Supporting Information

Branas et al. 10.1073/pnas.1718503115

SI Materials and Methods

Trial Design. A controlled, parallel-group, cluster randomized trial of standardized interventions for vacant lots of land was conducted citywide in Philadelphia. This trial was approved by the University of Pennsylvania Institutional Review Board and registered with the International Standard Randomized Controlled Trial Number (study ID ISRCTN92582209). All sections of this paper were written using the Consolidated Standards of Reporting Trials statement for the reporting of cluster randomized trials (1). Geographic information systems technology (ArcGIS 10; ESRI) was used throughout the trial to support cluster, lot, and participant selection and follow-up and spatial data calculations.

The trial used a random selection procedure followed by a stratified random assignment of eligible vacant lots into intervention and no-intervention arms matched within four city sections: north, south, west/southwest, and northwest. Intervention status was randomly assigned and matched within each of the four city sections to promote comparability between trial arms. Clearly delineated roadway and water boundaries were used to define these four geographically and demographically distinct city sections.

We also integrated a qualitative ethnographic component to the project. A team of ethnographers: (*i*) conducted conversational-style semistructured interviews, including oral histories of vacant lots from the perspective of neighbors; (*ii*) collected detailed in situ observational field notes in a representative range of microneighborhoods; and (*iii*) assembled archival gray literature, social media data, and longitudinal geographic maps, street-view images, and satellite photographs. This qualitative component documented the larger social and political economic community context of vacant lots, both historical and contemporary. It followed a previously tested protocol of direct, real-time participant-observation data collection on an ongoing randomized controlled trial designed to monitor logistics and generate causal hypotheses for the potential mechanisms producing a trial's outcomes (2). This robust qualitative component allowed us to monitor the consistency of time-sensitive field procedure logistics as well as staff and subcontractor fidelity to study protocols. On an analytical level, the ethnography qualitatively explored the particularities of microneighborhood characteristics, identified neighborhood typologies, potentially unexpected or unwanted effects, and documented potential causal mechanisms that might explain differential microneighborhood responses to the interventions (3). The ethnography also generated qualitative hypotheses for further quantitative stratification and analyses (4).

Two microneighborhoods were studied and consisted of ~300 square blocks, adding up to between 2.8 and 3.2 million square meters. We provide only an approximate measure for the number of square blocks and square feet comprising these microneighborhoods because Philadelphia blocks vary in size with small blocks inconsistently interspersed between longer, more standard-sized, 150-m-long blocks. Furthermore, the more precise social boundaries of these microneighborhoods track along various urban infrastructures, including railroad tracks, highways, elevated train lines, bus routes, parks, and public-use buildings. The one poorer microneighborhood was unaffected by rises in property values and dominated by open-air narcotics markets. The vast majority of the population in this microneighborhood was Puerto Rican with a significant presence of African Americans. The second microneighborhood was subject to accelerated development and was majority African American with rapidly growing White and Asian populations.

Random Sampling of Clusters and Participants. From among master lists of all vacant lots citywide available from city records in January 2011 (n = 44,768), vacant lots that were authorized by municipal ordinance as "blighted" and eligible for the intervention (n = 34,149) were randomly sampled for the trial. These authorized/eligible lots constituted 76.3% of lots citywide and were included if they specifically: (*i*) had existing violations signaling blight, including illegal dumping, abandoned cars, and/or unmanaged vegetation growth greater than a certain height; and (*ii*) had been abandoned, as confirmed through contact with the owner of record who was given 10 d to reply and did not; or (*iii*) had been authorized for the intervention by the owner of record (including the city itself for publicly owned lots) within the 10-d period. Excluded were lots that were not eligible because of insufficient blight or lack of authorization (n = 4,284), lots that were >510 m^2 (n = 3,755), and lots that were on existing private or commercial parking lots (n = 2,580). The ethnographic team also accompanied field staff to provide added fidelity to our sampling protocol.

Clusters were then formed as randomly selected places chosen to represent the entire city (5). All 34,149 eligible vacant lots were ordered based on the assignment of random numbers within the four city sections. Polygons representing each eligible vacant lot's parcel of land were assigned the longitude–latitude point of their centroids, or geometric centers. The first vacant lot in the randomly ordered list in each section was then chosen as an "index lot" and a 0.4-km radius buffer circle was generated around its centroid. All other eligible vacant lots within this radius were then used to form a cluster of vacant lots that summed to between 400 and 500 total square meters in area and were excluded from consideration as future index lots. Using satellite images and Google Street View photos, a group of eligible vacant lots that were as geographically close to one another as possible within each cluster was formed.

