Spatial Processes in Environmental Economics: Empirics and Theory

Amy Binner

For the Degree of Doctor of Philosophy

University of East Anglia

School of Environmental Sciences

December 2012

ABSTRACT

Economic activities are fundamentally influenced by their location in space, which determines the physical and natural environment in which they take place. Likewise, location defines the social context of economic activity prescribing the particular laws, regulations and social norms to which it should conform. Moreover, spatial location defines proximity, which shapes the costs of accessing factor inputs, product markets and other economic and social institutions. In fact, spatial location mediates most forms of interaction, intended and unintended, that may arise from communication and connections between economic agents. These spatial processes have important implications for estimation, policy evaluation and prediction in models of economic activity.

This thesis is comprised of two parts. Part I presents a broad range of issues that arise in estimation due to space and frames these as general spatial omitted variables. I explore the use of semi-parametric estimators to identify the parameters of interest in this general model and derive identification conditions for fixed and local adaptive spatial smoothing estimators. The properties of these estimators are contrasted to OLS and spatial econometric estimators.

Part II addresses issues in policy evaluation and prediction. I derive an equilibrium sorting model with endogenous tenure choice that can be used to evaluate the general equilibrium welfare effects of policies that affect local environmental quality. Using a series of simulations, motivated by a real world policy application, I contrast the welfare changes derived under this model to a conventional static approach. By allowing for rental and purchase markets the model I develop provides a far richer characterisation of the complex adjustments that propagate through the property market following policy changes and the contrary impact such policies can have upon renters and owners. The usefulness of the model for applied policy analysis is demonstrated through two applications: The Polegate Bypass and Mortgage Interest Deduction reform.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Economic activities are fundamentally influenced by their location in space, which determines the physical and natural environment in which they take place. Likewise, location defines the social context of economic activity prescribing the particular laws, regulations and social norms to which it should conform. Moreover, spatial location defines proximity, which shapes the costs of accessing factor inputs, product markets and other economic and social institutions. In fact, spatial location mediates most forms of interaction, intended and unintended, that may arise from communication and connections between economic agents. Furthermore, space also influences the way in which economic activities are measured, for example the spatial scale that information is collected at and the process of spatial aggregation (e.g. to census track, county and country levels) introduce a spatial dimension into economic data. Likewise, data can be affected by spatially correlated measurement error due to localised differences in the tools that are used to measure economic activity and their drivers. For example, local authority procedures may lead to regional differences in the classification of categorical variables and scientific equipment may be sensitive to localised fluctuations in wind speed, temperature or machine calibration. This suggests that the analysis of economic activities is affected by numerous spatial processes. These spatial processes have important implications for both the estimation of economic models and economic policy evaluation. This thesis examines these two issues sequentially.

Chapter 1 presents an examination of the issue of identifying the parameters of models with an omitted spatial data generating process (DGP). I identify a number of theoretical foundations for spatial processes and characterise these through a set of spatial data generating processes, which I show can each be represented as an additive omitted spatial variable in a general linear model. I then contrast the conventional estimation approach provided by spatial econometric models to the alternative semi-parametric approach of spatial smoothing. I derive a set of conditions under which a spatial smoothing estimator is able to identify the parameters of the model and discuss the efficiency of the estimator. The properties of a local polynomial regression based spatial smoothing estimator are illustrated through a series of Monte Carlo simulations, which demonstrate the ability of the estimator to deal with a broad class of spatial-DGPs. The chapter provides the first derivation of a set of clearly defined identification conditions that are also presented in terms of the underlying data in environmental economics.

Chapter 2 builds on the work of chapter 1 to present an intuitive extension to the SSE. In particular, I develop a locally adaptive spatial smoothing estimator (LASSE) in which the smoothing parameters are specified at each location in space using an optimality condition. The paper demonstrates when the LASSE is more efficient than a Fixed-SSE estimator and derives a relaxed set of conditions under which the parameters of the model are identified. While the formal properties of the LASSE are not derived analytically, I present the results of a Monte Carlo analysis demonstrating its properties under a variety of spatial-DGPs. This work presents the first consideration of these issues in terms of parameter identification within the context of environmental economics.

Part II explores the issue of policy appraisal and predicting welfare changes in the presence of spatial processes. This part of the thesis concerns itself with economic theory, particularly the development of economic models in which the choice of spatial location is a fundamental component of economic activity. The particular focus is residential location decisions and their response to exogenous policy shocks. The modelling framework that I adopt for this purpose is that of Equilibrium Sorting Models (ESMs) (Kuminoff et al., 2010). In brief, the ESM I develop treats the property market as a series of distinct neighbourhoods. Each neighbourhood is differentiated by its level of environmental quality. Households, with varying characteristics, judge the quality of a neighbourhood not only on the basis of its environmental quality but also on the basis of the characteristics of the other households living in that neighbourhood. Using numerical algorithms, the model is solved by finding a set of property prices that match the supply of property in each neighbourhood with the demand for property in that neighbourhood. A solution to the model, therefore, describes both the set of neighbourhood property prices as well as a sorting of the heterogeneous households across neighbourhoods.

