

## Webappendix

### Propensity Score Matching to define control districts

We identified control neighborhoods that are highly similar to the deprived neighborhoods in which the Dutch District Approach was carried out, by means of propensity score matching. Propensity score matching (PSM) is a statistical method to select similar control areas for intervention areas that are the subject of an observational study in which the researchers do not have any control over the assignment of the intervention (as opposed to a randomized experiment). By using propensity score matching for selection of control areas as well as adjusting statistical analyses for remaining confounders, the comparability between intervention and control areas is maximized, with exception of the intervention.

The propensity score indicates the degree of comparability between the possible control area and the intervention area with regard to the problems that are targeted by the interventions (Dehejia & Wahba, 2002; Kawachi & Subramanian, 2007; Oakes & Johnson, 2006). By calculating a propensity score, we predict the chance to have been selected for the intervention for each neighborhood in the dataset using multiple predictor variables. The propensity score is hence not predicting the effect of the intervention (in our case, health and health related behavior), but the score predicts the likelihood that the intervention would have been implemented in that neighborhood. The most important predictor for the implementation of the Dutch District Approach is the extent of problems regarding employment, education, living circumstances, social integration, and safety.

Neighborhoods that are very similar to the intervention neighborhoods with regard to this problems will have similar propensity scores. The basis of the matching will be the range of propensity scores of the group of intervention neighborhoods and not the 'individual' neighborhood level score. We did vary the degree of comparability between control areas and intervention areas by selecting different groups of control areas, using different ranges of propensity scores.

We used propensity score matching to select control neighborhoods that are similar to the intervention neighborhoods with regard to the neighborhood environment at the time of the start of the Dutch District Approach. Comparability with regard to demographics is guaranteed by including these socio-demographic variables as confounders in the statistical analyses. These variables, such as age, gender, and socio-economic indicators are therefore not included in the calculation of the propensity scores of neighborhoods. The calculation of the propensity score were based on variables indicating housing circumstances (housing conditions and physical neighborhood characteristics), social integration (social neighborhood characteristics), and safety. We used three different datasets to calculate propensity scores, i.e. the data that were used by the national government to select the 40 deprived neighborhoods that were included in the Dutch District Approach (VROM Syswov & WoON), the datasets used by the National Bureau of Statistics to monitor trends in outcomes of the District Approach (VROM Syswov and GBA), and the national Livability Monitor (Leefbarometer) (also see Table 1). We use the following variables, measured in the beginning of the year 2008 or as close to that date as possible (Table 1);

- Housing conditions
  - % small houses ( $\leq 3$  rooms) in the neighborhood
  - % old houses (built  $\leq 1970$ )
  - % cheap houses (social housing)
  - % residents that reported to be satisfied with their own home

- % apartments in the neighborhood
- average number of persons per address
- % of houses that have been demolished
- Safety and social neighborhood environment
  - % residents that reports nuisance from direct neighbors
  - % residents that reports nuisance from other neighborhood residents
  - % residents that is afraid to be harassed or robbed in their neighborhood
  - Public order disturbances, violent crimes, theft from cars, and nuisance
- Physical neighborhood environment
  - % residents that reports graffiti in the neighborhood
  - % residents that reports vandalism in the neighborhood (demolition of telephone booths, bus or tram stops, etc.)
  - % respondents that reported to be satisfied with their living environment
  - Vandalism

Table 1. Data sources for variables used to calculate the propensity score in order to select control neighborhoods, measured on January 1<sup>st</sup>, 2008 and available for all postal code areas in the Netherlands