This process then cycled to the next randomly ordered index vacant lot on the list that was at least 0.4 km away from the edge of prior clusters until a total of 110 clusters were formed. These clusters contained 541 vacant lots that were ultimately enrolled into the trial. This process guaranteed that no clusters overlapped, reducing later spillover and contamination effects across trial arms. Of these 541 vacant lots, 201 were assigned to the main vacant land intervention, 174 to the second, more basic vacant land intervention, and 166 to the no-intervention condition. "Any intervention" was defined as having had either the main vacant land intervention or the second, more basic vacant land intervention.

Randomly chosen cluster locations then served as the basis from which outcome data were collected. The outer-bounding polygon and its centroid were calculated for each grouping of vacant lots per cluster. This centroid represented the point location that was mathematically closest to all of the study vacant lots in each cluster. The address of the closest building to this point location was then determined as the starting point for house-to-house random sampling and enrollment of survey participants.

At each starting address, a two-person survey team walked in a predetermined, randomly chosen direction and path on the city block of the address and then on randomly chosen, adjacent city blocks within the cluster, until a total of five participants had been identified, consented to the study, and interviewed. Full and valid informed consent was obtained from all participants as reviewed and specified by both the University of Pennsylvania and the University of California, Los Angeles Institutional Review

Boards. Only one participant per household was chosen; in households with multiple eligible participants, the individual with the most recent birthday was chosen. Both English and Spanish speaking individuals 18 y and older were administered the survey by the two-person team in the language of their choice. Only two Spanish language surveys were administered.

Households with individuals that refused to participate or did not qualify to participate in the survey were marked as such and surveyors moved on to the next closest household. Upon completion of a first in-person interview, participants were asked to then participate in three subsequent in-person interviews. The first two, preintervention interview waves were conducted in the 18 mo from October 2011 to March 2013 and the last two postintervention interviews were conducted in the 18 mo from June 2013 to November 2014. Each participant was compensated $25 per interview and surveys took an average of 39.6 min to complete. Based on standard formulae, our survey response rate was 47.4% (6). Our participant response rate matched or exceeded that of other surveys and was high enough to produce a reasonably representative sample of our target population (7–9).

A total of 445 participants were interviewed during the preintervention period and 343 of these original participants were interviewed during the postintervention period. This amounted to a 22.9% loss-to-follow up; 78.4% of these 102 lost participants moved and 21.6% refused to participate in subsequent waves. All 102 participants that were lost to follow-up were replaced with additional randomly selected individuals living in their same cluster so that a total of 445 survey participants were ultimately analyzed. Of these 445 participants, 148 were assigned to the main vacant land intervention, 147 to the second, more basic vacant land intervention, and 150 to the no-intervention condition.

We determined sample size taking into account anticipated intracluster correlation, participant response prevalence, and number of events, effect size, and power. We calculated the minimally detectable effect size given 80% power for the participant-level outcomes and four time points based on the group pre- vs. postinteraction test for any pairwise comparison among the randomly allocated groups of lots. The minimally detectable effect size was the smallest Cohen's effect size (group pre- vs. postinteraction/SD of outcome) that was significant with 80% power under the following assumptions (10): within-participant correlation (ρ_y) for participant-level outcomes = 0.70; within-lot correlation (ρ_y) for participant-level outcomes = 0.20; within-lot correlation (ρ_x) for the −1,1 dummy variables for group and pre/post indicator variables = −0.33; and an $\alpha = 0.05$. Given these assumptions, we computed a minimally detectable effect size of 0.50 under a nested random effects model to account for the within-lot and within-participant correlations. This is a medium effect size based on Cohen (11). From this, and predicting a 25% loss-to-follow up rate, we estimated that we would maintain >80% power if we randomly surveyed three people per cluster for 35 vacant lot clusters per trial arm, twice before and twice after the intervention.

Ethnographic Microneighborhood Sampling. The ethnographic team collected preliminary data in all relevant neighborhoods of Philadelphia and then selected two working-class microneighborhoods heavily impacted by vacant lots and crime, but with dramatically different relationships to Philadelphia's changing real estate market during the trial years to understand and explore community responses to the greening of vacant lots. The precise boundaries of these socially identifiable microneighborhoods proceeded along well-known urban infrastructural features, including railroad tracks, highways, elevated train lines, major avenues with bus routes, parks, and public use buildings.

