Positioning Paper

Peer Reviewed

Cost-effective methods for evaluation of neighbourhood renewal programs

authored by

Gavin Wood and Melek Cigdem

for the

Australian Housing and Urban Research Institute

RMIT Research Institute

September 2012

AHURI Positioning Paper No. 151

ISSN: 1834-9250

ISBN: 978-1-922075-11-6

AHURI

Australian Housing and Urban Research Institute

Authors	Wood, Gavin	RMIT University
	Cigdem, Melek	RMIT University
Title	Cost-effective methods for evaluation of neighbourhood renewal programs	
ISBN	978-1-922075-11-6	
Format	PDF	
Key words	Neigbourhood renewal, residential property value, price impact, spillover, benefits	
Editor	Anne Badenhorst	AHURI National Office
Publisher	Australian Housing and Urban Research Institute Melbourne, Australia	
Series	AHURI Positioning Paper, no. 151	
ISSN	1834-9250	
Preferred citation	Wood, G. and Cigdem, M. (2012) *Cost-effective methods for evaluation of neighbourhood renewal programs,* AHURI Positioning Paper No.151. Melbourne: Australian Housing and Urban Research Institute.	

ACKNOWLEDGEMENTS

This material was produced with funding from the Australian Government and the Australian states and territory governments. AHURI Limited gratefully acknowledges the financial and other support it has received from these governments, without which this work would not have been possible.

AHURI comprises a network of universities clustered into Research Centres across Australia. Research Centre contributions, both financial and in-kind, have made the completion of this report possible.

The provision of the Property Valuations (2008) dataset and the Property Transactions dataset by the Office of the Valuer-General, Victoria (VGV) is also gratefully acknowledged. VGV is a division within the Department of Sustainability and Environment. It is the state's independent authority on property valuations, and oversees government property valuations and council rating valuations.

AHURI POSITIONING PAPER SERIES

AHURI Positioning Papers is a refereed series presenting the preliminary findings of original research to a diverse readership of policy-makers, researchers and practitioners.

PEER REVIEW STATEMENT

An objective assessment of all reports published in the AHURI Final Report Series by carefully selected experts in the field ensures that material of the highest quality is published. The AHURI Final Report Series employs a double-blind peer review of the full Final Report – where anonymity is strictly observed between authors and referees.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AHURI	Australian Housing and Urban Research Institute Limited
AITS	Adjusted Interrupted Time Series
ASGS	Australian Statistical Geography Standard
CBD	Central Business District
CDs	Collection Districts
DID	Difference-in-Differences
HUD	Department of Housing and Urban Development (USA)
ICSEA	Index of Community Socio-Educational Advantage
NiB	Neighbourhoods in Bloom
NR	Neighbourhood Renewal
SLA	Statistical Local Area
UR	Urban Renewal
US	United States
VGV	Valuer-General Victoria

EXECUTIVE SUMMARY

In Australia and several other Western countries, neighbourhood and urban renewal schemes have assumed increasing prominence in state and federal government policy agendas. This reflects growing concerns about the gap between disadvantaged neighbourhoods and middle-to-high income neighbourhoods. One such initiative is Victoria's Neighbourhood Renewal (NR) program led by the Housing and Community Building Division in the Department of Human Services. Introduced by the state government in 2001, the objective of NR is to tackle disadvantage in neighbourhood areas with concentrations of public housing and to address the problem of poverty and social exclusion. The program develops an integrated approach by combining the resources of government, residents, the community sector and local businesses. Currently, there are 21 established projects in metropolitan and regional Victoria which have been progressively launched since 2002 and selected for renewal on the basis of numerous key indicators of disadvantage.

This Positioning Paper is the first output of a research project which evaluates the impact of the NR scheme on Melbourne's neighbourhoods. Impacts are measured using property transactions data to compare changes in property values within or proximate to a renewal site (treatment group) with changes in comparable property values outside the site (control group). The idea behind this approach has the premise that renewal programs typically aim to improve the physical appearance of properties and communal facilities, while also investing in community services that make neighbourhoods more attractive, and strengthen the human and social capital of residents. If successful in meeting these goals, the demand for housing in and around neighbourhood renewal sites will increase, and house prices rise by more than would otherwise be the case. A critical aim of the research program is estimation of this neighbourhood renewal related price premium.

Research questions

There are four main research questions that we propose to investigate in the Project:

1. What are the direct effects of the NR program on residential property values that are transacted within the boundaries of renewal sites?
2. What are the indirect effects of the program on the property values of residential properties that border on or are adjacent to renewal sites?
3. Do the indirect (spillover) benefits from NR projects rapidly decay as distance from their boundaries increase?
4. What are the aggregate costs and benefits of the NR program?

To answer the first research question, we examine the degree to which NR programs impact on the prices of privately owned properties that are located within the boundaries of NR sites. Because these properties are located within areas that are designated as NR sites, they have an uninterrupted exposure to the program, and can therefore be used to estimate the direct effects of NR. The sample used to investigate the second and third research questions includes properties that lie outside of but adjacent to an area designated for NR. Drawing on US studies we choose a buffer that is within 2000 feet of the boundaries of NR sites. The properties located within this buffer are then used to determine whether program impacts 'spill over' to nearby properties (also referred to as indirect effects). Finally, in question four, we will use estimates of the price premiums associated with NR to aggregate dollar measures of returns/benefit.

Data sources

The analysis draws on two main datasets obtained from the Office of the Victorian Valuer-General (VG), namely the Valuation and the Property transactions datasets. The valuations

data is a cross-sectional dataset that is the key source of detailed property-level and neighbourhood characteristics as at year 2008. It is confined to land and buildings in metropolitan Melbourne. Variables that are contained in the dataset include characteristics such as number of bedrooms, distance from various amenities (e.g. CBD, train stations, schools & major activity centres) and local zoning codes and overlays.

The second data source is the property transactions dataset which is a repeated cross-sectional time series dataset containing transaction records for every property that has been sold over the years 1990–2011. Information in the transactions dataset includes property address, sales date and price. It does not contain any property-level characteristic information. The transactions dataset is merged with the valuation dataset. This means that sales records are matched with their corresponding property-level information such as property characteristics, location in relation to amenities, principal and major activity centres (areas designated by planning authorities as focal points for employment growth, transport nodes and urban amenities), and planning regulations such as zoning and overlay areas, as sourced from valuation records. The merged dataset is then imported into a GIS environment via the latitude and longitude fields to identify all properties lying within the boundaries of an NR site (to carry out research Question 1), and adjacent to the site (to carry out research Questions 2 & 3).

Methodological approach

To address the main research questions, we will exploit quasi-experimental techniques that have been used extensively in the US to measure the impacts of urban renewal (Briggs et al. 1999; Ellen et al. 2002). Specifically, we will employ the 'difference-in-differences' (DID) method to compare the pre and post-NR price levels and trends in NR sites with those of comparable properties that lie outside of NR site boundaries. This approach yields a dollar-value estimate of the NR program's benefits in the form of house price premiums. The validity of estimates hinges on the assumption that changes in the property prices of control properties represents a reliable measure of the counterfactual change in NR sites' property prices that would have eventuated in the absence of renewal programs.

To construct suitable control groups, two separate quasi-experimental methods will be employed:

- → the near neighbour approach
- → the 'revealed preferences method' (Cigdem 2012).

The near neighbour approach creates a control group using properties that belong to the same Statistical Local Area (SLA) as the NR site, but outside its boundaries. The revealed preference method uses residential properties in locations that will become a renewal site at some time in the future (but transacted before NR implementation) as controls for properties that are already exposed to the program. Matching with the use of propensity scores justifies the comparability of the control group with the treatment group by ensuring that the pre-program characteristics of the treatment and control housing samples are as similar as possible (Dehejia & Wahba 2002; Black & Smith 2003).

Sample numbers

The quasi-experimental method assigns house transaction records into one of two groups: the treatment group, which includes transaction records for properties located within the boundaries of NR sites introduced from 2001–06; and the baseline control group comprising property transactions that are located outside of NR site boundaries but within the SLAs that NR sites populate. There are close to 141 600 transaction records in the sample dataset with 13 527 observations in the treatment group and 128 041 observations in the baseline control group. Mean prices for the treatment sample are around 30 per cent lower than mean prices in the baseline control sample, as is to be expected since NR sites have

concentrations of public housing; consequently private housing in the NR site is likely to sell at a discount.

Preliminary findings

Because the NR scheme was introduced in several stages over the sample period, we estimate three separate difference-in-difference models. The first estimate is based on NR sites that were introduced in 2002. The treatment group in this case is properties that are located within the boundary of the two NR sites that were launched in 2002, and the control group contains properties that lie outside the NR boundaries but within the same SLAs that these NR sites are located in. Similarly, for the second and third models, the treatment group comprises property transactions within NR sites that were introduced in 2003 and 2006, respectively, and the control sample is formed from the relevant SLAs. If the expenditures on public housing and amenities (e.g. parks) in NR areas lift their capital values by an equal amount, our estimates of house price premiums can be regarded as measures of net benefits or net returns to investment outlays.

The regression results endorse the view that NR is indeed responsible for net benefits. While properties located within NR boundaries sold for between 18 per cent and 40 per cent less than properties in the control sample, this differential narrowed by between 1 per cent and 4 per cent in the period following the introduction of NR. At the median price of $144 000 (in the NR sites) these estimates translate into price premiums of between $1469–$5184. These results should be regarded with some caution however, as they are naïve models based on a crude version of the control sample. The second stage of this project will refine this approach by using the propensity score method to design a more robust matching sample of control transactions.

1 INTRODUCTION

1.1 Motivation and aims of the project

In the Australian context, renewal of public housing estates has assumed nation-wide prominence on policy agendas as state and federal administrations have become increasingly aware of problems in the public housing sector over the past 20 years (Randolph & Judd 2006). With poorer neighbourhoods hosting larger concentrations of disadvantaged households, the widening gap between disadvantaged neighbourhoods and middle–high-income neighbourhoods has manifested itself in the form of fewer employment opportunities in the poorer neighbourhoods, alongside widespread health problems, family breakdown, lower educational achievement, increasing use of drugs, and associated crime and social stigma (Klein 2004). In response to these problems, state governments have intervened to narrow this neighbourhood divide, with a substantial amount of public and private funds being allocated to regenerating disadvantaged Australian public housing neighbourhoods (Hughes 2004).

In Victoria, one such program is Neighbourhood Renewal (NR), which was introduced by the state government in 2001 as a place-based response focusing on the regeneration of marginalised communities. The objective of NR is to tackle disadvantage in neighbourhood areas with concentrations of public housing and to address the problem of poverty and social exclusion. Led by the Housing and Community Building Division in the Department of Human Services, the program develops an integrated approach by combining the resources of government, residents, the community sector and local businesses.

NR projects have been progressively launched since 2002. Currently, there are 21 established projects in metropolitan and regional Victoria selected for renewal on the basis of numerous key indicators of disadvantage. All project sites have a clearly defined geographic boundary and are characterised by their relatively large concentrations of public housing[1].

Research studies evaluating the renewal program's effectiveness in meeting these objectives have predominantly been qualitative in nature (Shield, Graham & Taket 2011; Kelaher, Warr & Tacticos 2010; Klein 2004; Department of Human Services 2005, 2008). While qualitative methods generally provide a richer understanding of the social and behavioural effects of renewal programs as perceived by program recipients and nearby communities, they are more costly to implement than quantitative approaches, and often suffer from small sample sizes that make it difficult to generalise findings. Also, qualitative methods do not offer 'value-for-money' dollar measures that can inform resource allocation decisions. In the USA there has been extensive quantitative research sponsored by the US Federal Department of Housing and Urban Development (HUD) offering financial measures of the impact of urban revitalisation policies (Santiago et al. 2001; Galster et al. 2004b; Castells 2010). These studies are based on a market failure interpretation of the rationale for such policies. Neighbourhood decline adversely affects all who live in the vicinity of concentrations of poverty, not just the poor. If intervention reverses decline there will be external benefits (reduced crime, improved health, increasing job opportunities etc.) that are reflected in house price premiums. Extracting precise estimates of these price premiums from the analysis of housing market transactions has been central to this innovative program of research.

[1] A more comprehensive discussion of Victoria's Neighbourhood Renewal Scheme will be provided in the Final Report as a context for interpretation of the findings. The Positioning Paper instead concentrates on the methods we have chosen to employ in the evaluation of Victoria's Neighbourhood Renewal Scheme.

The rationale for the use of house prices is straightforward; if the intervention succeeds in making an area more attractive to live in (whether for aesthetic or economic reasons) there will be a stronger demand for housing, and so (all else equal) house prices rise. It is not an ideal measure, because it does not reveal the source of the price premium e.g. has it come about because the program created employment opportunities and therefore in migration into the area as people seek the new jobs? Or is it attributable to fewer vacant and vandalised housing units as rehabilitation investments improve the 'look and feel' of the area and stronger demand for housing results?[2]

Quantitative-based evaluation studies of renewal programs in Australia have received very little attention in either academic or government circles. Randolph and Judd (2006, p.98) argue that 'while there is an emerging body of evaluation and research that has attempted to assess the outcomes of renewal programs and policies, it can be argued that there is still a relatively poor level of general understanding of what aspects of renewal are effective or what outcomes have actually been achieved'.

This is a particularly important time for place-based policies such as neighbourhood renewal as there are a growing number of sceptics. The US economist, Ed Glaeser, is in the vanguard of such thinking. In his view, governments should not try to stall urban change by seeking to reverse urban decline with place-based policies that often take the form of investment in buildings (e.g. housing construction & renovation) and the environment (e.g. parks). These projects invariably fail to help the poor people who live in declining neighbourhoods; since people leave distressed neighbourhoods those left behind tend to have an ample amount of housing and other buildings and do not need any more. According to this perspective declining neighbourhoods need more human capital. Helping poor people is far easier to justify and more effective than helping poor places. Rather than investing resources revitalising the built environments in poor neighbourhoods, policy should focus on giving disadvantaged people the skills they need to compete, wherever they choose to live, rather than encouraging them to stay in those locales where renewal programs are targeted (Glaeser 2011; Glaeser & Gottlieb 2008).

On the other hand, there is a case favouring neighbourhood renewal. In contemporary cities the poor can be concentrated (Wulff & Reynolds 2011) in enclaves with negative externalities such as crime, ill health and so on, the consequence; these social ills can spill over and spread to previously vibrant adjacent communities. The spillover effects can cause decline to set in on an even bigger scale. Interventions that succeed in quarantining or even reversing spillover effects can offer persuasive evidence favouring renewal interventions.