<b>Variables</b>	<b>Before used for</b>	<b>Dataset</b>	<b>Obtained thru</b>
<b><i>Housing conditions</i></b>			
% small houses ( $\leq 3$ rooms)	Selection deprived neighborhoods for District Approach & Outcome Monitor District Approach	VROM Syswov 2008	Ministry of the Interior and Kingdom Relations
% apartments	Outcome Monitor District Approach	VROM Syswov 2008	Ministry of the Interior and Kingdom Relations
% old houses (built $\leq 1970$ )	Selection deprived neighborhoods for District Approach	VROM Syswov 2008	Ministry of the Interior and Kingdom Relations
% cheap houses (social housing)	Selection deprived neighborhoods for District Approach & Outcome Monitor District Approach	VROM Syswov 2008	Ministry of the Interior and Kingdom Relations
% residents satisfied with their own home	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
average number of persons per address	Outcome Monitor District Approach	GBA 2008	Not available
% houses demolished	Leefbarometer (dimensie publieke ruimte)	VROM Syswov 2008	Ministry of the Interior and Kingdom Relations
<b><i>Safety and social neighborhood environment</i></b>			
% residents that reports nuisance from direct neighbors	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
% residents that reports nuisance from other neighborhood residents	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS

Variables	Before used for	Dataset	Obtained thru
% residents that is afraid to be harassed or robbed	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
public order disturbances	Leefbarometer (composite safety indicator*)	Leefbaarometer	Ministry of the Interior and Kingdom Relations
violent crimes	Leefbarometer (composite safety indicator*)	Leefbaarometer	Ministry of the Interior and Kingdom Relations
theft from cars	Leefbarometer (composite safety indicator*)	Leefbaarometer	Ministry of the Interior and Kingdom Relations
nuisance	Leefbarometer (composite safety indicator*)	Leefbaarometer	Ministry of the Interior and Kingdom Relations
% residents that reported graffiti	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
% residents that reported vandalism	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
% residents satisfied with living environment	Selection deprived neighborhoods for District Approach	WOoN 2006	DANS
vandalism	Leefbarometer (composite safety indicator*)	Leefbaarometer	Ministry of the Interior and Kingdom Relations

\*Only the composite safety indicator was available to us. This one safety indicator summarized the data on safety issues, such as public order disturbances, violent crimes, theft from cars, nuisance, vandalism, etc.

We base the selection of control areas on the propensity scores calculated, using a narrow and a wide definition;

- The narrow definition selects control neighborhoods that fall within the same range of propensity scores than the 83 postal code areas that constitute the deprived intervention neighborhoods, that is 0.06170 to 0.999. There are 78 districts with a propensity score  $\geq 0.06170$  that are not intervention neighborhoods.
- The wide definition selects neighborhoods that belong to the 10% of all areas with the highest propensity scores (= 10% with the highest odds to have similar problems as the intervention deprived neighborhoods), that is propensity score  $\geq 0.00036$ . This selection produces 321 control neighborhoods.

Of those 78 control neighborhoods that have been selected according the narrow definition, 59 are situated in the same cities as the intervention deprived neighborhoods. Three of the 18 cities that have one or more intervention neighborhood are however not included in this list. This means that in these three cities there is no other neighborhood that resembles the intervention neighborhood with regard to housing conditions, the physical and social neighborhood environment, and safety. This group of control neighborhoods (selected using the narrow definition and confined to the same municipalities than the intervention neighborhoods) is therefore not suitable as a control group for the evaluation.

Of the 321 control neighborhoods that have been selected using the wide definition, 181 are located in the same cities as the intervention neighborhoods. All 18 cities with one or more intervention

neighborhood are represented in this selection. This group of 181 control neighborhoods will therefore be added to the selection of groups of control neighborhoods for our evaluation.

## **References**

Dehejia RH, Wahba S. Propensity score-matching methods for nonexperimental causal studies. *The review of economics and statistics* 2002;84(1):151-161.

Kawachi I, Subramanian SV. Neighbourhood influences on health. Outstanding issues in the neighbourhood research agenda. *Journal of Epidemiology and Community Health* 2007;61:3-4.

Oakes JM, Johnson PJ. Propensity score matching for social epidemiology. In; Oakes JM, Kaufman JS. *Methods in social epidemiology*. New York; Jossey-Bass, 2006.