The first microneighborhood was 312 square blocks and the poorer of the two. It was located in the heart of the region's decaying industrial factory district and several of its census tracts were among the poorest in the city. It was unaffected by the rise in property values that was occurring unevenly in the rest of the city. The vast majority of its population was Puerto Rican, with a significant presence of African Americans. The second microneighborhood was 328 square blocks located in a part of the city that was under more rapid development. Property values were increasing, new construction was occurring, and abandoned housing was being renovated. The majority of its population was African American, with a growing presence of White and Asian residents.

Interventions and Outcome Measures. The purpose of this trial was to determine the effect of vacant land restoration on violence and crime outcomes in the nearby areas of surrounding clusters, as well as on perceptions of fear and safety outcomes among participants who lived in surrounding clusters. A main "cleaning and greening" intervention of vacant lots was completed via standard, reproducible processes implemented by well-coordinated teams of landscape contractors, many of whom came from local urban neighborhoods. The Pennsylvania Horticultural Society and the Philadelphia Office of Housing and Community Development designed and coordinated these interventions with cost-savings and rapid implementation in mind. This main vacant lot intervention specifically involved removing trash and debris, grading the land, planting new grass using an economical hydroseeding method that can quickly cover large areas of land by spraying a slurry mixture of seed and mulch, planting a small number of trees to create a park-like setting, and installing low wooden perimeter fences to show that the lot was cared for and to deter illegal dumping. A second, more basic vacant lot intervention involved removing trash and debris and mowing existing grass on each lot. Both interventions were performed by the Pennsylvania Horticultural Society and included regular monthly maintenance of treated lots. All vacant lot interventions occurred over a 2-mo springtime period, from April to May 2013, to maximize the probability of survival for newly planted vegetation (Fig. 1).

All interventions were randomly assigned and applied at the cluster level. All vacant lots within a cluster received one of the two interventions, or no intervention as a control condition, allowing us to test the effects of the main intervention and any intervention on the clusters. Because individual participants lived within the 0.4-km radius clusters, we were also able to test the effects of the interventions on them. At the end of the postintervention period in November 2014, the vacant lots that were randomly assigned to the no-intervention group during the trial were also scheduled for cleaning and greening.

Both police-reported outcome measures and participant-level outcome measures were collected and analyzed in and around each cluster. Violence and crime data were collected from the Philadelphia Police Department and aggregated by month for 18 preintervention months and 18 postintervention months, for a total of 36 observation periods. These data included the dates and address locations of six outcomes: gun assaults, burglaries, robberies and thefts, narcotics possession, sales and trafficking, and nuisances. Nuisances were defined as the summation of curfew violations, disorderly conduct, public drunkenness, illegal dumping, loitering, noise violations, prostitution, and vandalism. An additional variable of all crimes was also calculated as the sum of the six measures of gun assault, burglaries, robberies, thefts, narcotics possession, and narcotics sales. The address location of each police-reported event was geographically assigned to a point-in-space and a kernel density estimate was used to calculate events per square kilometer for all outcomes at the centroid point of each vacant lot. A kernel density gradient is a smoothed surface of values (such as number of crimes per square kilometer), given a set input of values (incidence of crimes) at specific latitude–longitude locations. Using a spatial bandwidth (or search area), the individual values are summed to create a smoothed probability density surface of observed phenomenon that can then be used to estimate density at any point within the space (12, 13).

Perceptions of violence, crime, nuisances, and fear for one's safety were surveyed from participants. The same questions were asked to all participants across all four waves of the survey. Participants were asked to focus their responses to their experiences within the past 30 d to avoid telescoping and overestimation by participants. The following survey questions were analyzed for changes from the pre- to the postintervention period: (*i*) "There is a lot of crime in my neighborhood"; (*ii*) "There is too much drug use in my neighborhood"; (*iii*) "Vandalism is common in my neighborhood"; (*iv*) "In my neighborhood, people watch out for each other"; (*v*) "Did you not go someplace in your neighborhood during the day because you felt you would not be safe?"; and (*vi*) "I spent time hanging out, relaxing, or socializing on porches, stoops or front yards in my neighborhood." Similar survey items have been successfully used in prior studies (14, 15).

