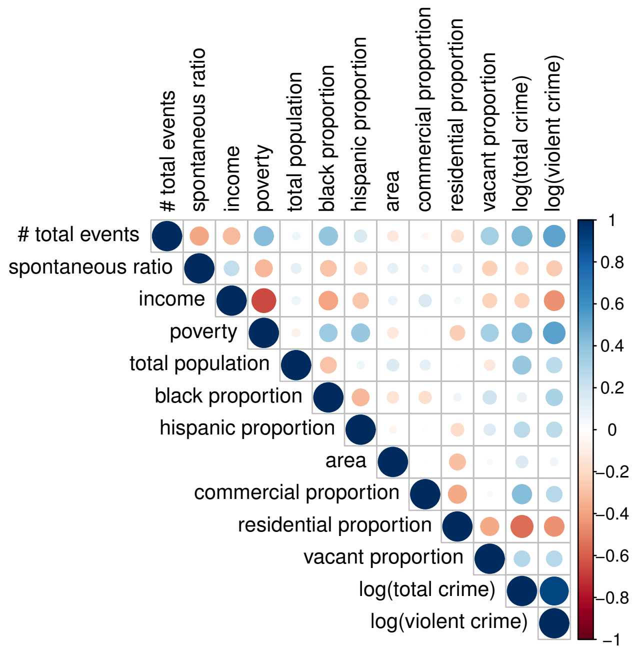


Figure 3. Correlations between community vibrancy, demographic, economic, land use and crime measures across all block groups in Philadelphia. Blue indicates a positive correlation, while red reflects a negative correlation. The darker the shade, the larger the magnitude of the correlation.

household income, poverty metric, and proportion of Black population between high and low vibrancy neighborhoods in Philadelphia.

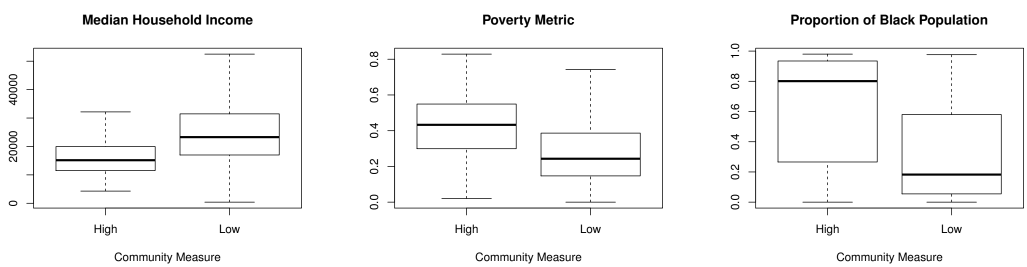


Figure 4. Distribution of median household income, poverty metric, and proportion of Black population between high and low vibrancy neighborhoods in Philadelphia.

In our investigation into the relationship between community vibrancy and safety, we will employ two different approaches to account for the imbalance in these other neighborhood characteristics: linear regression modeling and propensity score matching.

4 Crime Data in Philadelphia

Our crime data comes from the Philadelphia Police Department through the opendataphilly.org data portal and includes all reported crimes in Philadelphia from January 1, 2006 to December 31, 2015. For each reported crime, we have the date, time, and location in terms of GPS latitude and longitude (WGS84 decimal degrees). Each crime is also categorized into one of several types: homicide, sex crime, armed robbery, assault, burglary, theft, motor vehicle theft, etc.

We aggregate all reported crimes within the 1,336 neighborhoods (as defined by the US Census block groups) for which we have calculated measures of community vibrancy in Section 1. We provide maps of the spatial distribution of crime across Philadelphia in our supplementary materials.

Figure 5 (left) gives the distribution of total crimes over the entire time span across these 1,336 neighborhoods. Since that distribution is highly skewed, we will focus on the log transformation of crime in our analyses, which has the more symmetric distribution shown in Figure 5 (right).

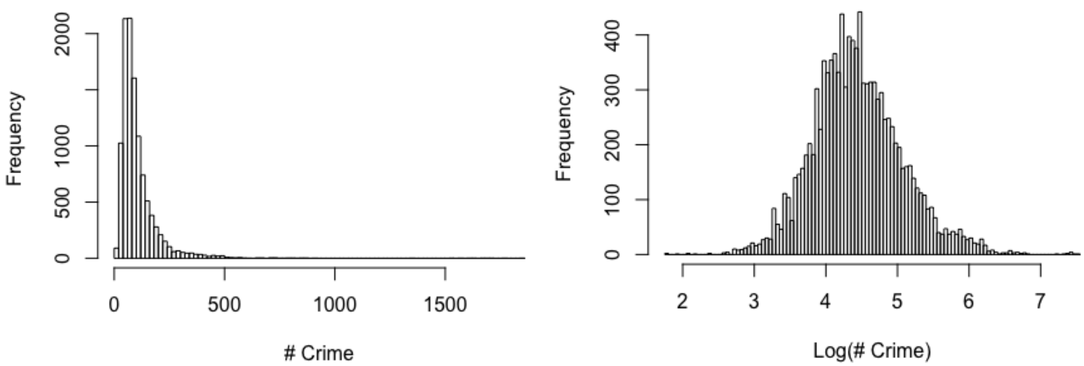


Figure 5. Left: Distribution of the total number of crimes by US census block group in Philadelphia and Right: Distribution of the logarithm of total crimes.

We also make a distinction between violent, non-violent (property), and vice crimes in our analysis. As defined by the Uniform Crime Reporting program of the FBI, violent crimes include homicides, rapes, robberies, and aggravated assaults whereas non-violent crimes include burglaries, thefts, and motor vehicle thefts. Vice crimes include drug violations, gambling, and prostitution.

In Figure 6, we examine the trend from 2006-2015 which are the years for which we also have complete block party data. We provide separate time trends for total crimes as well as subdividing crimes into three major categories: violent, non-violent, and vice crimes. We see in Figure 6 that each type of crime has generally declined over the time span of our data.

In Section 5, we examine whether there is an association between our measures of community vibrancy from Section 1 and total crime incidence and then in Section 6, we investigate the relationship between