1.2 Research questions

This Project will be, to our knowledge, the first empirical peer-reviewed research study to offer a quantitative measure of the effectiveness of Victoria's Neighbourhood Renewal Scheme in fulfilling its main policy objectives. The project's appeal from a policy perspective is first; that it is a cost effective alternative to qualitative methods, which typically involve a far more costly process of performing surveys, interviews and conducting focus groups. Second, we provide 'value for money' estimates of renewal programs based on property price premiums, an approach that has never been used to date in Australia. Third, we utilise widely used techniques like the propensity score matching method, whose contribution in this field is recognised as methodologically robust.

There are four main research questions that we propose to address:

[2] However, it is possible to shed some light on such questions when renewal programs differ in terms of their emphasis on investments in physical (that is buildings) and human (skills & qualifications) capital.

1. What are the direct effects of the NR program on residential property values that are transacted within the boundaries of renewal sites?
2. What are the indirect effects of the program on the property values of residential properties that border on or are adjacent to renewal sites?
3. Do the indirect (spillover) benefits from NR projects rapidly decay as distance from their boundaries increase?
4. What are the aggregate costs and benefits of the NR program?

1.3 Methodological approach

To address the above questions, we compare changes in property values within or proximate to a renewal site (treatment group) to changes in comparable property values outside the site (control group). The validity of estimates hinges on two assumptions. First, the benefits of NR will be reflected in property price premiums. Second, the assumption that changes in the property values of control properties represents a reliable measure of the change in property values for treatment properties in the absence of renewal programs. In other words: Can the control properties offer a credible 'counterfactual' profile for the trajectory of house prices had renewal not been implemented?

This quasi-experimental approach has been used extensively in the US to measure the impacts of urban renewal (Briggs et al. 1999; Ellen et al. 2002) as well as labour market programs (Chapman 1993) and social policy interventions (see Gruber 2000). In a housing policy context, the literature regards the approach developed by Santiago et al. (2001) as the most methodologically sophisticated because it overcomes many of the shortcomings plaguing earlier studies, thereby setting a new standard for conducting quantitative housing research. In their study into the effects of dispersed public housing sites on nearby property values, Santiago et al. (2001) estimate a 'difference in differences' (DID) model to compare the pre and post-policy intervention price levels and trends in treatment neighbourhoods, with those in control areas, where no policy intervention was present.

We will employ this DID method to estimate the dollar value of the NR program's benefits to residential households that are exposed to the program, both directly and indirectly[3]. We will therefore test whether the implementation of renewal programs causes a discernible break in the pre-NR or post-NR intervention price trends compared with the control neighbourhood price trends, where the controls are selected to be as similar as possible in terms of property and neighbourhood characteristics.

To establish the direct impact of the program (Question 1), the treatment group will comprise all private residential property transactions within the boundaries of a renewal site. The treatment group used to estimate indirect spillover effects of the program (Question 2) will comprise all private residential properties that are within 2000 feet of a renewal site. To test if the renewal impact on proximate properties decline as distance from the site increases (Question 3), we will apply what is referred to by Santiago et al. (2001) as the 'spatial fixed effects model', which uses the DID model to test for price variations across different spatial distances (e.g. 500ft, 1000ft & 2000ft). Finally, we will apply the above results in a cost–benefit analysis (Question 4) by adapting the Schwartz et al. (2006)

[3] By the term 'indirect', we do not refer to impacts on non-shelter outcomes like health, employment, crime etc., but instead to properties that are located outside of NR site areas, and yet adjacent to (e.g. within 2000 ft) their boundaries. Because these properties are not located within NR site boundaries, they cannot be 'direct' recipients of renewal programs (e.g. property upgrades). However, they may receive indirect benefits as a result of spillovers from NR areas.

method of estimating the aggregate increase in property values generated by each renewal site in the period following program implementation.[4]

To construct suitable control groups, two separate quasi-experimental methods will be employed:

- → the near neighbour approach
- → the 'revealed preferences method' (Cigdem 2012).

The near neighbour approach creates a control group using properties that belong to the same Statistical Local Area (SLA) as the NR site, but outside its boundaries. The revealed preference method uses residential properties in locations that will become a renewal site at some time in the future (but transacted before NR implementation) as controls for properties that are already exposed to the program. Matching with the use of propensity scores justifies the comparability of the control group with the treatment group by ensuring that the pre-program characteristics of the treatment and control housing samples are as similar as possible (Dehejia & Wahba 2002; Black & Smith 2003).

[4] Research questions 2 and 3 relating to the spillover effects of NR sites are put to one side in this Positioning Paper and will be tackled in the Final Report. A brief outline of the methodological approach that is employed to measure spillover effects and the costs and benefits of the NR scheme is presented in Section 5 of the Positioning Paper.

2 LITERATURE REVIEW

Neighbourhood Renewal programs have been a recurring feature of housing policy in Australia over the last twenty years. Despite their prominence, we could find few research studies evaluating these programs' outcomes and their effectiveness in achieving policy objectives. Victoria's Department of Human Services has released two evaluation reports to date that rely on community surveys and administrative data to measure program success.[5] But there seems to be a complete absence of quantitative studies using modern quasi-experimental techniques. Randolph and Judd put major gaps in the Australian literature down to the 'complete absence of any national policy interest in estate renewal outcomes over much of the last decade'. The result has been a significant lag in the development of evaluation methodologies and, as a result, Australia lags behind other developed countries (Randolph & Judd 2006, p.97). In New South Wales, Groenhart (2007) has conducted cost benefit analyses on urban renewal centred around public housing using a before/after type methodology. In Victoria, the Department of Human Services evaluates its neighbourhood renewal programs using a mix of survey-based evidence and administrative data such as crime rates, unemployment, household income and various other indicators. We describe the Victorian evaluations in our Final Report.

In the US, on the other hand, there has been extensive quantitative research over the last twenty years offering financial measures of efficacy.[6] Importantly, the credibility of these evaluations has improved over the last decade as a result of considerable advances in econometric techniques, which have progressed from simple before and after studies, through cross-section comparisons of treatment and control, to more sophisticated difference-in-differences (DID) estimators and most recently, the application of an Adjusted Interrupted Time Series (AITS) model. They have also benefited from the application of matching methods (e.g. the propensity score technique) that have advanced the sample design of treatment and control comparisons.

The early US empirical studies on the impact of urban renewal (UR) programs were criticised for failing to provide persuasive enough evidence that improvements in outcome measures could be attributed to the effects of UR. So, for example, before and after studies that revealed improvement in outcome measures in UR sites lacked credibility in the view of critics because favourable but unmeasured background factors (e.g. the economy) could be responsible, and the simple before and after methodology is too crude to rule this out.

According to Galster et al. (2004b), two important methodological developments distinguish recent studies from their earlier counterparts, and offer the prospect of more reliable estimates:

1. The difference-in-differences method, which compares the pre and post-policy intervention difference in treatment (areas the subject of renewal programs) and control group outcomes.
2. The adjusted interrupted time series model (AITSM), which extends the difference-in-differences model to make pre and post-intervention comparisons of both the level and trajectory of the outcome indicator of interest.

Consider house prices as an outcome measure; UR could affect both the level and trajectory of house prices, where change in the trajectory might (say) take the form of accelerating house price appreciation in the post-intervention phase. The level of house

[5] The Department's neighbourhood renewal evaluation reports can be found at the following link: www.neighbourhoodrenewal.vic.gov.au/evaluation.

[6] See Nourse 1963; Schafer 1972; Galster et al. 2006; Bair & Fitzgerald 2005; Ellen et al. 2002; Schwartz et al. 2006; Zielenbach et al. 2010, Ki et al. 2010.

prices is not only higher, their pace of change also quickens. The AITSM is designed to detect both sources of change. An additional advantage to the AITSM method is that researchers are able to detect whether benefits are sustained in the medium to long run.

We first consider the more numerous DID studies of UR programs and their impacts. The DID methodology was motivated by concerns that naïve before and after intervention comparisons can attribute improvements in outcomes to neighbourhood renewal interventions, when in fact they are due to metropolitan and economy-wide changes that have equally benefited all areas. The difference-in-differences method addresses this limitation by designing a sample that contains control sites that are as similar as possible to UR sites, and would therefore be exposed to the same metropolitan and economy wide influences as the UR sites. Before and after intervention outcome indicators are then contrasted in treatment and control areas. This quasi experimental approach draws its inspiration from medical research where it is common to evaluate the effectiveness of drugs, surgical procedures and so on by comparing a group of patients receiving treatment (e.g. a drug) with a control group of patients sharing the same condition, but given a placebo. In medical trials, random assignment of patients between treatment and control groups can replicate laboratory type experimental conditions. In the social sciences, random assignment is rarely possible and it is therefore more difficult to replicate these experimental conditions.[7] Study designs and statistical techniques are instead used to mimic randomised trials.

A difference-in-differences approach is employed by Schwartz et al. (2006) who apply hedonic price models to estimate the effect of publicly subsidised place-based housing investment from 1987–2000.[8] The authors define the treatment group as properties that are sold within 2000 feet of a subsidised investment in an apartment block (housing investment site). A novel feature is measurement of the rate at which treatment effects decay as distance from the investment site increases. This is achieved by adding a continuous distance variable, as measured by the Euclidean distance between a property and the nearest project site, to a hedonic regression model.[9] The authors also test for neighbourhood heterogeneity by detecting whether the impacts of housing investments vary with the typical incomes of residents in the neighbourhoods surrounding housing investment sites.

Houses and neighbourhoods are comprised of complex multidimensional features that are difficult to fully capture in regression model specifications. The hedonic price regression model is therefore vulnerable to bias due to omitted variables. An alternative estimation method is based on pairs of transactions—the repeat sales regression model—this

[7] For a rare example of a study in Australia that applies a randomised experiment with a housing focus, see Johnson G, Parkinson S, Tseng, G and Kuenle D (2011), *Long-Term Homelessness: Understanding the Challenge*, Sacred Heart Mission, St Kilda.

[8] The hedonic technique uses the prices struck in property transactions as the dependent variable in a regression model that relates these prices to property attributes—size, type, presence of various amenities etc.—and neighbourhood characteristics—density, access to urban amenities, crime rates and so on. The coefficient estimates break down the transaction price into a series of implicit prices for the various attributes and characteristics that make up the bundle of housing and location services supplied by the property and neighbourhood. These implicit prices can be interpreted as marginal valuations that housing consumers place on the various service components. They will not always be positive. Neighbourhood characteristics such as crime and air pollution for instance, exert a negative influence on housing prices and negative implicit prices will be estimated for these attributes (see Malpezzi, 2003).

[9] The distance variable is interacted with indicator variables that are 'switched on' when the housing investment program is 'active', and a treatment variable indicating whether within 2000 feet of an investment site. These variable specifications allow the impacts of housing investment programs to decay with distance. Non-linearity in the distance effect is also allowed for by adding a squared distance term or in alternative models by transforming the continuous variable into a categorical variable. Schwartz et al. (2006) transforms the distance variable into four distance intervals: 0–500 feet, 501–1000 feet, 1001–1500 feet and 1501–2000 feet.

approach is less vulnerable to omitted variable bias.[10] Both methods are invoked by Schwartz et al. 2006). However, regardless of whether hedonic or repeat sales models are estimated, they generate findings which suggest that the rate of house appreciation for properties within 2000 feet of a project site is significantly larger than that of areas outside of the ring. Also, the magnitude of the spillovers from the project site decrease with distance from the housing investments. Thus, their findings suggest that 'place-based housing investment may well be warranted to correct for market failures in urban housing markets'.

Castells (2010) employed similar methods to estimate the extent to which Baltimore's three completed HOPE VI[11] redevelopment sites had positive neighbourhood spillover effects on nearby property values. Using geographically coded property sales and structural characteristics data for Baltimore City from 1990–2006, Castells examines the extent to which price levels in the area immediately surrounding HOPE VI sites (micro-neighbourhoods) deviate from price levels in the same neighbourhood but outside the sites (macro-neighbourhoods). In creating micro-neighbourhoods, Castells uses a single ring around each HOPE VI project site, with the ring distance varying with the size of the project. For the relatively small sites, the study uses a 1500 foot ring to define micro-neighbourhood boundaries, but with larger sites a 2000 foot ring is applied. Results from this study show that of the three redevelopment sites, one showed signs of significant positive effects on property values within the micro-neighbourhoods, while the other two showed weakly positive but nevertheless statistically significant impacts.

Galster et al. (2004b) argue that studies of this kind are flawed because they compare micro- and macro-neighbourhood price levels before and after the intervention, but do not control for the trend change in prices. A relative increase in the post-intervention micro-neighbourhood price levels is therefore attributed to HOPE VI. But in the pre-intervention period prices could have been increasing faster in the micro-neighbourhoods, and a continuation of this trend would generate a relative increase in their price level even in the absence of HOPE VI interventions. Galster et al. (2004b) concludes that failure to control for both price level and trend can lead to misleading conclusions.

Galster et al. (2004b) tackle this problem by employing what is formally termed the adjusted interrupted time series (AITS) method in their study of the impacts that the Neighbourhoods in Bloom (NiB) revitalisation program[12] has on targeted areas.[13] Areas targeted for rehabilitation were designated by an NiB team organised by the Community Development Department who determined the precise boundaries of target areas, and developed a two-year work plan which specified buildings to be acquired and rehabilitated or demolished and stipulated where new housing was to be developed (Galster et al. 2004b). To measure the

[10] Repeat sales methods use transactions in properties that have been traded more than once. A strong assumption of the method is that property characteristics and their location attributes remain unchanged between sales. Its main appeal for researchers lies in its ability to assess price changes over the sample time period without the detailed characteristics data required by hedonic methods (see Malpezzi, 2003 for a review of hedonic & repeat sales methods).

[11] The Hope VI Program was introduced in 1992 and entailed the provision of competitive grants to housing authorities for the revitalisation of severely distressed public housing developments. Administered by HUD, it has issued around 254 grants- valued at over $6.1 billion to 132 public housing authorities from the time that it first began issuing grants in 1993–August 2010 (Levy, 2012). The Hope VI program is targeted on public housing and so has a particular relevance to the Victorian Neighbourhood Renewal scheme that is also targeted on public housing. Levy (2012) contains a helpful review of the HOPE VI program.