Documentation of the changing conditions in all study vacant lots was also recorded in the pre- and postintervention periods. Teams of individuals who were independent of the household interview teams took field video footage of study vacant lots and downloaded Google Street View images of study vacant lots over time (16, 17). These video and street-view images were then graded using a 1–10 scale of orderliness, with 1 being no disorder and 10 being high order. This scale was averaged and then differenced between the pre- and the postintervention periods to grade each vacant lot in terms of whether it had: (*i*) deteriorated (a negative pre/post scale difference), (*ii*) experienced minor improvement (a positive pre/post scale difference of up to 5), or (*iii*) experienced a major improvement (a positive pre/post scale difference of 5 or greater). These three categories were separated using tertile breaks.

Random Allocation and Blinding of Interventions. All 110 clusters were stratified within the four city sections and then assigned computer-generated random numbers. The clusters within each city section were then randomly allocated to the main vacant lot intervention A ($n = 37$ clusters), the second, more basic vacant lot intervention B ($n = 36$ clusters), or the no-intervention ($n = 37$ clusters) arms of the trial using a repeat randomization procedure (18). The repeat randomization procedure functioned under a predetermined protocol agreed upon by the study team. This protocol used repeated random allocation of the three study arms to ensure that no variable was different by more than a level that would occur by chance ($P < 0.05$). These variables were: total area, mean separating distance, and outer bounding polygon area of the study vacant lots in each cluster, as well as the total vacant lots, resident population, and number of serious crimes (violent and property crimes) in each cluster.

Only the study principal investigator (PI) had access to the randomization codes and the final random assignment of each cluster into one of the three trial arms. Randomization codes were securely filed in electronic format and inaccessible to maintain blinding of the other members of the study team, the field interviewers and staff, the contractors implementing of the different interventions, and the study participants. Contractors were given only the addresses of the vacant lots that were in each of the two intervention groups and instructed as to which intervention should be performed over the 2-mo intervention period. Field interviewers were only given street addresses from which to begin their household interviews, with no mention of vacant lots in the surrounding neighborhoods. Study participants were told that they were responding to a survey about urban health and their local environments, with no mention of specific vacant lots in their neighborhoods.

Our ethnographic team's documentation of logistical processes, protocol fidelity, and community context noted the importance of clearly explaining to field staff the scientific principle of random sampling in the early phase of the project when lots were being selected for eligibility. The field staff from our community-based and municipal partners were also reminded to disregard former priorities for selecting lots based on prior municipal contracts that had included, among other criteria, enhancing commercial corridors and school zones (19).

Statistical Methods and Analyses. The units of analysis for the violence and crime outcomes were a balanced panel of 541 vacant lots with monthly observations measured over 38 mo. The units of analysis for the perceptions of violence, crime, and safety outcomes were a balanced panel of 445 survey participants with observations taken in four survey waves, during pre- and postintervention periods, over 38 mo. Baseline individual and cluster-level variables were also inspected for balance between the randomly allocated arms of the trial.

ITT analyses of vacant lots and survey participants were conducted according to the intervention group to which they had been randomly allocated. Some vacant lots that were randomly assigned to the intervention groups for improvement actually deteriorated and some that were assigned to the no-intervention group naturally improved instead of deteriorating over the course of the study. Using the change in orderliness gradations that had been calculated for each vacant lot over time, CA-ITT analyses were also completed in accounting for the level of improvement that had actually occurred in each lot, regardless of its random assignment. These CA-ITT analyses used two-stage instrumental variables regressions with random treatment assignment as the instrument (as it was orthogonal to the outcomes studied) and provided complementary information in terms of adjusting for treatment nonadherence and avoiding as-treated and per protocol analytic biases (20–22).

Pairwise comparisons were completed for all study outcomes between the main intervention group and the no-intervention group, as well as the any-intervention group and the no-intervention group. These pairwise comparisons were tested for statistical significance (defined as $P < 0.05$) using random effects, cross-sectional time series regressions that accounted for the cluster design of the study. Regressions accounted for the clusters and month fixed effects for police-reported outcomes and wave fixed effects for participant-reported outcomes. All statistical analyses were conducted using Stata 14.1. Analytic data, associated protocols, analytic programming code, and material descriptions are available upon email request to C.C.B.

Difference-in-differences were calculated as interaction terms of 1–0 intervention-control differences multiplied by 0–1 pre/post differences. These difference-in-differences interaction terms were the primary independent variables of interest interpreted as the true effect of the interventions on the various outcomes studied. In addition, using the previously fit regression models, marginal effects where the difference-in-differences β-coefficients = 1 and 0 were also estimated. These marginal effects were differenced to obtain absolute magnitudes of reduction for each outcome in the post-period. Absolute magnitudes of reduction were then divided by the total magnitude of occurrence for each outcome in the postperiod to obtain percentage reductions (13, 23, 24).