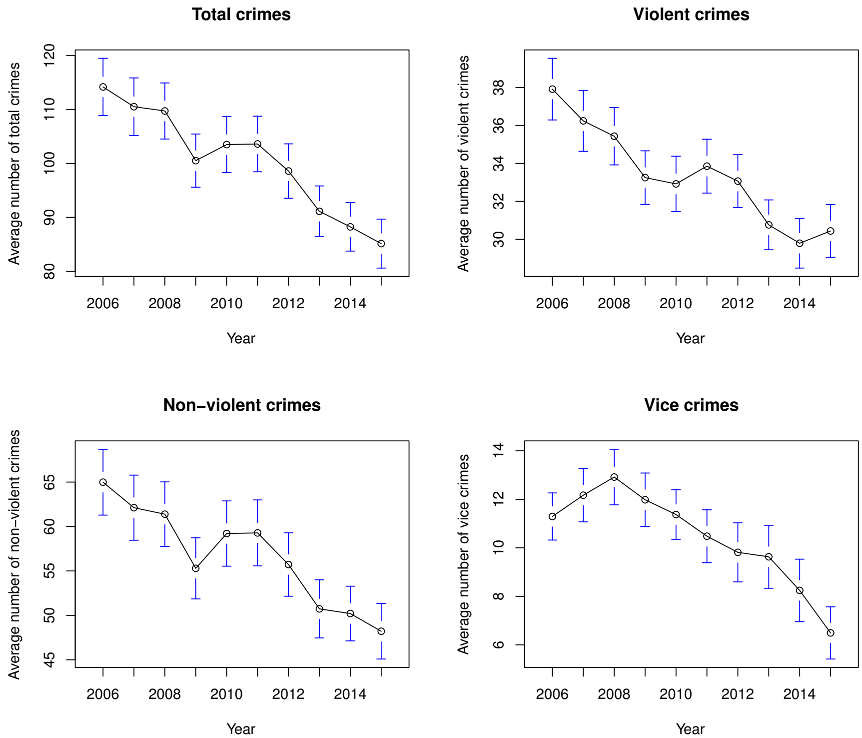


Figure 6. Trends over time in different types of crimes. Y-axis is the average number of crimes per year across all block groups in Philadelphia. Blue lines are the 95% confidence intervals for those yearly averages.

trends over time in community vibrancy and trends over time in crime at the neighborhood level in Philadelphia.

5 Association between Overall Community Vibrancy and Crime

In this section, we investigate the relationship between crime incidence and our measures of community vibrancy at the neighborhood level over the entire 2006-2016 time span of our crime and block party permit data. We will employ two different analyses in order to account for other characteristics of Philadelphia neighborhoods: regression modeling and propensity score matching.

5.1 Linear Regression Analysis of Total Crime and Community Vibrancy

In this regression approach, we consider total crime incidence from 2006-2016 within each neighborhood as our outcome variable and we are interested in whether our measures of community vibrancy are significant predictors of this outcome while controlling for other neighborhood characteristics.

Specifically, we consider the following linear model for the logarithm of total crime incidence y_i in block group i:

$$\log(y_i) = \beta_0 + \boldsymbol{\beta} \cdot \mathbf{X}_i + \phi \cdot C_i + \epsilon_i \tag{1}$$

where $\mathbf{X}_i$ are the demographic, economic, and land use characteristics of neighborhood i as outlined in Section 3 and C_i is one of our community vibrancy measures, either the number of total block party permits or the spontaneous ratio for neighborhood i. We are specifically interested in whether the coefficient ϕ is non-zero, which would imply that our measure of community vibrancy is predictive of total crime incidence beyond the other neighborhood characteristics included in the model.

We use a log transformation of total crime incidence y_i since Figure 5 suggests that the log scale for crime is a more reasonable fit to the assumption of normally distributed errors ϵ_i. However, we also consider an alternative regression approach where the total crime incidence y_i is directly modeled as a negative binomial random variable that is a linear function of the same predictor variables as in eqn (1).

We will compare the results from four different regressions that represent each combination of our two community vibrancy measures and our two regression model specifications:

1. Ordinary least squares (OLS) regression of the logarithm of total crime incidence $\log(y_i)$ on the number of total events C_i and other neighborhood characteristics $\mathbf{X}_i$
2. Ordinary least squares (OLS) regression of the logarithm of total crime incidence $\log(y_i)$ on the spontaneous ratio C_i and other neighborhood characteristics $\mathbf{X}_i$
3. Negative binomial regression of total crime incidence $\log(y_i)$ on the number of total events C_i and other neighborhood characteristics $\mathbf{X}_i$
4. Negative binomial regression of total crime incidence $\log(y_i)$ on the spontaneous ratio C_i and other neighborhood characteristics $\mathbf{X}_i$

As detailed in Section 3, our set of other neighborhood characteristics $\mathbf{X}_i$ for each block group i consist of the total population and fraction of White, Black, Asian and Hispanic residents, our poverty metric and the log of mean household income, and the total area and fraction of that area that is zoned as vacant, commercial or residential. Table 1 displays the parameter estimates and model fit statistics for the four regression models outlined above.

We see in Table 1 that the partial effects for most neighborhood characteristics are significant predictors of crime, which suggests that each of these economic, demographic, and land use characteristics have an association with crime, even after accounting for the other characteristics included in the model. Higher levels of poverty and larger commercial proportions are associated with higher levels of total crime in each of the four models. Higher proportions of vacant land also appear to be associated with more crimes, though the effect is only significant when spontaneity ratio is used as a measure of vibrancy.

However, our primary interest is the association between our measures of community vibrancy and crime, having controlled for these other neighborhood characteristics. In Table 1, we see that the number of total permits is significantly positively associated with total crimes (models 1 and 3), whereas the spontaneous ratio is negatively associated with total crimes (models 2 and 4). We found highly similar results when we ran regression models with (a) just violent crimes, (b) just non-violent crimes, or (c) just vice crimes as outcome variables. Tables and details for these additional regression models are given in our supplementary materials.

Table 1. Regression model summaries for four different models with the number of total crimes as the outcome variable

	Dependent variable:			
	Log number of total crimes		Number of total crimes	
	OLS		*negative binomial*	
	(1)	(2)	(3)	(4)
# Permits	0.003^{***} (0.0003)		0.003^{***} (0.0003)	
Spontaneity ratio		-0.348^{*} (0.166)		-0.315 (0.163)
Log income	0.011 (0.031)	0.017 (0.032)	0.020 (0.031)	0.026 (0.032)
Poverty	0.650^{***} (0.093)	0.727^{***} (0.097)	0.713^{***} (0.092)	0.782^{***} (0.096)
Population (100)	0.049^{***} (0.002)	0.055^{***} (0.002)	0.049^{***} (0.002)	0.056^{***} (0.002)
Black	0.287^{***} (0.038)	0.422^{***} (0.038)	0.272^{***} (0.038)	0.398^{***} (0.037)
Hispanic	0.449^{***} (0.070)	0.572^{***} (0.073)	0.441^{***} (0.070)	0.540^{***} (0.072)
Area (10^6)	0.119^{***} (0.022)	0.092^{***} (0.023)	0.139^{***} (0.022)	0.119^{***} (0.023)
Commercial	2.787^{***} (0.164)	2.724^{***} (0.171)	2.869^{***} (0.162)	2.821^{***} (0.168)
Residential	-1.139^{***} (0.075)	-1.220^{***} (0.078)	-1.260^{***} (0.074)	-1.348^{***} (0.076)
Vacant	0.251 (0.202)	0.606^{**} (0.209)	0.163 (0.200)	0.518^{*} (0.205)
Constant	6.547^{***} (0.340)	6.820^{***} (0.393)	6.553^{***} (0.336)	6.795^{***} (0.386)
Observations	1,265	1,265	1,265	1,265
R^2	0.695	0.667		
Adjusted R^2	0.693	0.664		
Akaike Inf. Crit.			19,765.700	19,863.790
F Statistic (df = 10; 1254)	285.829^{***}	251.313^{***}		
Note:				$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

It is interesting to see that our two measures of community vibrancy have very different associations with crime. Greater numbers of total permits are associated with a greater number of total crimes whereas a larger spontaneity ratio is associated with fewer total crimes. The opposing directions of these associations suggest that our two measures are capturing quite different aspects of community and the relationship between community and crime. To better understand these relationships, we employ a more sophisticated propensity score matching analysis in Section 5.2 below.