[12] The NiB program was first introduced in 1998 in Richmond, VA and set out to revitalise select high-poverty neighbourhoods in Richmond. The aim of the program was to essentially improve existing owner-occupied dwellings, rehabilitate blighted properties and develop new housing that would foster mixed-income homeownership environments.

[13] This method was first introduced in Galster, Tatian and Smith (1999).

outcomes of the NiB program, the authors use home sale prices[14] within target areas and compare them with home sale prices in control areas which are not subject to revitalisation, but are otherwise as similar as possible to target areas. To this end, the control group consists of all owner-occupied housing in Richmond census tracts with 1990 median values below $69 000. They go on to measure program impact by applying the AITS method.

The AITS approach makes pre and post-intervention comparisons of both the level and rate of change in the outcome measure (e.g. house prices). The AITS regression model has a complicated specification as follows:

First, the model controls for the time trend in prices in all low-income areas in the sample by specifying a time trend variable that is set to 1 for observations (transactions) during the first period (month, quarter or year) of the study, 2 for observations (transactions) during the second period of study, and so on.

Second, the post-intervention time trend in prices in all low income areas is captured by a time trend that equals 1 for observations during the first post- intervention period, 2 for observations during the second post-intervention period, etc., and zero otherwise. This innovation in model specification will detect any change in the trajectory of prices that is common to all low income areas.

Third, the time trend for house prices in the treatment group (where the intervention applies) is captured by setting a time trend variable equal to 1 for observations in the treatment area during the first period of the study, 2 for observations in the treatment area during the second period of study, and zero otherwise. By defining this time trend over the entire study period it will estimate the rate of change in house prices in the treatment areas relative to control areas in the absence of the intervention.

Finally, the model adds a post-intervention trend variable that equals 1 for observations in the treatment areas and during the first post-intervention period, 2 for observations in the treatment areas but during the second post-intervention period, etc., zero otherwise. It captures post-intervention deviations in price trends specific to the treatment area, and is the key variable.

The study finds that the post-intervention pace of house price appreciation is generally faster in treatment relative to control areas. The authors remark, however, that the reliability of the estimates produced by the AITS method depends on two considerations:

- substantial numbers of frequently recurring observations in treatment areas both pre and post-intervention[15]
- a well behaved trend in the indicator in the pre-intervention period[16] (Shadish, Cook & Campbell 2002).

This method of controlling for both price trend and level is also adopted by Schill et al. (2002) in their analysis of the impact of New York City's Ten-Year Plan—a government initiative introduced in the mid-1980s to spur revitalisation in the city's more distressed regions by constructing and rehabilitating housing. Schill uses the property values of

[14] Authors justify the use of sales data on the basis that 'home sales prices are well known to capitalise many changes in the underlying desirability of neighbourhoods, and thus represent a powerful summary measure of neighbourhood trajectory' (p.462).

[15] The time trend innovations 'eat-up' degrees of freedom and reliable estimates and therefore require a large sample of observations both before and after interventions.

[16] The authors state that 'even with sufficient numbers of observations of the outcome indicator before and after an intervention, the AITS method may flounder if the indicator (especially during the pre-intervention period) is volatile. The AITS fits distinct linear functions to the pre and post-intervention observations. If the underlying relationships are curvilinear or cyclical, however, these linear fits will be both imprecise and arbitrary, depending on the period over which observations are collected' (Shadish, Cook & Campbell (2002, p.532).

surrounding homes as an outcome measure, as sourced from property transactions data from 1980–1999. To perform the analysis, the authors identified all properties that were within 500 feet of housing units that were built or rehabilitated (treatment group), and compared them to properties that were located more than 500 feet from investment sites, but still in the same neighbourhood (control group). The transaction data is matched with an administrative data set that allows measurement of an array of property characteristics such as building age, square footage, and number of buildings on the lot, as well as 18 different building classifications. Property values are once again assumed to be an indirect measure of the intervention's impact on resident wellbeing and local economic development. The measurement technique rests on the use and estimation of hedonic price models.

As noted earlier, unmeasured housing and neighbourhood attributes are difficulties confronting researchers wishing to estimate these hedonic models. Unchanging but unobserved neighbourhood characteristics can be addressed by the inclusion of neighbourhood specific dummy variables. Schill et al. (2002) is a typical example of the use of hedonic price techniques in the UR context, though they go further and interact the neighbourhood (census tract) dummies with calendar quarter dummy variables to control for neighbourhood specific trends in prices. Detection of program impacts on price trends is achieved by specification of a post-completion trend variable that identifies property transactions in the treatment group and occurring in each quarter after completion of New York's Ten-Year Plan.[17] The study finds that housing investments have positive impacts on surrounding property values, their magnitude increases with the scale of the investment, and seem to be permanent rather than temporary. Finally, impacts are greater for properties surrounding homeownership projects compared with those near rental units.

In summary, the contemporary approach to NR evaluation in quantitative studies is quasi-experimental. Study designs contrast outcome measures for properties within and in the near vicinity of renewal sites with the same outcome measures for control properties that share the same characteristics, but are not exposed to the direct or indirect (spillover) effects of NR. The outcome measure employed in all studies reviewed is house prices; hedonic price and repeat sales regression models are frequently used to control for observable differences in property and neighbourhood characteristics between treatment and control samples. More modern matching techniques (e.g. the propensity score method) are an alternative approach to minimising the differences between treatment and control samples. Other recent innovations in method allow researchers to detect whether the post-intervention trajectory of outcome measures in NR sites deviate from those in control areas. A focus on trajectories offers more reliable estimates because changes in the relative level of outcome measures can be falsely attributed to NR, when it is in fact the continuation of pre-intervention price trends. The use of trajectories also allows researchers to investigate the permanence or otherwise of NR benefits. All studies reviewed choose treatment boundaries that define areas over which urban renewal programs impact; but this choice is typically arbitrary. For instance, in a study on the effect of a dispersed subsidy housing program on property values, Santiago et al. (2001) defines the treatment group as all properties within 2000 feet of a dispersed housing site. To detect whether impacts vary, the authors create a series of 'neighbourhoods' centred on each dispersed housing site, each one comprising one of several concentric rings: 0–500 feet, 501–1001 feet and 1001–2000

[17] They also include a dummy variable to denote properties within 500 feet of an existing or future project site to control for any systematic differences between properties that are proximate to a project site and those that are not. In an extension of the model, the authors control for differences between homeownership investment sites and rental units types by including separate ring variables for properties that are proximate to homeownership and rental developments. Also, authors generate dummies to distinguish between properties that are proximate to newly constructed project sites and those that are rehabilitated.

feet from the site. But practice varies across studies and so no firm guidelines can be gleaned from the literature on this important decision.[18]

All but one (Newell 2010) of the studies reviewed find that place-based housing investment interventions have positive impacts as detected by house price premiums. Two papers, those by Rossi-Hansberg, Sarte and Owens III (2010) and Galster et al. (2004b), using variants of a quasi-experimental approach to estimate impacts from the same neighbourhood renewal program in Richmond, USA conclude that this program works. Rossi-Hansberg, Sarte and Owens III (2010) show that gains exceed program costs, calculating that over a six-year period, a dollar of home improvement generated $2–6 in land value. Both Ellen et al. 2002 and Ellen & Voicu 2006 report positive findings for New York urban revitalisation programs. With one exception there are no studies of this type in other countries; the exception is Ki and Jayantha's (2010) study in Kowloon, Hong Kong, which again detects positive house price impacts. The empirical studies of this type seem to suggest that neighbourhood renewal 'works'.

But these studies are overwhelmingly US based and their findings could be location specific given differences in housing institutions and market features. For example, US housing markets have featured pronounced ethnic and racial segmentation that have motivated NR interventions as minorities are frequently over-represented in NR sites. Ethnic and racial 'fault lines' are not an important feature of Australian housing markets. These differences in context might make a difference such that findings from a country such as USA are not replicated in Australia, and so the same policy conclusions cannot be drawn. This is an important motivation for our research project.

[18] Schwartz et al. (2006) and Zielenbach et al. (2010) also confine the treatment group to properties within a 2000 foot radius. However, in areas where there were nearby public housing sites, Zielenbach et al. (2010) reduce the size of the treatment areas 'so as to minimize the influence of the non-selected sites'. Bair and Fitzgerald (2005) define areas surrounding HOPE VI project sites as those properties that had longitude and latitude coordinates within a radius of 1.5 miles (7920 feet) of the centre point of the HOPE VI site. Ki and Jayantha (2010), on the other hand, confine treatment areas to properties within 750 metres (2460 feet) of the centre of the redevelopment area. Castells (2010) alternates between 1500 feet and 2000 feet rings from each project site, depending on the size of the project. Ellen et al. (2002) defines treatment areas as properties within 500 feet of program sites.

3 METHOD

The idea behind our approach is a simple one; renewal programs typically aim to improve the physical appearance of properties and communal facilities, while also investing in community services that make neighbourhoods more attractive, and strengthen the human and social capital of residents. If successful in meeting these goals, the demand for housing in and around neighbourhood renewal sites will increase, and house prices will rise by more than would otherwise be the case. Capturing this increment or premium in house price movements due to neighbourhood renewal (NR) is more complicated than the idea motivating the research approach. House prices might well have increased in the absence of neighbourhood renewal; interest rates could fall and earnings and employment growth lift during the course of a renewal program. Price rises in NR sites may then merely reflect city-wide improvements in economic conditions.

The approach followed in this study mimics the treatment versus control methodology popular in medical research studies. Medical researchers are often in the advantageous position of being able to randomly assign patients between the two sample groups. There is then no reason to believe that patients selected to receive the medical treatment are more likely to recover regardless of the medical services they received. But when assignment is non-random this attribute of the research design is lost. Suppose, for example, that patients elect whether or not to receive treatment, and younger patients tend to 'self-select' into the treatment group. It is then hardly surprising to find that their rate of recovery/medical improvement is better than that of the older patients comprising the control group. The treatment/control outcome comparisons are then problematic, because the treatment group's relative improvement in medical condition could be due to their youth and greater recovery capacity, rather than the medical treatment.

In the settings that social scientists confront, random assignment is rarely possible. Study designs seek robust comparison groups that as far as possible mimic random assignment. Our design exploits the sequenced nature of the Victorian neighbourhood renewal program that allows a revealed preference method to be applied. The group of neighbourhood sites introduced at the onset of the program are compared to those sites that are actioned later in the program, but over a period of time prior to the control sites becoming part of the program. The control group of sites are an appealing choice because they must share similar characteristics to those that prompted selection of the treatment sites for neighbourhood renewal. Property price trajectories are then compared by matching each property in the treatment sites with a property from the control sites that is most similar in terms of property and area characteristics. A propensity score method will be used for this matching exercise.

But the comparison is over a time period before the control group benefit from the resources delivered by neighbourhood renewal. Thus, the revealed preference method has a limited timeframe over which to conduct treatment versus control comparisons.[19] We therefore use a second 'nearest neighbourhood' study design that selects control properties from the same Statistical Local Areas (SLAs)[20] that NR sites are located within; the propensity score method is again used to select control properties. Controls selected from the same SLA should share similar area characteristics, including broadly uniform market conditions. Post-intervention comparisons are longer because the controls are never exposed to NR programs. It is this second study design that we explain below. Since the same statistical

[19] It might also be flawed if the early neighbourhood renewal sites are chosen because they are most likely to be successful.

[20] The SLA is a geographic spatial defined in the Australian Statistical Geography Standard (ASGS). They are defined as the smallest spatial unit in non-Census years; in Census years, they are comprised of one or more Collection Districts (CDs).

techniques and estimation methods are used in both approaches there is little to be gained from a detailed outline of both nearest neighbourhood and revealed preference approaches.[21]

The rest of this Section is organised as follows. Data sources are described first, including a detailed description of the key characteristics featured in two datasets central to our modelling. This is followed by a description of the processes followed to form a final merged dataset, and how GIS software is used to identify urban renewal sites and the properties that are exposed to the direct effects of NR. Section 3.2 discusses the sample design for the treatment and control groups and presents the propensity score method which will later be used to select appropriate control observations for the implementation of our quasi-experimental methodology. Finally, Section 3.3 outlines the modelling approach that will be performed in this study.

3.1 Data

The analysis draws on two main datasets obtained from the Office of the Victorian Valuer-General (VG), namely the valuation and the property transactions datasets (supplied in a confidentialised format.)[22] These are described in detail below.

3.1.1 Property valuations data

Our first raw data source is 2008 property valuations data which is audited at the state level by the Valuer-General Victoria (VGV). Valuations data is collected by valuation officers who are, in turn, contracted by individual municipalities for the purposes of levying property rates (taxes).[23] The data is confidentialised via the removal of a number of fields (owner details and unimproved site values). While only a portion of the overall valuation dataset held by the VGV is released to the public, it nonetheless contains a rich set of property-level characteristics for each rateable property. The dataset is a point-in-time record of all rateable properties as at 2008: each property should appear, but can appear only once.

The valuations data is the key source of detailed property-level and neighbourhood characteristics. These characteristics are essential for the estimation of hedonic price models using property sales identified from the transactions dataset (see below). Relevant variables in the valuation dataset include:

- property address (main variable used to perform merger with transactions dataset)
- last sale date
- last sale price
- land use classification category (residential, commercial, industrial, agricultural)
- number of bedrooms
- year of construction
- construction material.

Spatial variables were also added to the valuation data using VicMap spatial reference datasets (Wood et al. 2012). These include:

- X and Y coordinates that locate properties on a map
- distance from designated principal and major activity centres

[21] Where there are material differences we have noted them below.

[22] This database was originally developed under AHURI project 30590 to analyse land use planning policies. We are grateful to Elizabeth Taylor who was responsible for the original design and creation of the merged dataset.

[23] At the time of conducting the analyses, the 2008 valuation records were the most recent available. Valuations are undertaken every two years.

- distance from railway stations
- nearest secondary state school
- distance to nearest secondary state school
- zoning codes that regulate land use
- overlays that identify neighbourhoods with land and buildings that have idiosyncratic characteristics e.g. environmentally significant landscapes or clusters of historical buildings. Areas and properties subject to overlays must comply with additional restrictions on the use of land and/or the design of buildings; e.g. a permit is required to remove vegetation in environmentally significant areas.