Additional subset analyses of all outcomes were also completed using the poverty level for Philadelphia in 2010. Pairwise statistical tests of the intervention conditions versus the no-intervention condition were then completed within neighborhood subsets below the poverty level.

Displacement analyses were also completed for the crime outcomes. Crime events were counted within a 0.1-km radius of each vacant lot and then between 0.1-km and 0.2-km distance from each vacant lot. Similar counts were obtained within a 0.2-km radius of each vacant lot and then between 0.2-km and 0.4-km distance from each vacant lot. This permitted "donut-hole vs. donut" spillover tests of the effect that the interventions were having to be conducted at two different spatial scales (25–28).

Ethnographic Data Archiving and Analyses. Audio interviews were reviewed and relevant time-coded excerpts transcribed. All field notes were written as electronic files. Social media and geographic data were reviewed online. Online tools, including Google Street View and Everyblock.com were used to archive, date, monitor, and visually explore community contexts, lines of sight, and changing crime statistics and infrastructure in the areas surrounding vacant lots. We bookmarked and notated online screenshots of critical web-based data for digital archiving and ongoing analysis. NVivo qualitative data analysis software (v11, QSR International) was used to code and retrieve passages relevant to physical infrastructure, neighborhood attitudes, crime, violence, race relations, law enforcement, and drug trafficking.

1. Campbell MK, Piaggio G, Elbourne DR, Altman DG; CONSORT Group (2012) Consort 2010 statement: Extension to cluster randomised trials. *BMJ* 345:e5661.
2. Messac L, Ciccarone D, Draine J, Bourgois P (2013) The good-enough science-and-politics of anthropological collaboration with evidence-based clinical research: Four ethnographic case studies. *Soc Sci Med* 99:176–186.
3. Lopez AM, et al. (2013) Interdisciplinary mixed methods research with structurally vulnerable populations: Case studies of injection drug users in San Francisco. *Int J Drug Policy* 24:101–109.
4. Bourgois P, et al. (2006) Reinterpreting ethnic patterns among white and African American men who inject heroin: A social science of medicine approach. *PLoS Med* 3:e452.
5. Boruch R, et al. (2004) Estimating the effects of interventions that are deployed in many places: Place-randomized trials. *Am Behav Sci* 47:608–633.
6. AmericanAssociationforPublicOpinion Research (2015) AAPOR response rate calculator. Available at https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx. Accessed November 13, 2015.
7. Groves R (2006) Nonresponse rates and nonresponse bias in household surveys. *Public Opin Q* 70:646–675.
8. Galea S, Tracy M (2007) Participation rates in epidemiologic studies. *Ann Epidemiol* 17:643–653.
9. Keeter S, Kennedy C, Dimock M, Best J, Craighill P (2006) Gauging the impact of growing nonresponse on estimates from a National RDD Telephone Survey. *Public Opin Q* 70:759–779.
10. Neuhaus JM, Segal MR (1993) Design effects for binary regression models fitted to dependent data. *Stat Med* 12:1259–1268.
11. Cohen J (1988) *Statistical Power for the Behavioral Sciences* (Erlbaum, Hillsdale, NJ), 2nd Ed.
12. Branas CC, Elliott MR, Richmond TS, Culhane DP, Wiebe DJ (2009) Alcohol consumption, alcohol outlets, and the risk of being assaulted with a gun. *Alcohol Clin Exp Res* 33:906–915.
13. Branas CC, et al. (2011) A difference-in-differences analysis of health, safety, and greening vacant urban space. *Am J Epidemiol* 174:1296–1306.
14. Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277:918–924.
15. Skogan WG (1990) *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods* (Univ California Press, Oakland, CA).
16. Cannuscio CC, et al. (2009) Visual epidemiology: Photographs as tools for probing street-level etiologies. *Soc Sci Med* 69:553–564.
17. Odgers CL, Caspi A, Bates CJ, Sampson RJ, Moffitt TE (2012) Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *J Child Psychol Psychiatry* 53:1009–1017.
18. Schulz KF, Grimes DA (2002) Generation of allocation sequences in randomised trials: Chance, not choice. *Lancet* 359:515–519.
19. Heckert M, Mennis J (2012) The economic impact of greening urban vacant land: A spatial difference-in-differences analysis. *Environ Plann Part A* 44:3010–3027.
20. Sommer A, Zeger SL (1991) On estimating efficacy from clinical trials. *Stat Med* 10: 45–52.
21. Brookhart MA, Rassen JA, Schneeweiss S (2010) Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf* 19: 537–554.
22. Sussman JB, Hayward RA (2010) An IV for the RCT: Using instrumental variables to adjust for treatment contamination in randomised controlled trials. *BMJ* 340:c2073.
23. Meyer B (1995) Natural and quasi-experiments in economics. *J Bus Econ Stat* 12: 151–161.
24. Branas CC, et al. (2016) Urban blight remediation as a cost-beneficial solution to firearm violence. *Am J Public Health* 106:2158–2164.
25. Kondo MC, Keene D, Hohl BC, MacDonald JM, Branas CC (2015) A difference-in-differences study of the effects of a new abandoned building remediation strategy on safety. *PLoS One* 10:e0129582.
26. Kondo M, Hohl B, Han S, Branas C (2016) Effects of greening and community reuse of vacant lots on crime. *Urban Stud* 53:3279–3295.
27. Han S, Branas CC, MacDonald JM (2016) The effect of a Sunday liquor-sales ban repeal on crime: A triple-difference analysis. *Alcohol Clin Exp Res* 40:1111–1121.
28. Weisburd D, et al. (2006) Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits. *Criminology* 44: 549–592.