5.2 Propensity Score Matching Analysis of Total Crime and Community Vibrancy

In Section 5.1,we used regression models to estimate the association between community vibrancy and total crime, while accounting for the demographic, economic, and land use characteristics of Philadelphia neighborhoods. Matching analyses are an alternative approach for isolating the relationship between community vibrancy and total crime from these other neighborhood characteristics.

In this approach, we create artificial experiments consisting of matched pairs of neighborhoods that have highly similar demographic and economic characteristics but differ substantially in terms of their measures of community vibrancy, which allows us to isolate the relationship between community vibrancy and crime.

We set up two different experiments to investigate each of our two measures of community vibrancy. In the first experiment, we categorize all Philadelphia neighborhoods into a “treatment” group vs. “control” group based on whether their total number of block party permits were above or below the city-wide median of 42.5 block parties. In the second experiment, we categorize all Philadelphia neighborhoods into a “treatment” group vs. “control” group based on whether their spontaneity ratio was above or below the city-wide median of 0.962.

Within each experiment, our goal is to create pairs of neighborhoods consisting of one treatment neighborhood and one control neighborhood that both share highly similar economic, demographic, and land use characteristics. These matched pairs allow us to evaluate the association between crime and our two community vibrancy measures based on within-pair comparisons that are balanced on these other neighborhood characteristics.

We create these matched pairs using a propensity score matching procedure (Rosenbaum and Rubin 1983). The propensity score for each unit (neighborhood) in our analysis is the estimated probability that a particular unit (neighborhood) receives the treatment (high community vibrancy) based on other neighborhood characteristics. We estimate these propensity scores using a logistic regression model with the treatment vs. control indicator as the outcome and the demographic, economic, and land use measures for each neighborhood as predictors.

Two neighborhoods with highly similar demographic, economic, and land use characteristics will have highly similar propensity scores. For each neighborhood in the treatment group (e.g., having a large number of block parties), we will create a matched pair by finding a neighborhood in the control group (e.g., having a small number of block parties) that has a highly similar propensity score. Thus, within each matched pair we have an “apples-to-apples” comparison of two neighbor- hoods that differ in terms of high vs. low community vibrancy but have highly similar other neighborhood characteristics.

In Figure 7, we evaluate the balance in other neighborhood characteristics that we have achieved with our propensity score matching procedure. Specifically, we compare the standardized differences in each neighborhood characteristic between high vs. low community vibrancy neighborhoods before matching to the standardized differences within our matched pairs. We give separate plots for our two different experiments where either the total number of block parties or the spontaneous ratio were used to define our high vs. low community vibrancy groups.

We see in Figure 7 that, for both experiments, our propensity score matching procedure has created matched pairs of neighborhoods with almost no difference in their demographic, economic, and land use characteristics. This balance in other neighborhood characteristics enables us to better isolate the relationship between our two measures of community vibrancy and total crime.

We then use our created sets of matched pairs to estimate the effect of having high community vibrancy on total crime at the neighborhood level. Specifically, when using the total number of block party permits as our measure of community vibrancy, we find that the average within-pair difference in log total crimes is 0.223 between the high vibrancy neighborhood and the low vibrancy neighborhood. The 95% confidence interval on that average within-pair difference is [0.173, 0.273]. So we find that log total crimes are significantly higher in neighborhoods with a large number of block party permits compared to their matching neighborhoods that have a small number of block party permits.

When using the spontaneous ratio as our measure of community vibrancy, we find that the average within-pair difference in log total crimes is -0.991 between the high spontaneous ratio neighborhood and the low spontaneous ratio neighborhood. The 95% confidence interval on that average within-pair

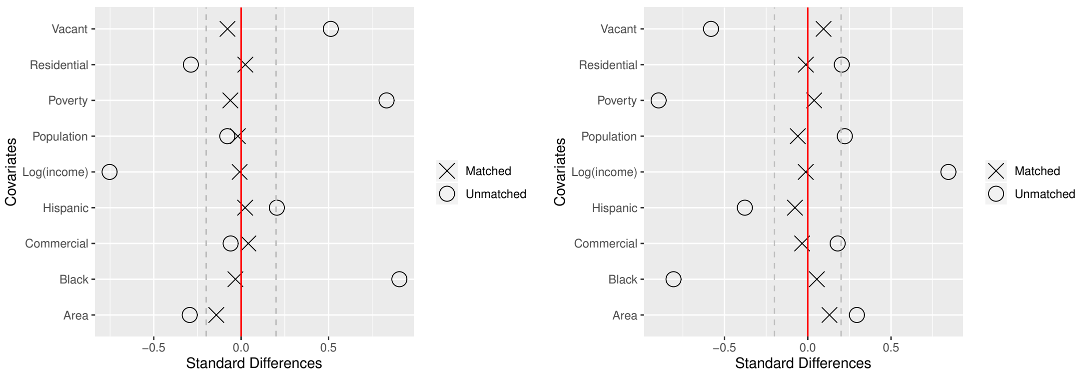


Figure 7. Standardized differences between neighborhoods with high vs. low community vibrancy, both before and after propensity score matching. Left plot is the experiment with total number of permits as the measure used to define the high vs. low community vibrancy group while the right plot is the experiment with spontaneous ratio as the high vs. low community vibrancy measure.

difference is [-0.148, -0.050]. So we find that log total crimes are significantly lower in neighborhoods with a high spontaneous ratio compared to their matching neighborhoods that have a low spontaneous ratio.

hese propensity-score matching results confirm the findings from our regression analysis: our two measures of community vibrancy have significant associations with total crime over the 2006-2016 time span of our data. However, these associations are in opposing directions with greater numbers of total permits associated with a greater number of total crimes and a greater spontaneity ratio associated with fewer total crimes. In the following section, we see how these measures of community vibrancy and crime have changed over time.

6 Trends in Block Parties and Crime over Time

In Section 5, we found significant associations between overall levels of crime and community vibrancy at the neighborhood level, when accounting for other characteristics of those neighborhoods. However, levels of crime and our measures of community vibrancy have all changed substantially over this time period across Philadelphia. In this section, we investigate the relationship between changes in crime incidence over time and the changes in community vibrancy over time at the neighborhood level.

As a reminder, we can compare the overall trends in yearly crime incidence in Figure 6 to the trends by year in our two community vibrancy measures in Figure 8. We see that both the number of permits and total crime incidence have a decreasing trend while the spontaneity ratio has an increasing trend over the time span of our data.