While rich in property and spatial variables, the valuations dataset does have certain limitations. First, it is cross-sectional point-in-time (2008) data. The transactions data, on the other hand, is a cross-section, time series data set comprising transaction information for every property sold over a 21-year period (1990–2011). The difference in data formats means that when merging the valuation data with the transactions dataset, all sold properties are assigned 2008 characteristics regardless of the year they are sold. This may result in inaccuracies if the property (e.g. renovations) and/or spatial (e.g. zoning) characteristics changed from 1990–2008. For instance, a two-bedroom house that was last sold in 1999 may have been subsequently renovated with improvements that include an extra bedroom. In the final merged dataset however, the house will be recorded as having three bedrooms even though it contained only two bedrooms at the time of sale (1999). These caveats are important, but concerns are assuaged by two observations. First, structural changes to houses and apartments are costly and executed infrequently. Second, zoning and overlay regulations also limit the changes that owners can make to the design, height and size of their homes. As a result the built environment changes slowly.

While the valuation dataset is confined to land and buildings in metropolitan Melbourne, the transactions dataset contains Victoria-wide sales records. This means that transaction records for properties lying outside of metropolitan Melbourne cannot be matched with property and neighbourhood characteristics, and are consequently omitted from the final merged dataset. This has an important implication for the study design as it restricts NR evaluation to those sites located in the metropolitan region.

3.1.2 Property sales data

The property sales data is an unbalanced panel,[24] it consists of one file for each year and residential property type (house, land, units/apartments). Property sales data is collected at the time of sale for taxation purposes. Being records of sale, a property will appear more than once if it is sold multiple times, hence the longitudinal (panel) nature of the data set—the transaction profile of each property is tracked over the 1990–2011 sample period. Note that a property will not appear if it has not been sold during the period. The residential property sales data contains the following information:

- property address
- suburb
- municipality
- sale price (nominal)
- date of sale

[24] We refer to the dataset as an unbalanced panel because while it includes consecutive observations on individual properties that sell one or more times, the number of observations in any one year will differ because some properties sell frequently, others infrequently.

- sale type (house, unit/apartment, vacant land)
- lot area (measured in metres squared)
- community district number
- Local Government Area code
- statistical Local Area Code
- dwelling type (e.g. single residential, detached home etc.)
- spatial coordinates (latitude/longitude) for each property.

3.1.3 Merged dataset

The valuations and transactions datasets are merged using property address fields that are available in simple format in both datasets. Address fields were augmented by the addition of lot numbers so that transactions in units/apartments could be matched with a corresponding valuation record. Approximately 70 per cent of the transactions in the property sales dataset were matched to key property and spatial characteristics in the valuations dataset. This means that successfully matched sales records were assigned with their corresponding property-level information such as property characteristics, location in relation to principal and major activity centres (areas designated by planning authorities as focal points for employment growth, transport nodes and urban amenities), and planning regulations such as zoning and overlay areas.[25]

Both the valuations and sale information are collected for the purposes of revenue collection and as a result offer a good level of coverage; there should also be a high degree of reliability with regard to sales values and other items used for the assessment and collection of stamp duties, land taxes and local government rates.

Table A1 provides a list of variables included in the final merged dataset along with their definitions and the unit of measurement.

3.1.4 Identification of neighbourhood renewal sites

A critical step in the study design is identification of the street-level location and boundaries of each neighbourhood renewal site in metropolitan Melbourne.[26] Across Victoria, 21 projects have been progressively launched since 2001, with 11 project sites located in metropolitan Melbourne and 10 project sites in regional Victoria. Table 1 below provides a list of the neighbourhood sites in metropolitan Melbourne and their locations.

Table 1: List of neighbourhood renewal sites in metropolitan Melbourne

Year of NR program commencement	NR sites (area)	Statistical Local Area (SLA)
2002	Braybrook (Braybrook and Maidstone)	Maribyrnong
	Collingwood	Yarra
2003	Fitzroy Atherton Gardens	Yarra
	Ashburton (Ashburton, Ashwood and Chadstone)	Boroondara and Monash
	Broadmeadows	Hume
	Werribee (Heathdale)	Wyndham

[25] The overlay boundaries are identified using VicMap database 2010 version.

[26] We would like to thank Olwyn Redshaw and Mark O'Driscoll from the Victorian Department of Human Services for their assistance.

	Doveton-Eumemmerring	Casey
2006	West Heidelberg	Banyule
	East Reservoir	Darebin
	Hastings	Mornington Peninsula
2009	Flemington	Moonee Valley

Because the final property dataset pertains to metropolitan Melbourne, our analysis omits project sites in regional Victoria. However, since the interventions are similar and regional sites belong to the same state and are consequently subject to the same institutional arrangements as metropolitan NR sites, we confidently expect the findings to have relevance to NR programs in non-metropolitan regions.

Fitzroy Atherton Gardens, East Reservoir and Flemington project sites are also left out of the sample frame; the Fitzroy NR site does not include any private sales; the East Reservoir NR site had too few property transactions to derive an estimate; and the Flemington site could not be analysed because NR was initiated late in the study timeframe, leaving too few post-treatment years for robust estimation of impacts. This leaves eight neighbourhood renewal sites making up our sample frame which still leaves us with a healthy coverage of the Victorian government's NR program. The following NR sites are included in the final data framework; Braybrook/Maidstone, Collingwood, Broadmeadows, Werribee, Doveton-Eumemmerring, Hastings, West Heidelberg and Ashburton, Ashwood and Chadstone.

3.2 Sample design for a quasi-experiment

A key research objective is the design of an appropriate counterfactual that is comparable with properties exposed to NR programs. Experimental or randomised methods are one way of potentially solving the problem of identifying an appropriate counterfactual, and are considered the most robust approach because of the unbiased nature in which the experiment is conducted. In an experimental setting there is random assignment of the program treatment to recipients, all of whom are drawn from the same population. Program impacts are then estimated by comparing outcomes of those exposed to the treatment and those in the control group who are not exposed to the program. While highly appealing, random assignment methods cannot be applied in this study, the sites selected for NR were not randomly chosen, but selected because of their 'relative disadvantage compared to other parts of the community and because there are high concentrations of low-income residents living in older and relatively neglected public housing' (Department of Human Services 2002). Consequently, quasi-experimental methods are drawn on to find a suitable counterfactual group.

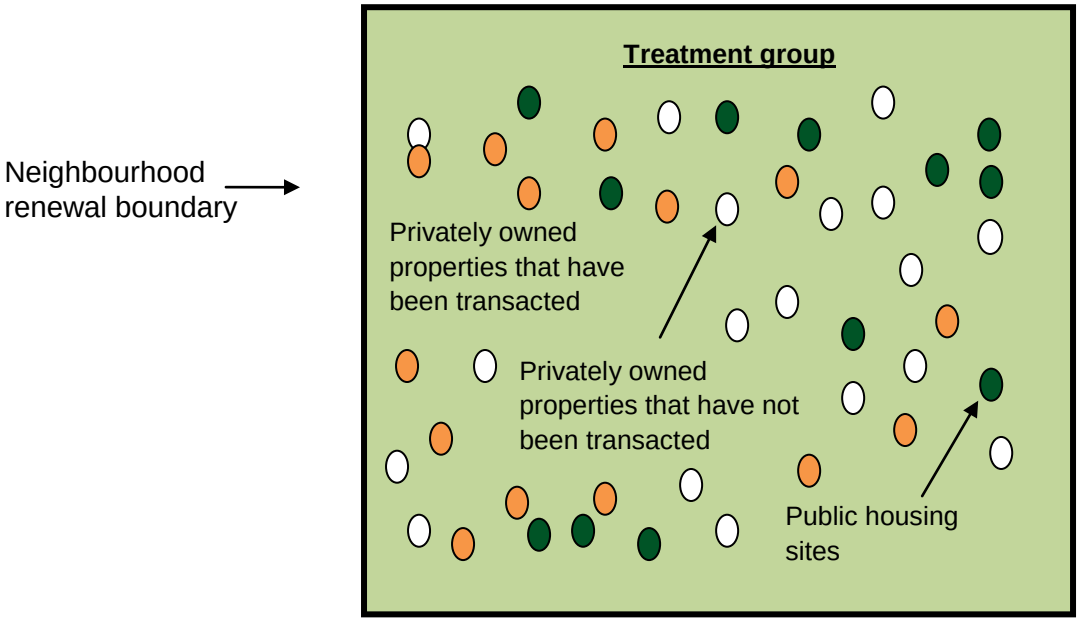
We begin this Section with a conceptual analysis of the problems endemic to the measurement of program impacts in non-randomised settings. We then outline our proposed solutions to these problems.

3.2.1 The design of treatment and control groups

GIS tools were used to create map layers that delineate the boundary of each neighbourhood renewal site. The transactions dataset was then imported into MapInfo Professional via the latitude and longitude fields to identify all property transactions lying within the boundaries of neighbourhood renewal sites; these properties form a 'treatment' sample containing privately owned housing units directly exposed to the neighbourhood improvements and upgrades executed in their immediate vicinity.

Figure 1 below illustrates the treatment sample's derivation. The dark border denotes the NR project site boundaries.[27] The orange circles represent privately owned dwellings and vacant lots that are located within each project site and have been sold at least once from 1990–2001; the white circles identify privately owned dwellings and vacant lots that have never been sold over the sample period; and the green circles denote public housing that benefits from rehabilitation. Because we can only observe the prices of properties represented by the orange circles, the estimates of price premiums due to NR are based solely on the sales records of privately owned properties that have been transacted. However, these premium estimates will also accrue to properties that have not been the subject of transactions. The Project's Final Report will report aggregate measures of benefits that use price premium estimates to impute gains that accumulate on those properties.

Figure 1: Identification of neighbourhood renewal sites and selection of the treatment group



3.2.2 Identification of the control sample of properties

To identify an appropriate comparison group, we first pinpoint all properties and land plots belonging to SLAs where a neighbourhood renewal site was initiated from 2002–09.[28] There are systematic SLA-level differences in levels of education, unemployment, average income, crime rates and other socio-economic and demographic profiles. This restriction is designed to eliminate these inter-SLA differences that could confound comparisons between otherwise similar properties. Figure 2 below presents a map of Melbourne and highlights those SLAs where an NR site is present; areas shaded in purple represent the boundaries of the NR sites in the study sample.

To ensure that treatment and control groups are as comparable as possible, we apply the propensity score method to match treatment properties with control properties (in the same SLA) that share similar neighbourhood and property characteristics. Thus, treatment

[27] This treatment sample captures the direct benefits of the NR program; it ignores any indirect or 'spillover' effects that may result from the program. A separate treatment sample will be reported in the Final Report that will also determine the indirect effects of the program.

[28] With the revealed preference method, we identify transactions in property within the boundaries of NR sites initiated toward the end of the study time frame. The control group is then formed by the application of matching techniques that are explained later in this Section.

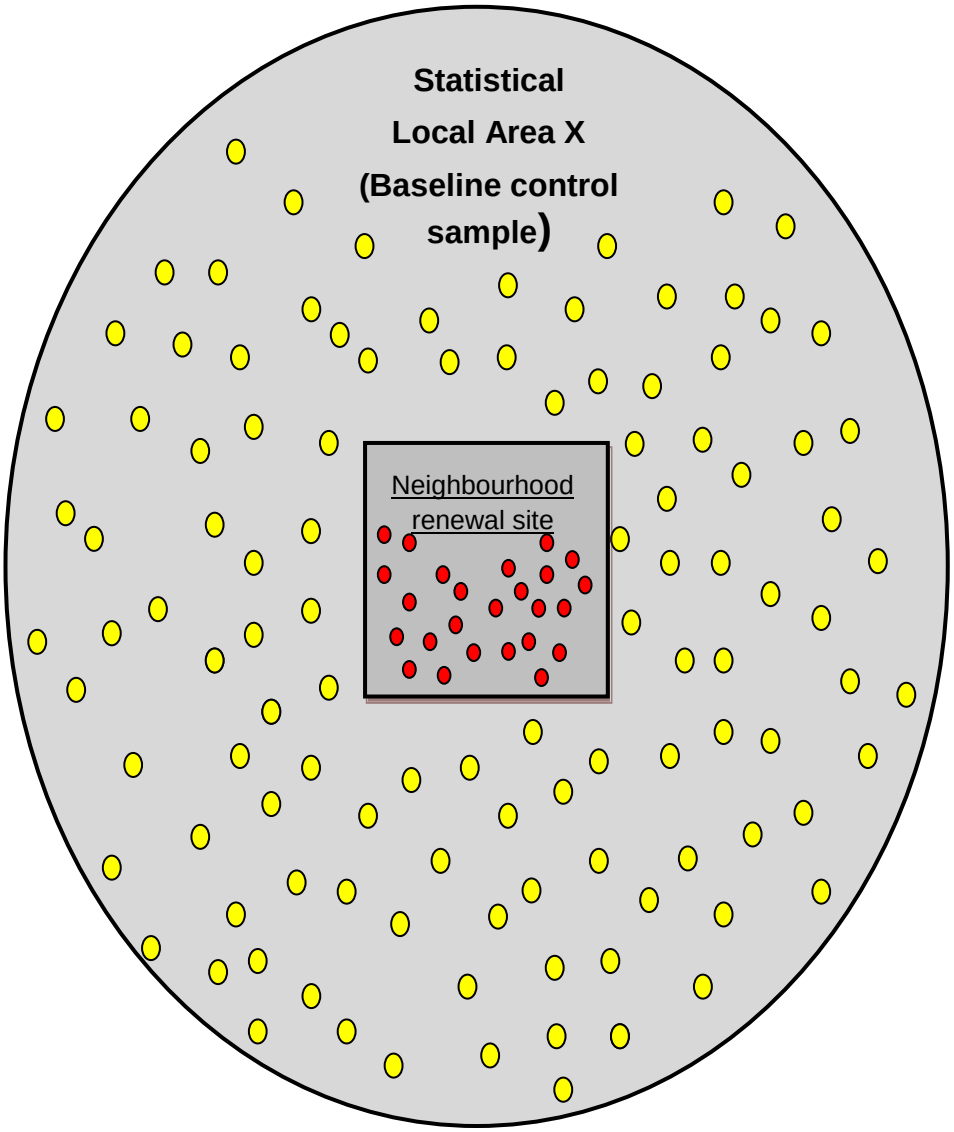
properties are matched not only on their SLAs but also on their property-level characteristics.