My research builds on the work of previous authors, particularly (Epple and Platt, 1998) extending those precedents in a number of key directions. In particular, the model I present is the first ESM to endogenise tenure choice. Households in my model can choose to rent or buy their homes according to conditions in parallel and interacting purchase and rental markets. I adapt the model to explore two public policy questions that together constitute the final two chapters of this thesis.

This thesis was part-funded by the UK Department for Transport (DfT). Many projects that the DfT undertakes result in localised changes in environmental quality. Examples include the construction of a new rail link or the building of a bypass around a rural town. Currently, those projects are evaluated by aggregating estimates of the willingness to pay (WTP) of the residents that are directly and immediately impacted by the local change in environmental quality. Chapter 3 demonstrates how a static analysis fails to capture the complex repercussions that emanate through the entire economy as a result of the local change. Over the medium term, rental and purchase prices adjust throughout the economy as households change their residential location and tenure decisions as a result of the new conditions in the property market. Accordingly, the motivating policy question for the first paper in the second part of my thesis concerns understanding how project evaluation using a static analysis compares to an analysis allowing for property market adjustments.

By allowing for rental and purchase markets my model provides a far richer characterisation of those differences. For example, I find that renters in neighbourhoods that experience environmental improvements may be disadvantaged as a result of rental prices rising in those neighbourhoods. In contrast, homeowners in such neighbourhoods enjoy an improvement in environmental quality and also benefit from capital gains that result from increases in the price of their homes. I adapt the model to examine how these distributional effects are impacted when I allow for the possibility that households in the community are also landlords and hence gain income from the rental payments of other households. The paper examines the key dynamics of the model in the context of a simple two-community model. Finally, in Appendix A a more complex seven-community model is developed and calibrated to the town of Polegate in East Sussex in 2001, the year in which a bypass was constructed round the town. The Polegate calibration is used to illustrate how the ESM developed in my thesis might be used to provide input in a real policy context.

Chapter 4 presents the second paper in Part 2. This chapter adapts the ESM with endogenous tenure choice to examine a policy question currently under debate in the USA: reform of Mortgage Interest Deduction (MID). The policy of MID allows taxpayers to subtract interest paid on a residential mortgage from their taxable income. MID is vigorously supported by the real estate industry on the grounds that it reduces the costs of purchasing properties and, as such, encourages homeownership. As well as homeownership being seen as a desirable outcome in its own right, it is also argued that a significant weight of empirical evidence supports the idea that increasing levels of homeownership creates positive spillovers, raising neighbourhood quality in a variety of ways (e.g. higher educational achievement, lower levels of crime etc.). In contrast, others, including the current administration, have argued that MID simply inflates the prices of properties, greatly favours wealthy homeowners and cannot be justified given the large federal deficit. In this paper I explore the MID debate using an ESM with simultaneous rental and purchase markets and endogenous tenure choice. The model is extended by additionally endogenising neighbourhood quality, such that the desirability of any particular neighbourhood is partly determined by its levels of homeownership. The public policy relevance of the model is shown through a calibration exercise for Boston, Massachusetts, which explores the impacts of various reforms to the MID policy. The simulations confirm some of the arguments made about reforming MID but also demonstrate how the complex patterns of behavioural change induced by policy reform can lead to unanticipated effects. The simulations suggest that it may be possible to reform MID whilst maintaining the prevailing rates of homeownership and reducing the federal budget deficit.

INTRODUCTION

This thesis is motivated by the importance of space, which is receiving growing attention from researchers in economics and is becoming an increasingly prevalent issue with the progressive availability and use of large spatially delineated datasets. Economic activities are fundamentally influenced by their location in space, which determines the physical and natural environment that surrounds them. Likewise, location defines the social context of economic activity prescribing the particular laws, regulations and social norms to which it should conform. Moreover, spatial location defines proximity, which shapes the costs of accessing factor inputs, product markets and other economic and social institutions. In fact, spatial location mediates most forms of interaction, intended and unintended, that may arise from communication and connections between economic agents.

The roots of spatial research stem from earlier developments in the regional science and urban economics literature, which sought to explain spatial spillovers, where outcomes in one location are related to outcomes in other proximate locations. This developed into the field of spatial econometrics, the expansion of which has shed light on a multitude of complexities that are introduced into the estimation and analysis of spatially organised data, hereafter referred to as "spatial data". In particular, through exploring the theoretical foundations of spatial models, the literature has simultaneously highlighted the potential spatial data generating processes that underpin and are embedded in spatial data. The interesting and challenging issue is that these spatial data generating processes are a feature of spatial data whether they themselves are of primary interest or not.