However, trends over time in either crime incidence or community vibrancy can vary substantially between different neighborhoods across the city. We are interested in the association between trends over time in crime incidence and trends over time in community vibrancy across these different neighborhoods. We will again employ two different analyses in order to account for other characteristics of Philadelphia neighborhoods: regression modeling and propensity score matching.

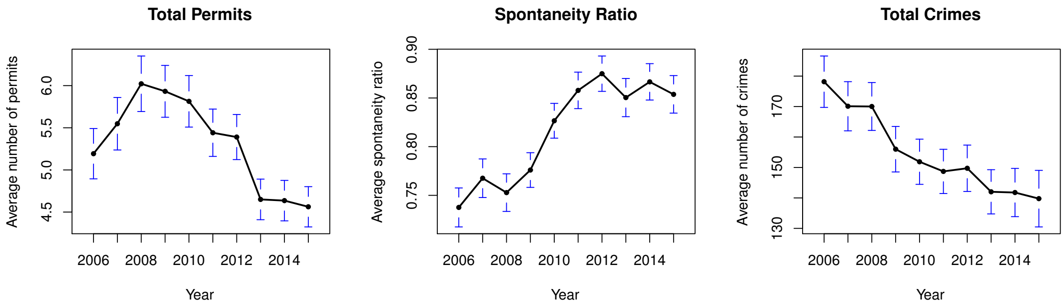


Figure 8. Trend over time in (left) total number of permits (center) spontaneity ratio and (right) total crime incidence. Y-axis is the average number of permits (or crimes) per year across all blockgroups in Philadelphia. Blue lines are the 95% confidence intervals for those yearly averages.

6.1 Regression Analysis of Trends over Time

. We summarize the trend over time in crime within each neighborhood by fitting a separate linear regression of the yearly number of total crimes within each neighborhood on year, and then classifying neighborhoods according to their slope on crime over time. Only 18 neighborhoods (1.4%) had a significantly positive linear trend in crime over time, whereas 540 neighborhoods (42.4%) had a significantly negative linear trend in crime over time.

Similarly, we summarize the trend over time in community vibrancy within each neighborhood by fitting a separate linear regression of the yearly number of block party permits within each neighborhood on year, and then classifying neighborhoods according to their slope on number of permits over time. Only 94 neighborhoods (7.4%) had a significantly positive linear trend in number of permits over time, whereas 184 neighborhoods (14.4%) had a significantly negative linear trend in number of permits over time.

We will focus our regression analyses on determining the neighborhood factors that are predictive of whether or not a neighborhood has a significant trend over time in either crime or our measures of community vibrancy. Specifically, we fit the four different logistic regression models enumerated below:

1. Logistic regression with significantly increasing trend in permits (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators trends in crime) as the predictors
2. Logistic regression with significantly decreasing trend in permits (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators trends in crime) as the predictors
3. Logistic regression with significantly increasing trend in crime (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators trends in permits) as the predictors
4. Logistic regression with significantly decreasing trend in crime (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators trends in permits) as the predictors

Table 2 displays the parameter estimates and model fit statistics for the four logistic regression models listed above, where we use the number of block party permits as our measure of community. We see that log income is a strong predictor of significantly increasing trends in block party permits (model 1) and that vacant proportion is a strong predictor of significantly decreasing trends in block party permits (model 2). We see that the Hispanic proportion is a strong predictor of a significantly decreasing trend in crime (model 4). It is worth noting that for predicting significantly increasing trends in either block party permits or crimes, there are so few cases of either of those two outcomes ($n = 94$ in model 1 and $n = 38$ in model 3), which gives us limited power to detect strong associations.

Table 2. Logistic regression model results for predicting neighborhoods with different types of significant trends over time

	Dependent variable:			
	Permits +	Permits −	Crimes +	Crimes −
	(1)	(2)	(3)	(4)
Log income	0.084^{***} (0.024)	−0.0003 (0.032)	−0.010 (0.016)	0.109^{*} (0.046)
Poverty	0.004 (0.071)	0.145 (0.097)	0.036 (0.047)	0.028 (0.136)
Population (100)	0.001 (0.002)	0.003 (0.002)	0.0004 (0.001)	0.006 (0.003)
Black	−0.027 (0.027)	0.017 (0.037)	0.041^{*} (0.018)	−0.058 (0.052)
Hispanic	−0.003 (0.053)	0.097 (0.072)	0.016 (0.035)	0.327^{***} (0.101)
Area (10^6)	−0.018 (0.017)	−0.016 (0.023)	0.008 (0.011)	−0.050 (0.033)
Commercial	−0.025 (0.126)	−0.223 (0.171)	0.185^{*} (0.083)	−0.227 (0.240)
Residential	-0.115^{*} (0.057)	0.110 (0.078)	−0.039 (0.038)	-0.257^{*} (0.110)
Vacant	−0.292 (0.153)	0.706^{***} (0.207)	0.116 (0.101)	0.512 (0.293)
Crimes +	0.012 (0.043)	−0.060 (0.059)		
Crimes −	0.004 (0.015)	0.012 (0.020)		
Permits +			0.003 (0.019)	0.016 (0.054)
Permits −			−0.016 (0.014)	0.033 (0.040)
Constant	-0.700^{***} (0.261)	−0.021 (0.354)	0.089 (0.173)	−0.594 (0.501)
Outcome = 1	94	184	38	589
Observations	1,265	1,265	1,265	1,265
Log Likelihood	−66.428	−452.955	458.784	−889.082
Akaike Inf. Crit.	156.857	929.910	−893.567	1,802.164
Note:				*p<0.05; **p<0.01; ***p<0.001

In Table 2, we see that trends in crimes are not predictive of trends in the number of block party permits and vice versa. We fit the same four logistic regression models but using spontaneous ratio as our measure of community and the results are given in Table S4 of our supplementary materials. In Table S4, we see that trends in crimes are also not predictive of trends in the spontaneous ratio and vice versa. These results suggest that there is not an association between trends over time in crime and trends over time in our two measures of community vibrancy. However, we further investigate this possibility with a propensity score matching analysis in Section 6.2.

6.2 Propensity Score Matching for Examining Trends over Time

Similar to our approach in Section 5.2, we create artificial experiments consisting of matched pairs of neighborhoods that have highly similar demographic and economic characteristics but the two neighborhoods within each pair differ substantially in terms of their trend over time in community vibrancy. This approach allows us to isolate the relationship between trends over time in community vibrancy and trends over time in crime.

For example, we can categorize all neighborhoods based on whether they have a significantly positive trend in the number of block party permits or not. We label neighborhoods with a significantly positive trend in the number of block party permits as the “treatment” group and label all other neighborhoods as the “control” group. Just as in Section 5.2, we fit a logistic regression with these treatment vs. control labels as the outcome variable and all other neighborhood factors (demographic, economic and land use) as predictor variables of that outcome. From this fitted model, the probability of a neighborhood being in the treatment group is called the *propensity score* for that neighborhood.