Figure 2: Map of Statistical Local Areas where a current or future NR site is present

Figure 3 below presents a hypothetical SLA (in the form of a circle) from which the control sample is drawn; the red circles represent transactions in property and land parcels that will form the treatment group. The yellow circles are transactions in property and land parcels that are outside the NR site boundaries but within the same SLA (the baseline control sample).[29] The control group is formed from these transactions using the propensity score method; it optimally matches the red circles with comparable yellow circles based on their property and neighbourhood characteristics.

[29] In forming this baseline control sample, indirect effects of the NR program are ignored for the moment and will be dealt with in the Final Report.

Figure 3: Selection of the baseline control sample



3.2.3 Propensity score method

The propensity score method is a systematic approach to the design of a control group that is optimal as a proxy for the counterfactual (Y_{0i} in Table 3). Interpreted in terms of Figure 3 above, the propensity score method optimally matches every red circle with one of the yellow circles. Central to the technique is estimation of a regression model that is capable of estimating the likelihood of being selected for treatment (NR). The probit and logit maximum likelihood estimators are suitable for this purpose. These estimators model the probability that a property (such as those depicted in Figure 3) is assigned to a NR site, given a vector of measured property and neighbourhood characteristics *X*. The models are capable of generating predicted values that can be transformed into predicted probabilities or propensity scores (p_j). A nearest neighbour algorithm is employed to match each property that has been exposed to NR, with the one property outside NR sites that is closest in terms of predicted probability of selection (Becker & Ichino 2002). More formally let C denote the set of properties in the control group, and $C(i)$ the control property matched to the treated

property i. One control property j with propensity score p_j is selected for every treated unit i. The nearest neighbour algorithm selects that control according to the rule

(3) $$C(i) = min \parallel p_i - p_j \parallel$$

The rule selects that property j with a propensity score p_j that is closest to p_i as the 'match'.

Nearest neighbour matching provides users with the option of matching with or without replacement. With replacement, the property transaction j that satisfies the algorithm in (3) is put back into the sample and is potentially available as a match for a different treatment property. Thus nearest neighbour matching with replacement imposes no restrictions on the number of times the same control unit is matched to more than one treated unit. Matching without replacement removes a selected control unit after every match. When performing nearest neighbour matching without replacement, the order in which matching is performed will be an important influence on $C(i)$ .Researchers are therefore advised to randomly sort observations prior to matching (Caliendo & Kopeing 2008; Piquero & Weisburd 2010).

There is also the option of using more than one nearest neighbour per treated unit (commonly referred to as 'oversampling'). For example, we could select δ > 1 nearest neighbours and compute the (unweighted) average of the δ property transactions to represent the control for modelling purposes. As δ increases the 'distance', separating each treatment from its nearest neighbour control becomes more uniform. With δ = 1 some treatments might be matched to 'close' near neighbours (very similar propensity scores), others to nearest neighbours that are nevertheless clearly different (very different propensity scores). The choice of δ involves a trade-off between a lower variance resulting from the additional observations used to construct the control group, and a larger bias, which is the result of generally weaker matches. When oversampling, the researcher must choose the number of matching partners (δ) allowed per treated unit, and the weight assigned to them.

A drawback of the nearest-neighbour matching algorithm is that all properties in the treatment group are assigned a control match regardless of the difference in propensity scores. We can then end up with some very weak matches. One response is the caliper method which imposes a tolerance level on the maximum distance between the propensity scores of treatment and potential control observations. This means that control observations are matched with treatment observations if they have the closest propensity score and the propensity score lies within a predefined calliper, or 'propensity range'. The advantages and disadvantages of the caliper method are the same as when allowing for replacement; it promotes better quality matches by avoiding bad matches, but may also generate a larger variance due to fewer matches. However, the difficulty facing researchers when applying the caliper method is deciding a priori what tolerance level is most suitable. A variant of the caliper method is the radius matching method, which uses all control observations that lie within the defined caliper rather than only those with the closest propensity score. The benefits of this approach are that it ensures quality matches by omitting observations that lie outside of the caliper and makes use of all observations that lie within the caliper (Caliendo & Kopeing 2008).

An alternative to the nearest neighbour method is Kernel Matching which exploits all control observations in the baseline control sample to construct a suitable counterfactual outcome for the treatment group. Program impact is evaluated by comparing the outcomes of treatment observations with the weighted average of the outcomes of all the untreated observations, the highest weight being assigned to observations with propensity scores that are closest to the treated observation. This means that strong matches attract a high weight and weak matches attract a low weight. An advantage of the Kernel approach is the smaller

variance that will result from using a larger data sample. The drawback, however, is that some of the observations used may be poor matches.[30]

Researchers using the propensity score matching method commonly assess the quality of the match produced by the matching algorithm. A standard tactic is the conduct of t-tests to detect statistically significant differences in the mean values of the property and neighbourhood characteristics in treatment and control groups. While differences are expected prior to matching, the distribution of the covariates should be balanced after matching and so no statistically significant differences should be found between the treatment and matched-control groups. One can also compare the Pseudo-R^2, which signifies how well the covariates in the probit regression explain the probability of participation for the matched sample before and after matching. We would expect no significant differences between the two groups in the distribution of covariates after matching, and therefore a relatively low Pseudo-R^2 (Caliendo & Kopeing 2008).

The propensity score method's reliability rests on two key assumptions (see Appendix 2 for a formal statement). First, NR sites are declared on the basis of socio-economic and demographic criteria that cannot be influenced by decision-makers and are independent. This requirement (commonly referred to as 'ignorable treatment assignment') would be violated if (say) policy-makers are tempted to select NR sites where they think the outcome will be favourable, rather than objective criteria reflecting priorities for intervention. Second, the 'common support assumption' requires that for any vector of property and neighbourhood characteristics X, the probability of exposure to NR lies between zero and one. To appreciate the intuition, consider Figure 3 above. If all the properties represented by the yellow circles (outside NR sites) have a probability of exposure equal to zero, it would be impossible to find any subset with a set of X characteristics similar to those evident among the properties located inside the boundaries of NR sites.

3.3 Modeling approach

Once suitable treatment and control group sample designs are defined, researchers use them to estimate the price premiums that are our outcome measure. In this Section we discuss the modelling approach to estimation of these NR impacts within the boundaries of NR sites (direct impacts). But before discussing preferred models we discuss the difference model used in early impact studies.

3.3.1 The difference model

The difference model is a regression using a before and after sample of observations on transactions within NR sites. It therefore uses what we have been calling a treatment sample—properties exposed to the NR intervention—but no control group.[31] The before-after sample design estimates the impact of the treatment (NR) by measuring the difference between price observations before treatment is introduced, otherwise known as the pre-test, and price observations after the treatment, that is, the post-test (Wholey et al. 2010).

Meyer (1995) demonstrates that the following linear regression will (under restrictive assumptions) yield an unbiased estimate of NR-related house price premiums:

(4) $$y_{it} = \alpha + \beta d_t + \epsilon_{it},$$

where y_{it} is the price of property i sold in time period t = 0,1; t=0 identifies a property sale pre-test, t=1 when sale occurs post-test, and d_t *indexes individual property transactions*

[30] This problem can be alleviated by imposing the common support condition (for details see Caliendo & Kopeing 2008).

[31] The before/after design can also include properties adjacent but close (say within 2000 feet) to the boundaries of NR sites to detect indirect spillover effects.

1,2,3…….n within NR sites. Variable d_t is a dummy variable taking the value 1 if t = 1 and 0 otherwise. The coefficient β gives an estimate of the impact NR has had on property prices within the NR site. There is an important but little recognised point to make about the exact interpretation of β. When the sample is restricted to transactions in properties that are within the boundaries of NR sites, but not the subject themselves of renewal (e.g. renovation or repairs), β are external benefits. If NR outlays on land and buildings raise their capital values by an equivalent amount, β is a net return. If the sample used for estimation includes the properties renovated or repaired by NR programs β must be compared with cost outlays to compute a net return.

A key assumption of this model is that property prices would remain unchanged in the absence of NR.[32] Subject to this condition, an unbiased estimate of β can also be calculated using the following equation:

(5)
$$\hat{\beta}_d = \Delta\bar{y} = \bar{y}_1 - \bar{y}_0,$$

where the bar signifies the mean price of properties and the subscript denotes the time period.

But there are various threats to the internal validity of this study design. For instance, the NR scheme may have been introduced at a time when the property market was experiencing an upswing. In this case, using the before-after method will lead to biased results as increases in property values will be attributed to the NR program and not the housing boom. In this case it will overstate the impact of NR. When property markets slump in (say) the aftermath of interest rate increases, we risk understating the impact of NR.

3.3.2 Difference-in-differences method

The difference-in-differences (DID) method overcomes this limitation by using the treatment and control group study design described above. It first computes the difference between the treatment and control group's outcome measures in the period before the program is introduced—this is referred to as the 'first difference'. The difference in outcome measures between treatment and control groups is again computed in the period after the program is implemented—this is called the 'second difference'. The difference-in-differences technique measures program impact by subtracting the second difference from the first difference. The rationale is that program impact is discerned by the difference in outcomes for treatment and control groups in the period after the program is implemented, net of any pre-existing differences in outcomes between the two groups that pre-date the program. So, for instance, house prices in the treatment NR site may have been lower than in the control for the period before the program is introduced; the DID method measures whether this price discount narrows in the period after the program is implemented.

The DID impact measures can also be obtained on estimating the following linear regression model:[33]

(6)
$$y_{it}^j = \alpha + a^1 d^j + \alpha_1 d_t + \beta d_t^j + \epsilon_{it}^j,$$

[32] The technical statement of this condition is $E\left[\epsilon_{it} \mid d_t\right] = 0$ —the conditional mean of the error term is independent of the value of *d* (Meyer 1995, p.154).

[33] If covariates do not balance after matching, they are removed from the probit and added into the OLS regression (equation 6). This procedure is referred to as the regression adjusted model (see Dehejia & Wahba 2002). The regression adjusted model will then include the property characteristics and neighbourhood features commonly entered into hedonic housing models. This is how models will be estimated but are omitted from equation 6 for the purposes of simplifying the exposition and presentation in Figures 4 and 5. It should also be pointed out that the dependent variable (property price) is typically measured as the natural logarithm of price.

Where again t=0 identifies a property sale pre-treatment (NR), t=1 identifies sales post-treatment but the outcome measure (house price) y is now also indexed by j which is equal to 0 if properties belong to the control group, and 1 for properties belonging to the treatment group (NR site). The intercept α captures the average price of properties in the control and in the period pre-dating the NR program. Variable d^j is dummy variable equal to 1 if the property is in the treatment group, and zero if in the control group. d_t is a time dummy equal to 1 if t is 1 and 0 otherwise; α_1 is then an estimate of the change in house prices common to all properties in period $t=1$. Coefficient a^1 captures differences in property prices that pre-date the NR program's implementation. The parameter of particular interest is β on the interaction term d_t^j which is equal to 1 if the property belongs to the treatment and the period after the program is implemented, 0 otherwise. β measures the impact of the NR program on property prices in the NR site. The key assumption is that the mean price difference between treatment and control groups of properties would remain unchanged in the absence of NR.[34]

This model is summarised in Table 2 where outcome could in principle refer to any measure of NR impact.

Table 2: Coefficients of interest in the difference-in-differences model

	Pre-neighbourhood renewal outcome	**Post-neighbourhood renewal outcome**	**Difference**
Properties in treatment group	$\alpha + a^1$	$\alpha + \alpha_1 + a^1 + \beta$	$a^1 + \beta$
Properties in control group	α	$\alpha + \alpha_1$	α_1
Difference-in-differences			β

Source: Adapted from Buckley, Jack and Yi Shang (2003), 'Estimating policy and program effects with observational data: the "difference-in-differences" estimator', Practical Assessment, Research & Evaluation, 8(24).

To obtain an unbiased estimate of β, the difference-in differences model can be used:

(7)
$$\hat{\beta}_{dd} = \Delta\bar{y}_t - \Delta\bar{y}_c = \bar{y}_t^1 - \bar{y}_t^0 - (\bar{y}_c^1 - \bar{y}_c^0),$$

where bar indicates an average over i, the subscript denotes the group, and the superscript denotes the time period. To perform this type of evaluation design however, we must have a sample of property sales before and after the initiation of the NR site in both the treatment and control groups.

The simple difference-in-differences method can also be depicted diagrammatically as in Figure 4.