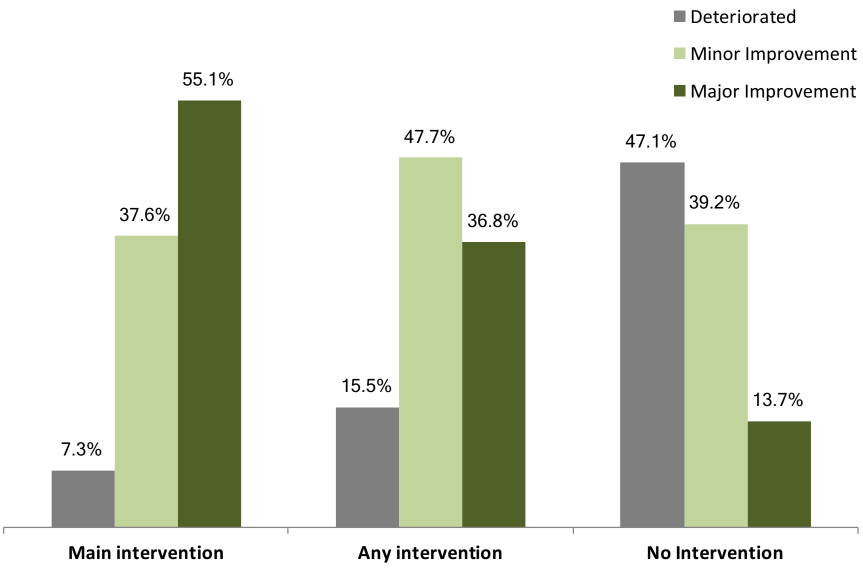

Fig. S1. Actual shifts in conditions of vacant lots despite random assignment to interventions and control.

Table S1. Baseline, preperiod balance among key participant and cluster-level variables (within a quarter-mile radius) compared as means between trial intervention conditions and a no-intervention control condition

Participant- or cluster-level means	Main intervention	Any intervention	No intervention
Participant-level means			
Age, y	44.1	45.2	45.9
Tenure in home, y	12.7	13.7	13.5
Female, %	63.2	63.8	59.2
Hispanic, %	9.7	9.1	11.2
Black, %	78.6	79.0	69.1
Any college, %	25.7	24.6	29.7
Unemployed, %	35.9	30.2	25.4
Family income <$25K, %	46.6	48.2	42.6
Cluster-level means			
Resident population	288 People	293 People	285 People
Serious crimes	16.5 Crimes	17.4 Crimes	17.1 Crimes
Serious crime rate (per 100,000)	5,729	5,939	6,000
Total eligible vacant lots	38.3 Lots	40.7 Lots	38.1 Lots
Prior treated lots	6.7 Lots	6.0 Lots	5.6 Lots
Study lots per cluster	5.4 Lots	5.1 Lots	4.5 Lots
Study lots total area per cluster, ft^2	4,844	4,890	4,872
Study lots mean separation, ft	75.6	73.5	73.5
Study lots bounding polygon area, ft^2	10,110	10,996	10,649