We then match up each neighborhood in the treatment group with a neighborhood from the control group with the closest possible propensity score. In this way, we form a set of matched pairs where each pair of neighborhoods have highly similar demographic, economic, and land use characteristics but one of those neighborhoods has a significantly positive trend in the number of block party permits and the other neighborhood does not.

Figure 9 compares the standardized differences between neighborhoods before and after propen- sity score matching for two of the experiments that we perform. In the left plot, the treatment group is neighborhoods that have a significantly increasing trend over time in block party permits whereas in the right plot, the treatment group is neighborhoods that have a significantly increasing trend over time in spontaneous ratio.

Figure 9. Standardized differences between neighborhoods that have a significantly positive trend over time in block party permits and neighborhoods that do not, both before and after propensity score matching. Left: Number of permits is used as a measure of vibrancy. Right: Spontaneity ratio is used.

We see that, for both experiments, our matching procedure has created pairs of neighborhoods with almost no difference in their demographic, economic, and land use characteristics, which makes for a more balanced comparison of crime between neighborhoods that have significantly positive trends over time in either of our two community vibrancy measures.

We considered 12 different experiments with this same type of propensity score matching design with each experiment being a different combination of four definitions for the treatment variable and three crime outcomes. The four treatment variables considered were: 1. having a significantly positive trend over time in block party permits, 2. having a significantly negative trend over time in block party permits, 3. having a significantly positive trend over time in the spontaneous ratio, and 4. having a significantly negative trend over time in spontaneous ratio. For each of these different treatments, we evaluated our matched pairs of neighborhoods for differences in three crime trend outcomes: 1. the slope on the trend over time in total crime, 2. an indicator for a significantly positive crime trend (or not), and 3. an indicator for a significantly negative crime trend (or not).

Table 3 gives the average within-pair differences between the treatment and control groups (and 95% confidence intervals for those averages) for all 12 combinations outlined above. We see that 11 of the 12 comparisons do not yield statistically significant results. However, we do find that neighborhoods with a significantly positive trend in their spontaneous ratio also have significantly more negative trends over time in total crimes. This is the only significant association we have been able to detect between trends over time in crime and trends over time in our two measures of community vibrancy.

Table 3. Average within-pair differences between the treatment and control groups (and 95% confidence intervals) for all 12 combinations of four treatment variables (columns) and crime outcomes (rows). For the “crime slope” outcome, the difference between slopes is provided, whereas for the “Crime +” and ”Crime -” indicators, the odds ratio is provided.

	Treatment			
Outcome	# Permits +	# Permits −	Spont +	Spont −
Crime slope	-1.4341 $[-4.7230, 1.8548]$	-0.6030671 $[-2.9085, 1.7024]$	-2.1928^{***} $[-3.8507, -0.5350]$	0.6705 $[-1.1154, 2.4563]$
Crime +	1.0085 [0.9356, 1.0509]	0.9933 [0.9588, 1.0290]	1.0053 [0.9724, 1.0393]	$0.9902^{\dagger}$ [0.9367, 1.0468]
Crime −	1.0393 [0.8852, 1.2201]	1.0503 [0.9364, 1.1781]	1.0558 [0.9715, 1.1475]	1.1560 [0.8970, 1.4899]

Note: *p<0.05; **p<0.01; ***p<0.001 with Wilcoxon
†Estimates from many to one matching rather than 1:1 due to imbalance

7 Summary and Discussion

In this paper, we explore the relationship between crime incidence at the neighborhood level and measures of community vibrancy created from a unique dataset of block party permit approvals in the city of Philadelphia. In order to properly analyze this relationship, we must account for the economic, demographic, and land use characteristics of these neighborhoods, which may also influence both community vibrancy and crime incidence. We employ two statistical techniques, regression modeling and propensity score matching, in order to isolate the association between crime and community vibrancy while controlling for other neighborhood characteristics.

We have found significant associations between aggregate levels of crime and our two measures of community vibrancy at the neighborhood level, while accounting for other characteristics of those neighborhoods. Neighborhoods with more block parties have a significantly higher crime rate, while those holding a greater proportion of spontaneous events have a significantly lower crime rate. We also find that neighborhoods with an increasing spontaneous ratio over time are significantly more likely to have a decreasing trend in total crime incidence over time.

However, the relationships between community vibrancy and public safety are subtle, nuanced, and presumably influenced by many types of neighborhood contexts. Thus, higher resolution data and measures of community vibrancy, such as direct measures of human occupancy and usage of public spaces, are needed for future study.

References

Bellair PE (1997) Social interaction and community crime: Examining the importance of neighbor networks. *Criminology* 35(4): 677–704.

Crutchfield RD, Geerken MR and Gove WR (1982) Crime rate and social integration the impact of metropolitan mobility. *Criminology* 20(3-4): 467–478.

Geeting J (2016) Mapping 10 years of Philadelphia block parties. *Plan Philly* : https://whyy.org/articles/mapping–10–years–of–philadelphia–block–parties/.

Humphrey C, Jensen ST, Small DS and Thurston R (2017) Urban vibrancy and safety in philadelphia. *Environment and Planning B: Urban Analytics and City Science* : 2399808319830403.

Rosenbaum PR and Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41–55.

Sampson RJ and Groves WB (1989) Community structure and crime: Testing social-disorganization theory. *American journal of sociology* 94(4): 774–802.

Shaw CR and McKay HD (1942) Juvenile delinquency and urban areas. .

Wu J, Ta N, Song Y, Lin J and Chai Y (2018) Urban form breeds neighborhood vibrancy: A case study using a gps-based activity survey in suburban beijing. *Cities* 74: 100–108.

Supplementary Materials for "Community Vibrancy and its Relationship with Safety in Philadelphia"

8. Spatial Distribution of Total Crimes in Philadelphia. Figure 9 is a map of the spatial distribution of total crimes per year (averaged over the years from 2006-2015) in Philadelphia, as well as the log transformation of total crimes per year.