[34] The technical statement of this condition is $E\left[\epsilon_{it}^j \mid d_t^j\right] = 0.$

Figure 4: Potential types of price impacts from NR: the DID method

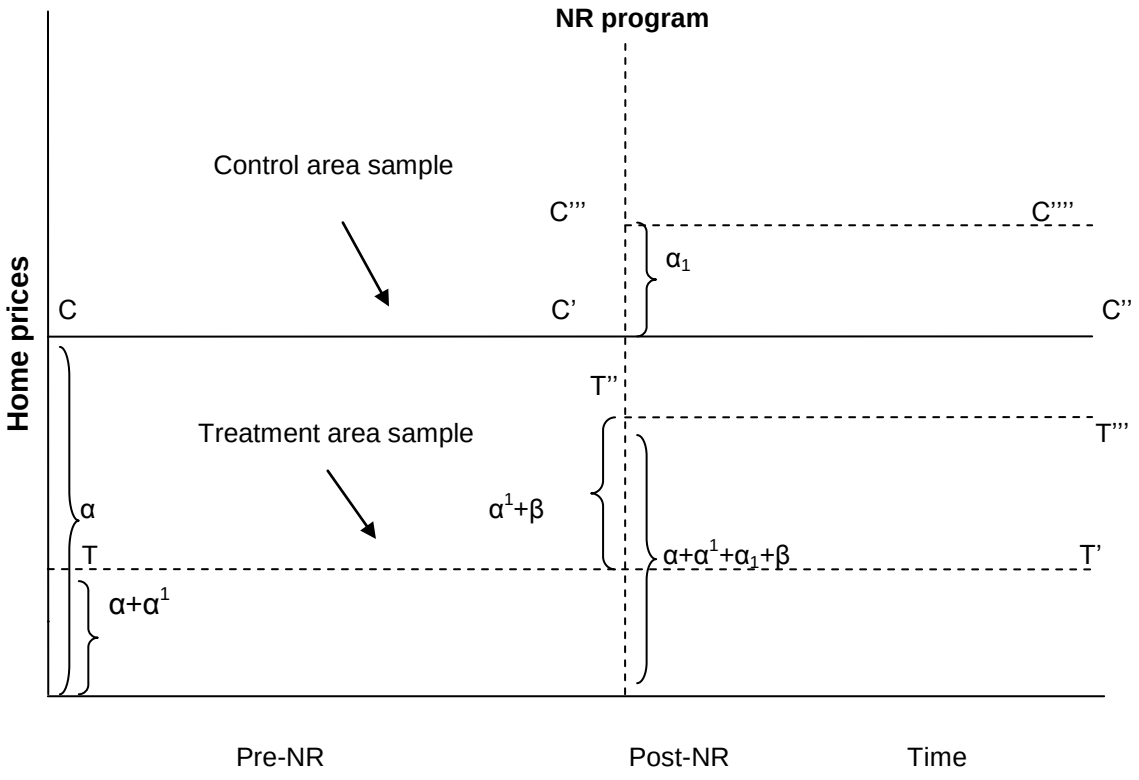
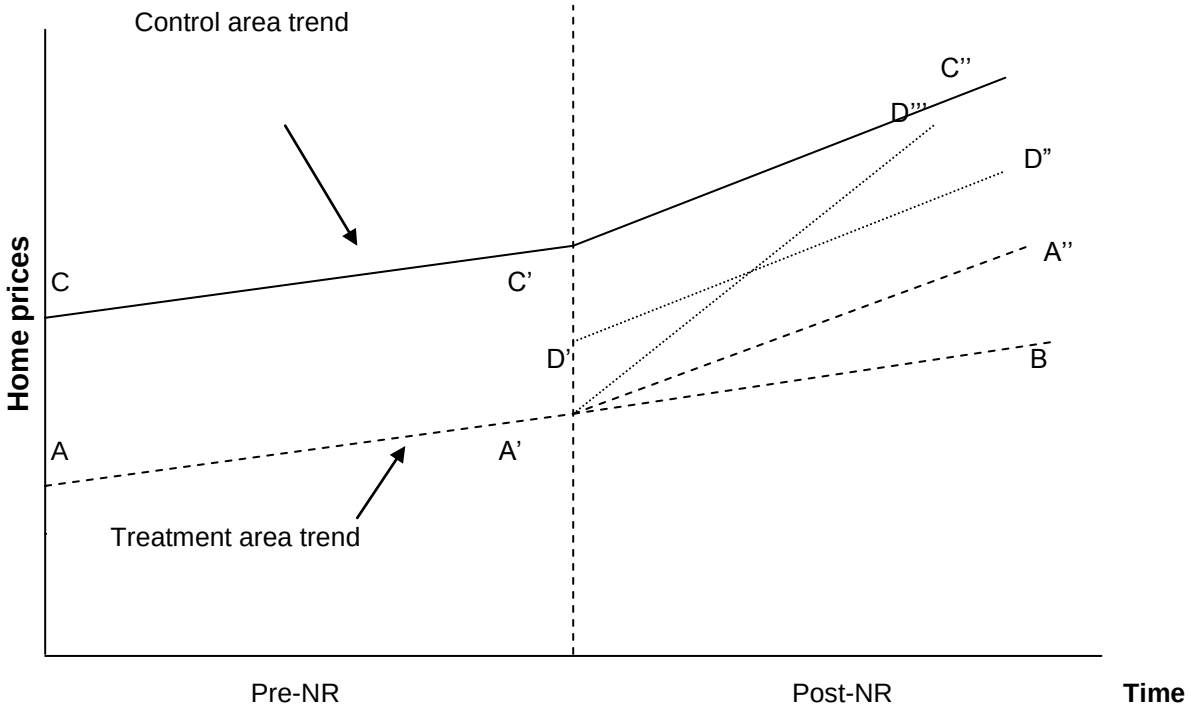


Figure 5: Potential types of price impacts from NR: the AITS method



Note: Adapted from Galster, Tatian and Pettit (2004a), Supportive Housing and Neighborhood Property Value Externalities, *Land Economics*, vol.80, no.1, p.41.

In the context of this study, the impact neighbourhood trend line represents the price trend of properties that are proximate to an NR site, and is distinguished by the dashed line. The control neighbourhoods comprise properties that are similar to the treatment group but outside of the NR site boundaries; they are represented by the solid line.

To determine the effect of the NR scheme on property values, the simple difference-in-differences method allows us to detect whether there is an upward shift in the price level in the treatment neighbourhoods in the period following the development of an NR site. Thus, it enables us to gauge whether, in the event of a price change in the post-NR period, the price shift was also prevalent in the general housing market (i.e. the control group) or whether it was encountered by the impact neighbourhoods alone. In Figure 4 above, we can see that the NR program results in an upward shift in property values in the treatment group (from T-T' to T''-T'''). Note that the shift is a once and for all lift in house prices in the NR site, so the price premium is a constant (β in equation 6 above). The trend in house prices is assumed to be unaffected and so T''-T''' remains parallel with C'''-C''''.

3.3.3 Adjusted interrupted time series (AITS) method

The AITS model is illustrated in Figure 5 above. The AITS model allows us to determine whether the impact of the NR program on house prices was merely a temporary change or more permanent. To derive a measure of the duration of the program impact, we must be able to observe the price trend in both the treatment and control areas and draw comparisons between the two. Suppose that the price trends were common in control and treatment areas before the introduction of the NR program. A-A' represents the price trend in the treatment sample and C-C' represents the price trend in the control areas. The introduction of the NR program is assumed to coincide with a favourable change in market conditions so that house prices increase at a faster rate in the control, e.g. C'-C". If the NR program has no impact, then the house price trend in the treatment area will be A'-A". The two separate potential sources of impact in the AITS model are illustrated by the line segments A'-D''' and D'-D". The segment D'-D" will eventuate if NR has a once and for all but permanent impact on the level of house prices in NR areas. D'-D'' is therefore drawn such that it is parallel with C'-C". The segment A'-D''' will eventuate if NR impacts favourably on the trend in house prices; note that it converges on C'-C'' as prices in NR sites accelerate faster than in the controls. Finally, if NR impacts both the level of and trend in house prices, the line segment will both shift up and have a steeper slope (not shown in Figure 5).

4 DESCRIPTIVE STATISTICS AND PRELIMINARY FINDINGS

This Section reports sample numbers from the treatment and baseline control sample as well as descriptive statistics on property values and neighbourhood and dwelling characteristics. Additional neighbourhood characteristics were added to the valuation dataset to enhance the amount of information available at neighbourhood level; derivation of these variables is explained in greater detail below. Finally, this Section reports some preliminary findings.

4.1 Descriptive statistics and sample numbers

Model estimates are based on transactions in residential property and vacant land transactions over 1990–2011 that are located within metropolitan Melbourne. The quasi-experimental method assigns transaction records into one of two groups: the treatment group, which includes transaction records for properties located within the boundaries of NR sites introduced from 2001–06,[35] and the baseline control group, which comprises property transactions that are located outside of NR site boundaries but within the SLAs that NR sites populate. Table 3 presents sample numbers and summary statistics for all property transactions in the treatment and baseline control areas. There are close to 141 600 transaction records in the sample dataset with 13 527 observations in the treatment group and 128 041 observations in the baseline control group. Mean prices for the treatment sample are around 30 per cent lower than mean prices in the baseline control sample, as is to be expected since NR sites have concentrations of public housing; consequently private housing in the NR site is likely to sell at a discount.

Table 3: Sample numbers and descriptive statistics for treatment and control groups

Group status	Count	Mean price (*thousands*)	Median price (*thousands*)	Standard deviation (*thousands*)
Treatment group	13,527	$181	$144	$131
Baseline control group	128,014	$234	$180	$184
Total	141,541	$229	$175	$180

Table 4 offers a more detailed description of individual NR sites that are grouped according to their vintage (date of introduction). Here, the treatment group is divided into the following three classifications:

1. property transactions that are located within the boundaries of NR programs introduced in 2002—namely, Braybrook and Maidstone and Collingwood;
2. property transactions located within the boundaries of NR sites in Ashburton, Ashwood and Chadstone, Broadmeadows, Werribee and Doveton-Eumemmerring. These sites were introduced in the following year 2003; and
3. properties within the boundaries of NR sites in West Heidelberg and Hastings introduced in 2006.

The baseline control sample comprises properties that lie outside of the NR site yet belong within the same SLA as the NR sites (see Table 1 for a list of the SLAs). In rows four and five we report the mean and median prices, respectively, for properties in the treatment and

[35] Assignment of the treatment and control groups will be different in the revealed preference method, a detailed explanation will be provided in the Final Report.

control groups. It can be seen from the records in row four that the average price of properties within an NR site boundary are lower than properties in the baseline control sample in all cases but one; properties within the boundary of the Broadmeadows NR site are priced slightly higher than those outside the boundary but in the same SLA.

Table 4: Breakdown of sample numbers and descriptive statistics by SLA

	Neighbourhood renewal sites	Count		Mean price		Median price	
		Treatment	Control	Treatment	Control	Treatment	Control
Properties near NR sites introduced in 2002	Braybrook and Maidstone	2,396	19,818	$203,770	$257,918	$178,000	$206,000
	Collingwood	89	7,163	$332,402	$362,036	$270,000	$300,000
Properties near NR sites introduced in 2003	Ashburton, Ashwood and Chadstone	1,988	20,788	$295,968	$372,936	$246,625	$300,000
	Broadmeadows	1,184	18,728	$167,776	$166,634	$149,975	$141,750
	Werribee	3,631	32,523	$142,393	$154,948	$120,000	$125,000
	Doveton-Eumemmerring	3,468	15,393	$144,053	$178,196	$115,000	$145,850
Properties near NR sites introduced in 2006	West Heidelberg	593	10,374	$163,874	$291,594	$135,000	$240,000
	Hastings	178	3,227	$180,880	$181,734	$119,000	$120,000
Total		13,527	128,014	$229,066	$229,069	$175,000	$175,000

The baseline control sample was constructed in such a way as to improve the comparability of the treatment and control groups by minimising any SLA-level differences between the two groups. But because NR sites are recognised as being areas of relative disadvantage, we nevertheless expect NR areas to have inferior accessibility to various local amenities compared with properties in the baseline control sample. Tables 5 and 6 below report neighbourhood amenities and locational attributes as well as some key property characteristics by treatment status (see Appendix 1 for definitions of key variables); and list descriptive statistics for continuous (indicator) variables. Transaction records include commercial properties, units and apartments, vacant land and residential houses.

Overall, accessibility measures are very similar for treatment and baseline control samples. However, it seems that properties within NR sites are typically more distant from the CBD (though closer to activity centres) than properties outside the boundaries of NR sites. From the property characteristic comparisons we find a large difference in floor area with smaller buildings a feature of properties transacted within NR sites.

Table 5: Descriptive statistics for continuous variables in the baseline dataset

Continuous variables	Mean		Median		Standard deviation	
	Treatment group	Control group	Treatment group	Control group	Treatment group	Control group
Sale Price	181,180	234,125	144,000	180,000	131,007	183,905
Distance to CBD (kms)	20.5	18.2	25.4	16.6	9.8	10.8
Distance to train station	1.9	2.3	1.7	1.6	1.9	3.0

Distance to activity centre	3.5	5.4	2.8	5.2	2.4	3.3
Distance to primary school	.69	.72	.65	.63	.36	.45
Distance to secondary school	3.7	3.8	1.5	1.6	12.9	13.8
Land area (square metres)	607	585	604	596	135	223
Floor area (square metres)	115	152	102	136	41.2	58.7
Number of bedrooms	2.9	3	3	3	.59	.68

Note: Sample numbers exclude extreme values at the 1st and 99th percentiles with respect to variables sales price and land area.

There are a couple of marked differences between the incidence of indicator variables in treatment and baseline control samples. The most noticeable is the environmental significance overlay variable. Only three property transactions in the treatment sample are located in an area subject to an environmental significance overlay, yet 1492 (1.2%) have this overlay status in the baseline control sample. Heritage overlays are also much more likely to apply in the locations where baseline control transactions occur (8.5% in baseline control sample, but only 1.3% in treatment sample).

Table 6: Descriptive statistics for indicator variables in the baseline dataset

Indicator variables	**Treatment**		**Control**	
	Count	**Frequency**	**Count**	**Frequency**
House	11,188	82.7%	100,456	78.5%
Land	1,320	9.8%	21,402	16.7%
Unit/Apartment	1,019	7.5%	6,156	4.8%
Residential zone dummy	13,491	99.7%	126,563	98.8%
Industrial zone dummy	1	.007%	237	.19%
Business zone dummy	29	.21%	358	.28%
Other zone dummy	6	.04%	880	.69%
Environmental significance overlay dummy	3	.02%	1,492	1.2%
Land subject to inundation overlay dummy	6	.04%	369	.29%
Heritage overlay dummy	181	1.3%	10,869	8.5%
Total number of observations	13,527		128,041	

Note: Sample numbers exclude extreme values at the 1st and 99th percentiles with respect to variables sales price and land area. 'Rural zone' and 'wildfire management overlay' dummy variables were also removed from the dataset as there were no observations for these variables in the treatment sample.

4.2 Preliminary findings

In this Section, we report some preliminary findings using the simple difference-in-differences method. We compare the average price of properties in the treatment sample before and after the NR sites were introduced, with the average price of those in the baseline control sample over the same sample period. This is performed by estimating the regression model specified in equation (6). Because the NR scheme was introduced in several stages over the sample period, we estimate three separate difference-in-differences models. The first estimate is based on NR sites that were introduced in 2002, namely,

Braybrook/Maidstone and Collingwood. The treatment group in this case is properties that are located within the boundary of the two NR sites, and the control group contains properties that lie outside the NR boundaries but within the same SLAs—these NR sites are located in (Maribyrnong & Yarra). Similarly, for the second and third models, the treatment group comprises property transactions within NR sites that were introduced in 2003 and 2006, respectively, and the control sample is formed from the relevant SLAs. As mentioned earlier, NR sites introduced after 2006 are not included in this research exercise as there are too few transactions in the post-treatment period and so estimates would be unreliable. Table 7 below presents preliminary findings for NR sites initiated in 2002, 2003 and 2006. The dependent variable is the natural logarithm of transaction prices and the 'right hand side' variables are a treatment dummy, a post-dummy to distinguish transactions before and after the introduction of NR and finally a treatment* post (difference-in-differences) interaction variable, which is the critical variable as far as NR impacts are concerned. Neighbourhood and structural characteristics are ignored in this preliminary investigation, but will be controlled for in the estimations presented in the Final Report. The findings reported in Table 7 should therefore be regarded as illustrative only.[36]

Table 7: Preliminary findings

Variables	Stage 1 (2002)	Stage 2 (2003)	Stage 3 (2006)
Treatment	-0.314*** (.014)	-.179*** (.008)	-.406*** (.027)
Post 2002	1.810*** (.006)	Na	na
Post 2003	na	1.468** (.004)	na
Post 2006	na	Na	1.548*** (.012)
Treatment*Post 2002	.036* (.020)	.031** (.013)	.0102 (.049)
Adjusted R-squared	0.5460	0.3507	0.3168
Number of Observations	29,466	97,703	14,372

Note: Standard errors in parentheses * denotes coefficient statistically significant at 10 per cent, two-tailed test; ** denotes coefficient statistically significant at 5 per cent, two-tailed test; *** denotes coefficient statistically significant at 1 per cent level, two-tailed test.
Regression results are based on a trimmed dataset which omits transactions lying in the 1st and 99th percentile of the price distribution due to extreme values.