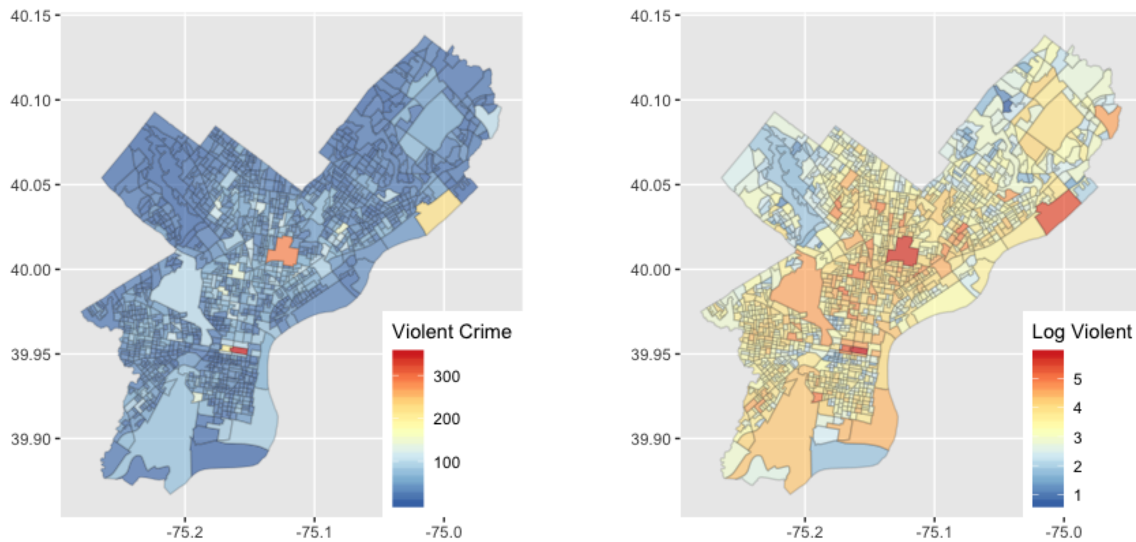


Fig 9: Distribution of violent crime over the block groups of Philadelphia. ***Left:*** *violent crimes per block group, averaged over the years from 2006 to 2015.* ***Right:*** *logarithm of violent crimes per block group, averaged over the years from 2006 to 2015.*

9. Additional Aggregate Regression Models. In our main paper, we compare the results from four different regressions that represent each combination of our two community vibrancy measures and our two regression model specifications,

1. Ordinary least squares (OLS) regression of the logarithm of total crime incidence $\log(y_i)$ on the number of total events C_i and other neighborhood characteristics $\mathbf{X}_i$
2. Ordinary least squares (OLS) regression of the logarithm of total crime incidence $\log(y_i)$ on the spontaneous ratio C_i and other neighborhood characteristics $\mathbf{X}_i$
3. Negative binomial regression of total crime incidence $\log(y_i)$ on the number of total events C_i and other neighborhood characteristics $\mathbf{X}_i$
4. Negative binomial regression of total crime incidence $\log(y_i)$ on the spontaneous ratio C_i and other neighborhood characteristics $\mathbf{X}_i$

In this section, we provide results from a similar set of regressions but with (a) just violent crimes, (b) just non-violent crimes or (c) just vice crimes as the outcome variable. Tables S1, S2, and S3 displays the parameter estimates and model fit statistics for the four regression models listed above but with violent crime, non-violent crime and vice crime as the outcome variable, respectively.

We generally observe similar results in Tables S1-S3 to Table 1 in our main paper where total crimes was the outcome variable. The partial effects for most neighborhood characteristics are

Table S1

Results from linear regression models with the number of violent crimes as the outcome variable

	Dependent variable:			
	Log number of violent crimes		Number of violent crimes	
	OLS		*negative binomial*	
	(1)	(2)	(3)	(4)
# Permits	0.003^{***} (0.0003)		0.003^{***} (0.0003)	
Spontaneous ratio		-0.422^{**} (0.190)		-0.322^{*} (0.185)
Log income	-0.197^{***} (0.036)	-0.191^{***} (0.037)	-0.203^{***} (0.035)	-0.204^{***} (0.036)
Poverty	0.377^{***} (0.108)	0.459^{***} (0.112)	0.494^{***} (0.105)	0.561^{***} (0.109)
Population (100)	0.047^{***} (0.003)	0.055^{***} (0.003)	0.049^{***} (0.003)	0.056^{***} (0.003)
Black	0.644^{***} (0.044)	0.789^{***} (0.044)	0.585^{***} (0.043)	0.711^{***} (0.043)
Hispanic	0.685^{***} (0.081)	0.816^{***} (0.084)	0.574^{***} (0.079)	0.666^{***} (0.082)
Area (10^6)	0.084^{***} (0.026)	0.055^{**} (0.027)	0.110^{***} (0.025)	0.083^{***} (0.026)
Commercial	2.345^{***} (0.189)	2.277^{***} (0.196)	2.275^{***} (0.185)	2.223^{***} (0.190)
Residential	-1.018^{***} (0.086)	-1.105^{***} (0.089)	-1.168^{***} (0.085)	-1.268^{***} (0.087)
Vacant	-0.294 (0.233)	0.088 (0.240)	-0.549^{**} (0.228)	-0.219 (0.232)
Constant	6.638^{***} (0.392)	6.980^{***} (0.451)	6.852^{***} (0.385)	7.176^{***} (0.439)
Observations	1,265	1,265	1,265	1,265
R^2	0.668	0.641		
Adjusted R^2	0.666	0.639		

Significance Levels: *p<0.1; **p<0.05; ***p<0.01

Table S2

Results from linear regression models with the number of non-violent crimes as the outcome variable

	Dependent variable:			
	Log number of non-violent crimes		Number of non-violent crimes	
	OLS		*negative binomial*	
	(1)	(2)	(3)	(4)
# Permits	0.002^{***} (0.0003)		0.002^{***} (0.0003)	
Spontaneous ratio		-0.258 (0.158)		-0.219 (0.157)
Log income	0.086^{***} (0.030)	0.090^{***} (0.031)	0.096^{***} (0.030)	0.101^{***} (0.030)
Poverty	0.561^{***} (0.091)	0.611^{***} (0.093)	0.602^{***} (0.091)	0.648^{***} (0.092)
Population (100)	0.053^{***} (0.002)	0.058^{***} (0.002)	0.054^{***} (0.002)	0.058^{***} (0.002)
Black	0.030 (0.037)	0.117^{***} (0.036)	0.008 (0.037)	0.086^{**} (0.036)
Hispanic	0.118^{*} (0.069)	0.196^{***} (0.070)	0.076 (0.068)	0.137^{**} (0.069)
Area (10^6)	0.115^{***} (0.022)	0.098^{***} (0.022)	0.142^{***} (0.022)	0.130^{***} (0.022)
Commercial	2.885^{***} (0.160)	2.845^{***} (0.163)	3.002^{***} (0.158)	2.970^{***} (0.161)
Residential	-1.250^{***} (0.073)	-1.303^{***} (0.074)	-1.379^{***} (0.073)	-1.435^{***} (0.073)
Vacant	0.163 (0.198)	0.391^{**} (0.199)	0.129 (0.196)	0.350^{*} (0.197)
Constant	4.817^{***} (0.333)	5.027^{***} (0.375)	4.813^{***} (0.330)	4.972^{***} (0.371)
Observations	1,265	1,265	1,265	1,265
R^2	0.680	0.667		
Adjusted R^2	0.677	0.664		

Significance Levels: *p<0.1; **p<0.05; ***p<0.01

Table S3

Results from linear regression models with the number of vice crimes as the outcome variable

	Dependent variable:			
	Log number of vice crimes		Number of vice crimes	
	OLS		*negative binomial*	
	(1)	(2)	(3)	(4)
# Permits	0.009^{***} (0.001)		0.008^{***} (0.001)	
Spontaneous ratio		-1.538^{***} (0.545)		-1.129^{***} (0.394)
Log income	-0.345^{***} (0.102)	-0.329^{***} (0.106)	-0.183^{**} (0.074)	-0.165^{**} (0.077)
Poverty	0.976^{***} (0.309)	1.193^{***} (0.320)	1.442^{***} (0.222)	1.586^{***} (0.232)
Population (100)	0.028^{***} (0.008)	0.048^{***} (0.008)	0.026^{***} (0.005)	0.048^{***} (0.005)
Black	1.655^{***} (0.127)	2.050^{***} (0.125)	1.316^{***} (0.091)	1.654^{***} (0.091)
Hispanic	1.651^{***} (0.233)	1.995^{***} (0.241)	1.880^{***} (0.166)	2.120^{***} (0.173)
Area (10^6)	-0.055 (0.074)	-0.135^{*} (0.077)	0.036 (0.053)	-0.042 (0.055)
Commercial	3.533^{***} (0.543)	3.348^{***} (0.563)	3.015^{***} (0.387)	2.985^{***} (0.405)
Residential	-2.038^{***} (0.248)	-2.283^{***} (0.256)	-1.670^{***} (0.179)	-1.897^{***} (0.186)
Vacant	0.043 (0.670)	1.080 (0.687)	0.420 (0.477)	1.403^{***} (0.493)
Constant	5.768^{***} (1.128)	7.082^{***} (1.294)	4.457^{***} (0.814)	5.372^{***} (0.939)
Observations	1,265	1,265	1,265	1,265
R^2	0.546	0.512		
Adjusted R^2	0.543	0.508		
Significance Levels:			*p<0.1; **p<0.05; ***p<0.01	

significant, suggesting that each of these economic, demographic and land use characteristics are associated with violent, non-violent and vice crime. Similar to the models for total crime in our main paper, we see that the number of total permits is significantly positively associated with total crimes, whereas the spontaneous ratio is negatively associated with total crimes.

10. Additional Regression Models for Trends over Time. In our main paper, we used regression models to explore the neighborhoods factors that are predictive of whether or not a neighborhood has a significant trend over time in either crime or our measures of community vibrancy. Specifically, we fit the four different logistic regression models enumerated below:

1. Logistic regression with significantly increasing trend in community (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators of trends in crime) as the predictors
2. Logistic regression with significantly decreasing trend in community (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators of trends in crime) as the predictors
3. Logistic regression with significantly increasing trend in crime (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators of trends in community) as the predictors
4. Logistic regression with significantly decreasing trend in crime (or not) as the outcome and neighborhood characteristics $\mathbf{X}_i$ (including indicators of trends in community) as the predictors

In our main paper, we used the number of block party permits as our measure of community. Table S4 displays the parameter estimates and model fit statistics for another four logistic regression models in which we use the spontaneity ratio as our measure of community. We find that several factors are significantly associated with a positive trend over time in spontaneity ratio: poverty, population size, proportions of Black and Hispanic residents, neighborhood's area, residential zone, and vacant land (column 1). On the other hand, the only significant predictors of a much smaller set of neighborhoods with a negative trend in spontaneous ratio are the proportion of Black residents and residential zone (column 2). There is no significant association between trends over time in crime and trends over time in the spontaneous ratio

TABLE S4
Logistic regression model results for predicting neighborhoods with different types of significant trends over time

	Dependent variable:			
	Spont +	Spont −	Crimes +	Crimes −
	(1)	(2)	(3)	(4)
Log income	0.046 (0.037)	0.014 (0.013)	−0.009 (0.016)	0.108* (0.045)
Poverty	0.320*** (0.112)	0.001 (0.040)	0.035 (0.047)	0.018 (0.137)
Population (100)	0.006** (0.003)	0.001 (0.001)	0.0004 (0.001)	0.006 (0.003)
Black	0.290*** (0.043)	−0.042*** (0.015)	0.041* (0.018)	−0.069 (0.053)
Hispanic	0.340*** (0.083)	−0.034 (0.030)	0.015 (0.035)	0.316*** (0.102)
Area (10^6)	−0.061* (0.027)	0.010 (0.010)	0.008 (0.011)	−0.049 (0.033)
Commercial	−0.345 (0.197)	−0.065 (0.070)	0.187* (0.083)	−0.216 (0.241)
Residential	−0.214* (0.090)	0.074* (0.032)	−0.042 (0.038)	−0.249* (0.110)
Vacant	0.764*** (0.239)	0.105 (0.085)	0.108 (0.101)	0.492 (0.293)
Crimes +	−0.006 (0.068)	−0.001 (0.024)		
Crimes −	0.029 (0.023)	0.003 (0.008)		
Spont +			−0.004 (0.012)	0.045 (0.035)
Spont −			−0.004 (0.033)	0.039 (0.097)
Constant	−0.484 (0.409)	−0.141 (0.146)	0.085 (0.172)	−0.578 (0.500)
Outcome = 1	313	27	38	589
Observations	1,265	1,265	1,265	1,265
Log Likelihood	−634.917	669.765	458.167	−888.535
Akaike Inf. Crit.	1,293.833	−1,315.530	−892.334	1,801.069

Note: *p<0.05; **p<0.01; ***p<0.001

11. Monthly and Daily Analysis. Our main paper focuses on measures of crime and community vibrancy at a yearly level. In this section, we provide some additional exploratory analyses of crime and community vibrancy at the monthly and daily level.

11.1. *Month-Level.* We aggregate the number of block party permits and the number of crimes to a month level. We find that the number of monthly parties is generally positively correlated with the number of monthly crimes. This is partially because more events are happening in the spring-summer-fall months. We observe a negative correlation between them from November to March (see Figure 10. Figure 11 illustrates the average numbers of crimes and block party permits for each month from 2006 to 2015. The figure on the right reveals an interesting insight. While there are unfortunately more crimes than parties every month, on average there are actually slightly more parties than crime. This is likely a result of consistency in permits. Crime has a higher variation over time.

11.2. *Day-Level.* We can further investigate whether recent or same-day block parties have any effects on crime. Figure 12 displays the average number of crimes and permits for each day of week (1 is Sunday and 7 is Saturday). Crime happens at a relatively same rate throughout the week with a slight peak on Tuesdays, while block parties happen most frequently on Saturdays.