The regressions endorse the view that NR is indeed responsible for external benefits. While properties located within NR boundaries sold for between 16 per cent and 33 per cent[37] less than properties in the control sample, this differential narrowed by between 1 per cent and 4 per cent in the period following introduction of NR. At the median price of $144 000 (in the NR sites), these estimates translate into price premiums of $1469–5184.[38] The price premium in the earliest (2002) cohort of NR sites is largest, but only weakly significant at 10

[36] In addition to property and neighbourhood characteristics, models will be augmented by a more sophisticated treatment of the time dynamics of house prices.

[37] The percentage change is obtained from $100^* \left[\exp\left(\hat{\phi}\right)-1\right]$. See Halvorsen and Palmquist (1980).

[38] Taking the lowest estimate ($1469) and summing over properties transacted in the NR sites, we obtain an aggregate estimate of close to $20 million. This ignores the price premiums in private properties not transacted, and excludes possible spillover (indirect) effects.

per cent. In the second cohort, the price premium is slightly lower in size, but statistically significant at 5 per cent. The most recent cohort has a statistically insignificant and small positive premium; as this model has the shortest post-treatment period we can expect less robust estimates. These are naïve models as they are based on a crude version of the control sample. In the second half of the project more refined techniques will be applied to generate robust impact estimates. The next steps in our program of research are now outlined.

5 NEXT STEPS

The preliminary findings reported in Chapter 4 above use a baseline control sample that includes all property transactions from the SLA that NR sites are located in. The second stage of this project will refine this approach by using the propensity score method to design a more robust matching sample of control transactions. Each transaction within the boundaries of NR sites will be matched (using the propensity score method) to that transaction belonging to the same SLA, but the closest match in terms of property and location characteristics. The quality of the match will be evaluated by the detection or otherwise of statistically significant differences between NR site transaction characteristics, and matched control transaction characteristics.

A standard difference-in-differences (DID) model specification will then be estimated in which all post-NR transactions within NR site (treatment) boundaries are identified by a dummy variable, and its coefficient estimate will be the key outcome measure (see equation 6 & Figure 4 above). The standard specification has a number of important features; first, it will detect whether the difference in average house price levels before introduction of NR change after their introduction, as will be the case if NR is responsible for a price premium that boosts relative house prices at the treatment sites. But NR might also impact the trajectory taken by house prices in the post-treatment period, not just the levels. Second, in the standard specification, price premiums due to NR are assumed to be contained within the boundaries of NR sites. But spillover effects to adjacent property are conceivable. Third, the NR program is in fact introduced on a staggered basis, with two sites first introduced in 2002, four sites in 2003, and two sites in 2006. The standard DID specification deals with the NR program as if all sites were implemented in the same year. We explore each of these features by introducing increasingly sophisticated measurement techniques, sample designs and model specifications.

5.1 The adjusted interrupted time series (AITS) model

The standard DID model will not accurately detect impacts if NR affects both the level and trajectory of house prices at NR sites. It is also unable to convey any information on the permanence or otherwise of NR impacts. These points are illustrated in Figures 4 and 5 above. Galster, Tatian and Pettit (2004) and Galster et al. (2004b) propose the AITSM model specification to detect relative changes in the trajectory of house prices. The second stage of the project will estimate a model specification along these lines with a view to shedding more light on the dynamics of NR impacts.

A policy change that is staggered over time rather than introduced on a one-off ‘overnight’ basis raises some challenges for impact evaluations. In the standard DID regression model it is common to identify post-intervention property transactions by the use of a single dummy variable. If all NR sites were introduced on the same date the use of this dummy variable is valid. But suppose the declaration of NR sites is staggered as in the Victorian case. The post-intervention transactions in the treatment and baseline control samples defined for early NR sites will then be spread over a different time period from those defined for later NR sites. But the standard DID dummy variable described above will fail to distinguish between these different post-treatment periods, and as a result impact estimates will be biased. We will follow the ‘tactic’ advocated by Bertrand, Duflo and Mullainathan (2004); in their paper a two-stage estimation method is advocated to address non-uniform before and after time periods. Details will be described in our Final Report.

5.2 Spillover (indirect) effects

The propensity score method of matching each transaction in NR sites with a control omit transactions outside NR sites but within 2000 feet of the boundaries of NR sites from the

baseline control sample. This is a deliberate research strategy as previous research studies (Galster et al. 1999; Santiago et al. 2001; Ellen & Voicu 2006; Castells 2010) have reported significant spillover effects beyond the boundaries of NR sites. Two questions will be addressed through modifications to the DID and AITSM model specifications. First, in the Victorian NR program can we detect significant spillover effects in neighbouring communities? Second, if spillover effects are detected, do they rapidly decay with distance from NR site boundaries? It has become common practice to uncover these indirect spatial effects by drawing a buffer around the edges of NR sites; following US protocols we will experiment with buffers that have different dimensions, but starting with the most commonly used—2000 feet. Transactions within this 2000 feet buffer will be tagged according to their distance from the boundary of the nearest NR site, and regressions models re-estimated to obtain relevant measures (see Galster, Tatian & Smith (1999) for a typical example of this approach).

5.3 Aggregating benefits and impacts on state government revenue streams

As is evident from the above description, considerable effort is devoted to the estimation of house price premiums that can be attributed to NR interventions. This is because premium estimates are the crucial bit of information needed in order to measure aggregate dollar measures of benefits/returns. In the standard DID model, the price premium estimate (assuming the detection of statistically significant premiums) will take the form of an average percentage lift in house prices. We will assume that any increase applied to both private property bought and sold in a post-intervention period, as well as other private residential property within NR sites but not transacted in that period. Dollar price increments will be computed using the percentage premiums, discounted (to translate into present value measures) and summed to obtain an aggregate measure of community benefits. The aggregation method ignores the public housing dwellings and common land areas (e.g. playgrounds, parks) in NR sites that have been the subject of renewal investment. But if we assume that such investment lifts the capital values of the assisted housing and land by an equivalent amount, our aggregate measure can be interpreted as a net return due to the external benefits of NR.

The price premium on properties bought and sold in NR sites after-intervention will in fact boost state government revenues from stamp duty. The overall net return estimate will be augmented by measurement of these increases in stamp duty revenues that represent revenue return to the government on its investment. Contemporaneous stamp duty schedules will be used to measure this potential boost to state government revenues.

5.4 The revealed preference approach

There is an alternative revealed preference approach to the measurement of NR impacts that exploits the staggered introduction of NR sites. As described earlier in this Positioning Paper, later cohorts of NR sites offer an appealing control sample of property transactions that is designed to include transactions within the boundaries of these sites, but before their introduction. From this baseline control the propensity score method can again be used to match each transaction in treatment samples with the transaction in the baseline control that is an optimal match in terms of property and location characteristics. The strength of this approach is the idea that NR sites will be selected according to similar if not unchanging criteria; transactions occurring within the boundaries of later cohorts of NR sites, but before their introduction, will therefore share many of the strengths we attribute to randomised experiments in medical research. The controls will share the same ‘disease’ as the treatment sample because they eventually receive the same treatment (NR). But the selection of property transactions before the introduction of NR ensures that they have yet to receive the treatment.

Estimates using this revealed preference approach will be reported in our Final Report. The method has at least one potentially serious drawback that does not afflict our SLA-based treatment versus control sample design; the period following introduction of the early cohort of NR sites 'treatment' but before the introduction of the later cohort of NR sites is necessarily limited. If the impacts of NR are cumulative and accrue slowly, the revealed preference method could underestimate impacts. We nevertheless plan to execute this research exercise as a useful illustration of an evaluation technique that might have helpful applications to a range of other housing and urban policy programs.

REFERENCES

Apel, RJ & Sweeten, G 2010, 'Propensity score matching in criminology and criminal justice'. in Alex R Piquero and David Weisburd (eds), Handbook of quantitative criminology, pp.543–562, Springer. New York.

Bair, E & Fitzgerald, JM 2005, 'Hedonic estimation and policy significance of the impact of HOPE VI on neighborhood property values', Review of Policy Research, vol. 22, no. 6, pp.771–786.

Becker, Sascha O & Ichino, Andrea 2002, 'Estimation of average treatment effects based on propensity scores', Stata Journal, StataCorp LP, November, vol. 2, no. 4, pp.358–377.

Bertrand, M. Duflo, E. and Mullainathan, S. (2004), How Much Should We Trust Differences-in-Differences Estimates?" Quarterly Journal of Economics, Vol. 119, No. 1, pp. 249-75

Black, D & Smith. J 2003, 'How robust is the evidence on the effects of college quality? Evidence from matching', Journal of Economics, vol. 121, no. 1, pp.99–124.

Briggs, X, Darden, JT, Aidala, A 1999, 'In the wake of desegregation: Early impacts of scattered-site public housing on neighbourhoods in Yonkers, New York', Journal of the American Planning Association, vol. 65, no. 1, pp.27–49.

Buckley, Jack & Shang, Yi 2003, 'Estimating policy and program effects with observational data: The "difference-in-differences" estimator', Practical Assessment, Research & Evaluation, vol. 8, no. 24, http://pareonline.net/getvn.asp?v=8&n=24 Caliendo, M & Kopeinig, S 2008, 'Some practical guidance for the implementation of propensity score matching', Journal of Economic Surveys, vol. 22, no. 1, pp.31–72.

Castells, N 2010, 'HOPE VI neighborhood spillover effects in Baltimore', Cityscape: A Journal of Policy Development and Research, vol. 12, no. 1, pp.65–97.

Chapman, PG 1993, The economics of training, LSE Handbooks in Economics, Harvestor, Wheatsheaf, New York and London.

Cigdem M 2012, Do brothels affect house prices? Using quasi-experimental estimators, (mimeo paper), Latrobe University, Melbourne, 2012.

Dehejia, RH, & Wahba, S 2002, 'Propensity score-matching methods for nonexperimental causal studies', The Review of Economics and Statistics, vol. 84, no. 1, pp.151–161.

Department of Human Services. 2002. Neighbourhood Renewal- Growing Victoria Together.

Office of Housing, Department of Human Services, Victoria.

Department of Human Services, Victoria 2008, Neighbourhood renewal: Evaluation report 2008, Neighbourhood Renewal Unit, Department of Human Services, Victoria.

Department of Human Services, Victoria 2005, Neighbourhood renewal: Evaluation report 2005, Neighbourhood Renewal Unit, Department of Human Services, Victoria.

Ellen, IG, Schill, MH, Schwartz, AE & Susin, S 2002, Building homes, reviving neighbourhoods: Spillovers from subsidised construction of owner-occupied housing in New York City, Journal of Housing Research, no. 12, pp.185–206.

Ellen, IG & Voicu, I 2006, 'Nonprofit housing and neighborhood spillovers', Journal of Policy Analysis and Management, vol. 25, no. 1, pp.31–52.

Galster, G, Tatian, P & Smith, R 1999, The impact of neighbours who use section 8 certificates on property values', Housing Policy Debate, vol. 10, no. 4, pp.879–917.

Galster, G, Tatian,P & Pettit, K 2004a, 'Supportive housing and neighbourhood property value externalities', Land Economics, vol. 80, no. 1, pp.33–54.

Galster, G, Temkin, K, Walker, C & Sawyer, N 2004b, Measuring the impacts of community Development initiatives: A new application of the adjusted interrupted time-series model, Evaluation Review, vol. 28, no. 6, pp.502–538.

Gruber, J 2000, 'Disability, insurance benefits and labour supply', The Journal of Political Economy, vol. 108, no. 6, pp.1162–1183.

Glaeser, E 2011, Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier, The Penguin Press, New York.

Glaeser, EL & Gottlieb, JL 2008, 'The economics of place-making policies', Brookings Papers on Economic Activity, no. 1, pp.155–253.

Groenhart, L. E., (2007), 'Economic Modelling of Public Housing Estate Renewal, European Network of Public Housing Estate Renewal, ENHR International Conference, 25-28 June, Rotterdam.

Halvorsen, R & Palmquist R 1980, 'The interpretation of dummy variables in semilogarithmic equations', American Economic Review, vol. 70, no.3, pp 474–75.

Hughes, M 2004, Community economic development and public housing estates, Shelter New South Wales, Sydney.

Johnson, G, Parkinson, S, Tseng, Y & Kuenle, D 2011, Long-term homelessness: Understanding the challenge—12 months outcomes form the Journal to Social Inclusion pilot program, Sacred Heart Mission, St Kilda.

Judd, B & Randolph B 2006, 'Qualitative methods and the evaluation of community renewal programs in Australia: Towards a national framework', Urban Policy and Research, vol. 24, no.1, pp.97–114.

Kelaher, M. Warr, DJW & Tacticos, T 2010, 'Evaluating health impacts: Results from the neighbourhood renewal strategy in Victoria, Australia', Health and Place, no. 16, pp.861–867.

Ki, CO & Jayantha, WM 2010, 'The effects of urban redevelopment on neighbourhood housing prices', International Journal of Urban Sciences, vol. 14, no. 3, pp 276–294.

Klein, H 2004. 'Neighbourhood renewal: Revitalising disadvantaged communities in Victoria', Public Administration Today, September-November, pp.20–29.

Levy, D. (2012). HOPEVI in Smith, S., (General Editor) International Encyclopaedia of Housing and Home, Elsevier, Amsterdam.