Fig 10: Correlation between permits and crimes per month

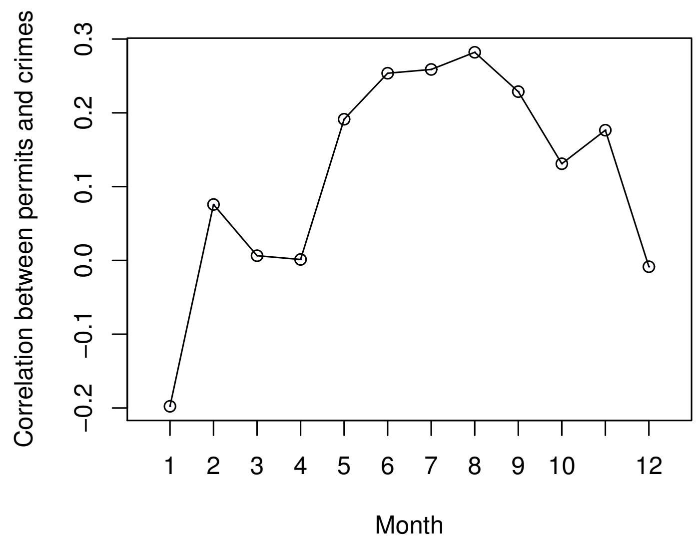


Fig 11: Average number of crimes and permits per month

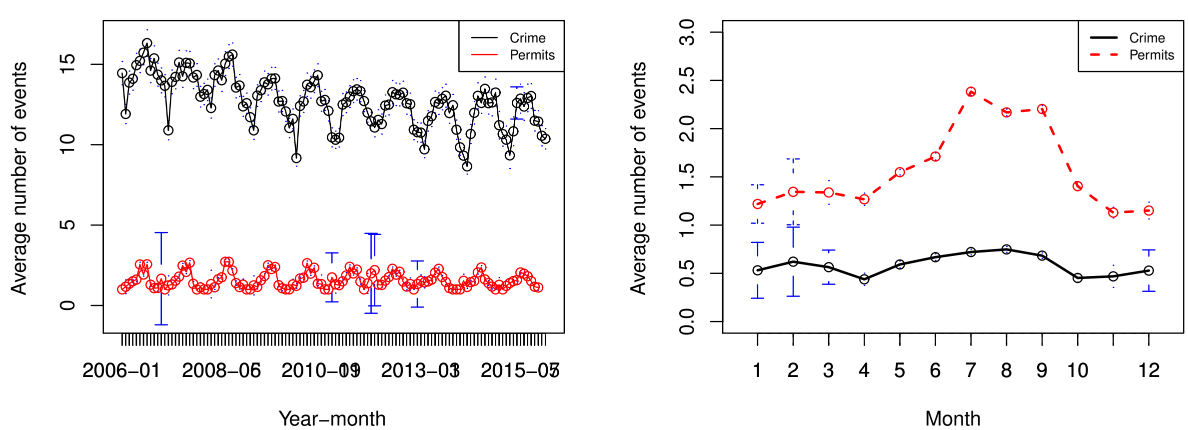
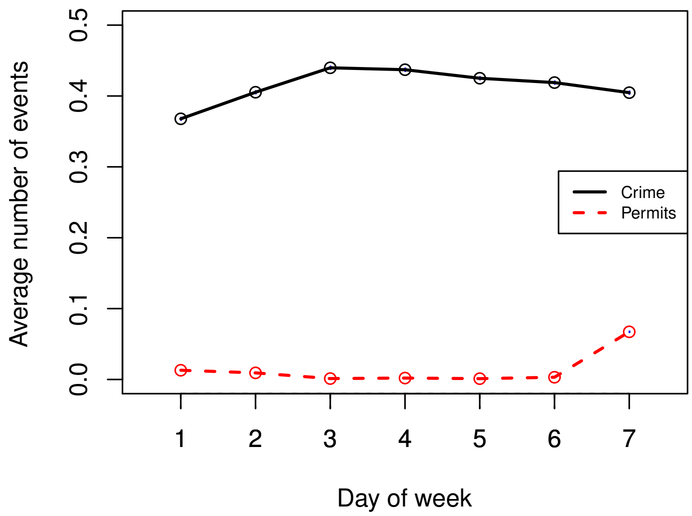


Fig 12: Average number of crimes and permits per day of week



We also observe that whether or not having a party, the crime rates are not significantly different. The likelihood of same-day crime is 28.28% for days without parties, while it is 28.32% for days with parties.

11.3. *Level of In/Activities.* We introduce four new measures that represent recent events. *Cumulative parties* or *party days* is the number of days of consecutive parties in the block group up until the focal day. Such measure gets reset to zero every time the streak of parties breaks. Similarly, *cumulative crimes* or *crime days* is the number of days of consecutive crimes up until the focal day. These two measures represent a level of activities in recent history. On the other hand, we can also compute *quiet days*, which is the number of consecutive days without any parties in the block group up until the focal date, and *safe days*, which is the number of consecutive days without any crimes. These two measures instead reflect level of inactivities in the neighborhood.

The majority of block group-days have 0 party days, meaning that most parties (within the same block group) did not happen on consecutive days. The mean number of party days is 0.012, while the range is from 0 to 7 days. The distribution for crime days is more spreaded out, but still right-skewed. On average, crime happened for 0.549 days in a row and this measure ranges from 0 to 551 days. The distributions for inactivities are even further spreaded out. Average length of quiet days is 303 days (almost a year!) with a median of 130 days, while the average safe days is 346.8 days with a median of only 5 days, unfortunately.

The correlations among these measures and whether there was a crime on the focal day are fairly small. Correlation between crime indicator and crime days is 0.092, followed by correlation between crime indicator and quiet days which is -0.023 (e.g., longer consecutive quiet days are associated with smaller chance of crime). We then perform regression models to relate either the crime binary or the number of crimes per day with whether there was a party on the same day and the four measures of in/activities. Table 8 reports the estimates.

Summarizing these results, we find that aggregating the number of block party permits and the

Table 8

	Dependent variable:			
	Crime indicator		Number of crime incidents	
	(1)	(2)	(3)	(4)
is_party	-0.006^{***} (0.002)	-0.019^{***} (0.002)	0.0003 (0.003)	-0.027^{***} (0.003)
cumin	-0.00001^{***} (0.00000)	-0.00001^{***} (0.00000)	-0.00002^{***} (0.00000)	-0.00002^{***} (0.00000)
cumsafe	-0.00003^{***} (0.00000)	-0.00003^{***} (0.00000)	-0.0001^{***} (0.00000)	-0.0001^{***} (0.00000)
cumparty	-0.006^{***} (0.002)	-0.017^{***} (0.002)	-0.008^{***} (0.003)	-0.030^{***} (0.003)
cumcrime	0.004^{***} (0.00005)	0.004^{***} (0.00005)	0.015^{***} (0.0001)	0.015^{***} (0.0001)
Month FE	No	Yes	No	Yes
Observations	4,879,072	4,879,072	4,879,072	4,879,072
R^2	0.002	0.004	0.007	0.009
Adjusted R^2	0.002	0.003	0.007	0.008
Note:				*p<0.1; **p<0.05; ***p<0.01

number of crimes to a month level, we find that the number of monthly parties is generally going in the same direction as the number of monthly crimes. This is partially because more events are happening in the spring-summer-fall months. On a day level, crime happens at a relatively same rate throughout the week with a slight peak on Tuesdays, while block parties happen most frequently on Saturdays. However, whether or not having a party, the crime rates are not significantly different.

Wichinpong Park Sinchaisri
Department of Operations, Information, and Decisions
The Wharton School
University of Pennsylvania
400 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, PA 19104
Email: swich@wharton.upenn.edu

Shane T. Jensen
Department of Statistics
The Wharton School
University of Pennsylvania
400 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, PA 19104
Email: stjensen@wharton.upenn.edu