Malpezzi, S. (2003), Hedonic Pricing Models: A Selective and Applied Review in O'Sullivan, T. and Gibb, K. Housing Economics and Public Policy, Oxford, Blackwell.

Meyer, BD 1995, 'Natural and quasi-experiments in economics', Journal of Business & Economic Statistics, vol. 13, no. 2, pp.151–161.

Newell, T. A. 2010. 'Development and neighborhood revitalization: The effects of residential investment on property values in Durham, NC', The Michigan Journal of Business, Vol. 3, no. 2: 97-120.

Piquero, A & Weisburd, D (eds) 2010, Handbook of quantitative criminology, Springer, New York.

Randolph, B. & Judd, B. (2006) 'The Role of Qualitative Analysis in the Evaluation of Community Renewal Programs'. Urban Policy and Research, Vol 24, No 1. pp 97-114.

Rosenbaum, PR & Rubin, DB 1983, 'The central role of the propensity score in observational studies for causal effects', Biometrika, vol. 70, no. 1, pp.41–55.

Rossi-Hansberg, E., Sarte, P-D. and Owens, R., (2010) Housing Externalities Journal of Political Economy, Vol. 118, no. 3: 485-535.

Rubin, D 1977, 'Assignment to a treatment group on the basis of a covariate', Journal of Educational Statistics, no. 2, pp.1–26.

Santiago, AM, Galster, G & Tatian, P 2001, 'Assessing the property value impacts of the dispersed subsidised housing program in Denver', Journal of Policy Analysis and Management, vol. 20, no. 1, pp.65–88.

Schill, MH, Ellen, IG, Schwartz, AE & Voicu, I. 2002, 'Revitalizing inner-city neighborhoods: New York City's Ten Year Plan for Housing', Housing Policy Debate, vol. 13, no. 3, pp.529–566.

Schwartz, AE, Ellen, IG, Voicu, I & Schill, MH 2006, 'The external effects of place-based subsidised housing', Regional Science and Urban Economics, no. 26, pp.679–707.

Shadish, WR, Cook, TD & Campbell, DT 2002, Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin, Boston, MA.

Shield, M, Graham, M & Taket, A. 2011, 'Neighbourhood renewal: An effective way to address social inclusion', Journal of Social Inclusion, vol. 2, no. 2, pp. 4-18.

Smith, J & Todd, P 2005, 'Does matching overcome Lalonde's critique of nonexperimental estimators?', Journal of Econometrics, vol. 125, no. 1-2, pp.305–353.

Wholey, JS, Hatry, NP & Newcomer, KE (eds).2010, The handbook of practical program evaluation, Third Edition, Jossey Bass, San Francisco.

Wood, G, Ong, R, Cigdem, M & Taylor, E 2012, The spatial and distributional impacts of the Henry Review recommendations on stamp duty and land tax, Australian Housing and Urban Research Institute, Final Report, no 182, Melbourne.

Wulff, M & Reynolds, M 2011, Housing, inequality and the role of population mobility, Australian Housing and Urban Research Institute, Final Report, no 135, Melbourne.

Zhao, Z 2004, 'Using matching to estimate treatment effects: Data requirements, matching metrics, and Monte Carlo evidence', The Review of Economics and Statistics, vol. 86, no. 1, pp.91–107.

Zielenbach, S, Voith, R & Mariano, M 2010, 'Estimating the local economic impacts of HOPE VI', Housing Policy Debate, vol. 20, no. 3, pp.485–522.

APPENDICES

Appendix 1: List of key variables in the baseline dataset

Table A1: List of key variables in the baseline dataset

Variable name	Definition	Measurement
Y	Dependent variable, log of the sales price of land plot or property.	Nominal dollars.
Dtreatment	Dummy variable indicating properties lying within the NR site.	Equal to 1 if property sold is within the NR sites, regardless of whether sale occurs before or after its introduction; zero otherwise.
DPost_2002	Dummy variable indicating properties that are transacted after the first stage of the NR program.	Equal to 1 if property is sold in 2002 onwards, zero otherwise.
DPost_2003	Dummy variable indicating properties that are transacted after the second stage of the NR program.	Equal to 1 if property is sold in 2003 onwards, zero otherwise.
DPost_2006	Dummy variable indicating properties that are transacted after the third stage of the NR program.	Equal to 1 if property is sold in 2006 onwards, zero otherwise.
C	Vector of continuous and dummy variables capturing structural and locational characteristics (see below).	See below.
Number of bedrooms (log)	Continuous variable indicating number of bedrooms contained in each sold property.	Log of number of bedrooms variable.
ICSEA score	Continuous variable indicating the Index of Community Socio-Educational Advantage (ICSEA) value.	Linear value.
Age of building (log)	Continuous variable indicating the age of the building in years.	Log of age of building variable.
SLA	Vector of dummies indicating the statistical local area that each property transaction belongs to.	Equal to 1 if property is in SLA *x*, zero otherwise.
Distance to CBD (log)	Continuous variable indicating distance from property *i* to the CBD.	Log of distance to the CBD in km.
Distance to train station (log)	Continuous variable indicating distance from property *i* to the nearest train station.	Log of distance to nearest train station in km.
Distance to activity centre (log)	Continuous variable indicating distance from property *i* to the nearest principal or major activity centre.	Log of distance to nearest activity centre in km.
Distance to primary school (log)	Continuous variable indicating distance from property *i* to the nearest state primary school.	Log of distance to nearest primary school in km.
Distance to secondary	Continuous variable indicating distance from property *i* to the nearest	Log of distance to nearest

Variable name	Definition	Measurement
school (log)	state secondary school.	secondary school in km.
Land size (metres squared) (log)	Continuous variable indicating size of the property.	Log of the size of the land plot in hectares.
Rural zone dummy	Dummy variable indicating properties located in area that is zoned for rural development.	Equal to 1 if the property is in an area zoned as residential, zero otherwise (omitted category).
Residential zone dummy	Dummy variable indicating properties located in area that is zoned for residential development.	Equal to 1 if the property is in an area zoned as residential, zero otherwise.
Industrial zone dummy	Dummy variable indicating properties located in area that is zoned for industrial development.	Equal to 1 if the property is in an area zoned as industrial, zero otherwise.
Business zone dummy	Dummy variable indicating properties located in area that is zoned for commercial/business development.	Equal to 1 if the property is in an area zoned as commercial/business, zero otherwise.
Other zone dummy	Dummy variable indicating properties in an area that is zoned for other land uses (e.g public use zone, comprehensive development zone etc.).	Equal to 1 if the property is in an area zoned as comprehensive development zone, road zone, public park and recreation zone and special use zone, zero otherwise.
Environmental significance overlay dummy	Dummy variable indicating properties with environmental significance.	Equal to 1 if land is in area regarded as environmentally significant, zero otherwise.
Land subject to inundation overlay dummy	Dummy variable indicating property in an area prone to flooding.	Equal to 1 if land is in flood area, zero otherwise.
Heritage overlay dummy	Dummy variable indicating areas regarded as places of natural, historical or cultural significance.	Equal to 1 if land is in heritage area, zero otherwise.

Appendix 2: Outline of the propensity score method

The second methodological approach that is used to identify a comparison group is the commonly used matching estimators. Unlike randomised experiments where the treatment and comparison groups are drawn from the same population, matching methods allow for the potential comparison group to be drawn from a different population to the treatment group so long as the observed pre-treatment covariates are the same for the two groups. To determine this, the propensity score method first estimates the probability that an individual is assigned to the treatment group, given a set of observed characteristics *X*, and then matches the treated to the comparison group on the probability of receiving treatment (i.e. the propensity scores). Where $D = 1$ for the treated population, Rosenbaum and Rubin (1983) define the propensity score as the following:

(8) $$p(X) = prob(D = 1 \mid X),$$

where $D = 1\ (= 0)$ represents the treated (control) group. The propensity score can be obtained parametrically through such methods as a logit or probit model.

To obtain a true estimate of the population treatment effect for the treatment however, a critical assumption in the propensity score model is that of the ignorable treatment assignment. This assumption posits that, conditional on the observable pre-treatment variables, assignment of study units to treatment and comparison groups is independent of their potential outcomes. To convey this proposition algebraically, let Y_0represent the outcome of the untreated group, and Y_1 represent the outcome for the treated group. Rubin (1977) proposes that if covariates X can be observed for each individual and

(9) $$(Y_0, Y_1) \| D \mid X,$$

where $\|$ represents statistical independence, then the population treatment effect for the treatment can be obtained by estimating the treatment effect conditional on covariates and on assignment to treatment, $\tau \mid_{D=1}$, averaged over the distribution $X \mid D = 1$ (Dehejia & Wahba 2002, p.152). Markedly, this is one of the differences between the nonrandomised method and the randomised method in that the latter method includes all covariates, both observable and unobservable, in the assignment process (Rosenbaum & Rubin 1983, p.43). However, as pointed out by Dehejia and Wahba (2002, p.153), when conditioning on observable covariates, assignment of the treatment can be treated as a random process like the randomised experiment where the treated group and the counterfactual group are equally disposed to the intervention given the similarity in their observable characteristics. The inference here is that comparing two individuals of different treatment status, but subject to the same observable characteristics, is essentially like comparing those two individuals in a randomised experiment (Dehejia & Wahba 2002, p.153). It follows then that the conditional treatment effect, $\tau \mid_{D=1}$, can be obtained by first obtaining an estimate of $\tau \mid_{D=1,X}$, and then averaging over the treated group with covariates *X* (Dehejia & Wahba 2002, p.153).

A second assumption in the propensity score model is the 'common-support assumption' (Zhao 2004, p.92) which asserts that the probability of assignment of the treatment to the treated is positive conditional on the covariates. That is,

(10) $$0 < prob(D = 1|X) < 1$$

Assumption 5.16 implies that, for all observable characteristics X, the probability of either being treated ($D = 1$) or being untreated ($D = 0$) is positive, which further suggests that a match can be identified for all individuals where $D = 1$ (Smith & Todd 2005, p.313).

According to Rosenbaum and Rubin (1983), treatment assignment is 'strongly ignorable' when these two assumptions are satisfied, intimating that, in principle, experimental and non-experimental designs yield the same parameters under these conditions. Incorporating these assumptions into the ATT equation presented in Section 3.3.1, we get:

(11)

$$\Delta_{TT} = E_{x|D=1}\{E[Y_1|D=1, X=x] - E[Y_0|D=1, X=x]\}$$
$$= E_{x|D=1}\{E[Y_1|D=1, X=x] - E[Y_0|D=0, X=x]\}.$$

Substituting equation (8) defining propensity scores into equations (9) and (10), the first assumption becomes:

(12) $$(Y_0, Y_1) \perp\!\!\!\perp T \,|\, p(X)$$

and the second assumption becomes:

(13) $$0 < prob(T = 1 \,|\, p(X)) < 1.$$

Subject to these two assumptions, Zhao (2004) presents the formula for estimating the treatment effect of the treated, ΔTT, which takes the following form when incorporating the propensity score:

(14)

$$\Delta_{TT} = E_{p|D=1}\{E[Y_1|D=1, X=x], p(X)=p] - E[Y_0|D=1, p(X)=p]\}$$
$$= E_{p|D=1}\{E[Y_1|D=1, p(X)=p] - E[Y_0|D=0, p(X)=p]\},$$

where the estimates for the first term can be derived from the treatment group and those in the second term derived from the mean outcomes of the matched (on X) control group (Smith & Todd 2005, p.313).

However, given that the propensity score $p(X)$ is a continuous variable, the probability of two separate units generating precisely the same propensity score is zero (Becker & Ichino, 2002, p.4). Hence, the propensity score alone is not sufficient to estimate the effect of a treatment on the treatment group. The objective then becomes to implement appropriate matching methods to achieve the next best propensity score match between the treated and control units. There are several matching algorithms that have been developed to exploit the propensity score method. The methods most commonly used include nearest-neighbour matching, radius matching, kernel matching and stratification matching. These estimators contrast in their demarcations of the neighbourhood for treated units and the weights assigned to them, and also the way in which they approach the common support problem. For the sake of relevance, nearest neighbourhood matching will be discussed below, as this is the technique that will be later implemented.

Nearest neighbour matching

To estimate the treatment effect for the treated, the nearest-neighbour matching method finds control units that are closest to the treatment unit and constructs a match, conditional on the propensity score. Using the notation of Becker and Ichino (2002), we can express this algebraically. Let C denote the properties in the control group, and i denote the property in the treated group with an estimated propensity score value $p_i C(i)$therefore denotes the set of control properties matched to the treated property i, and is set such that only one control unit j is selected for every treated unit i with an estimated propensity score of p_i, unless there are multiple nearest neighbours:

(15) $$C(i) = min \parallel p_i - p_j \parallel$$

The control unit with a propensity score p_j that is closest to p_i is selected as a match.

Nearest-neighbour matching provides users with the option of matching with or without replacement, as well as allowing for matching with more than one nearest neighbour. Nearest-neighbour matching with replacement imposes no restrictions on the number of times the same control unit is matched to a suitable treated unit, whereas matching without replacement requires that a control unit be eliminated from consideration after every match. When performing nearest-neighbour matching without replacement, it is often suggested that the order in which matching is performed is of primary importance as it determines the final estimates. Users are therefore advised to sort observations randomly prior to matching.

It is also suggested that using more than one nearest neighbour per treated unit, or 'oversampling', enhances estimates as it reduces the variance. However, this again involves a trade-off between a lower variance resulting from the additional observations used to construct the control group, and a larger bias, which is the result of generally weaker matches. When applying oversampling, it is up to the researcher to specify the number of matching partners allowed per treated unit, and the amount of weight that is assigned to them.

One of the main drawbacks of the nearest-neighbour matching scheme is that all units in the treatment group receive a control match, regardless of the size of the discrepancy between the propensity scores, which can potentially result in some very weak matches. As a consequence, the nearest neighbour matching method is often coupled with at least one other method so as to alleviate the shortcomings of the former.

AHURI Research Centres

Queensland Research Centre
RMIT Research Centre
Southern Research Centre
Swinburne-Monash Research Centre
Sydney Research Centre
UNSW-UWS Research Centre
Western Australia Research Centre

AHURI
Australian Housing
and Urban Research Institute

Australian Housing and Urban Research Institute
Level 1 114 Flinders Street, Melbourne Victoria 3000
Phone +61 3 9660 2300 Fax +61 3 9663 5488
Email information@ahuri.edu.au Web www.ahuri.edu.au