The field of spatial econometrics was conceived over thirty years ago. In his recent review Anselin (2010) examines the evolution of the subject, dividing it into three phases: growth, take-off and maturity. The growth phase is taken to have begun roughly thirty years ago following the publication of Paelinck and Klassen's (1979) volume *Spatial econometrics*, Bartels and Ketellapper's (1979) *Exploratory and explanatory statistical analysis of spatial data* and Bennett's (1979) *Spatial time series*.

The early work of Paelinck and Klassen (1979) specified five principles to guide the formulation of spatial econometric models. These five principles illustrate potential issues associated with space, which continue to apply to the analysis of spatial data and provide a concise summary of the multifaceted importance of space. The five principles are: i) the role of spatial interdependencies, ii) the asymmetry in spatial relations, iii) explanatory factors located in other spaces, iv) differentiation between *ex post* and *ex ante* interactions and v) explicit modelling of space.

Although this work was originally conceived and interpreted within the realm of regional science models, as the literature has developed it has become evident that spatial processes have far reaching implications and the same challenges are faced by all analysts dealing with spatial data. In this thesis I draw on the characterisation of spatial issues from the spatial econometrics literature and consider it in the context of the analysis of general economic models. In this vein, it becomes clear that space has wide reaching implications for the analysis of many economic models. Drawing upon Paelinck and Klassen's five principles, spatial processes arise in economic models when i) the economic agents interact with each other across space, ii) there are asymmetric spatial effects for example through regional policies and jurisdictional boundaries, iii) there are unobserved or immeasurable spatial explanatory variables, iv) economic outcomes in one area may influence outcomes in another area and moreover the spatial environment may be endogenously determined by behaviour and activities in the area and v) the economic activity itself occurs in a geographically defined space.

Anselin (2010) categorises the contributions of spatial econometrics over the last thirty years into four key groups, i) specification of spatial models, ii) estimation, iii) specification testing and iv) prediction. Significant advances have been made in each of these areas over the last thirty years. Consequently, awareness and popularity of spatial econometrics has increased amongst analysts in regional science and broader fields of economics. However, it is argued that there is a gap between the theoretical models and applied problems and, moreover, that spatial parametric models have been adopted as a solution with little theoretical guidance, thus making the results difficult to interpret and justify (Gibbons and Overman, 2012). The Journal of Regional Science, a journal which has been at the forefront of promoting spatial econometrics, recently celebrated its 50^{th} anniversary issue with a focus on reviewing spatial econometrics (Partridge et al., 2012) including papers by McMillen (2012), Corrado and Fingleton (2012) and Gibbons and Overman (2012) reflecting upon the current condition of the literature and consider how spatial econometrics should be used. These papers, along with McMillen (2010), Pinkse and Slade (2010) and Brady and Irwin (2011) consider the limitations of the current spatial econometric methods when it comes to practical applications.

Following from this, what is needed now is a greater synergy between the requirements of applied analysis and theoretical developments. This involves marrying the insights and tools provided by spatial econometrics with knowledge from other fields, such as non-parametric estimation, and theoretical considerations. In this thesis I focus on the issues of estimation and prediction in the presence of spatial processes within the remit of environmental economics. To achieve this, the thesis is organised in two parts: Part I deals with the issue of parameter identification through the estimation of spatial models. Part II compliments this by addressing the issue of policy appraisal and prediction in spatial models.

PART I

INTRODUCTION TO PART I

Part I of this thesis considers the estimation challenges that arise in the presence of unaccounted-for spatial data generating processes (DGPs). This research was motivated by the problem of parameter identification in the presence of omitted spatially correlated variables. The importance of this spatial context is evident across the spectrum of empirical economic applications including studies of growth (Foster and Rosenzweig, 2003), human capital (Corneo and Jeanne, 1999), infrastructure (Baum-Snow, 2007), competition between firms (Berry, 1994, Davis, 2006, Epifani and Gancia, 2006) product demand (Jones, 2005), location choice (Hoff and Sen, 2005), economic migration, school choice (Gibbons and Machin, 2006, Harris and Johnston, 2008), crime rates (Sah, 1991, Anselin et al., 2000, Puech, 2004) and health (McIntosh, 2008). In environmental economics, studies of land use depend on spatially organised data such as soil types, climate, water quality and proximity to market (Leggett and Bockstael, 2000). Spatial factors such as the proximity to amenities and dis-amenities as well as the characteristics of neighbouring households are important in the analysis of residential property prices (Kim and Goldsmith, 2008, Grimes and Liang, 2009). The location of emitters and receptors of pollutants (Zabel and Kiel, 2000, Day et al., 2004, Day et al., 2007, Bayer et al., 2009), and likewise the location of environmental amenities, such as parks, beaches, mountains and rivers, and their users are important factors in valuation studies and are all inherently spatial (Rehdanz and Maddison, 2008, Hoshino and Kuriyama, 2010, Waltert et al., 2011). It is evident therefore that the challenges posed by spatial-DGPs have far reaching influences. In fact, any analysis that relies upon spatially organised data is vulnerable to omitted spatial variable problems and other spatial data problems, which can undermine the estimation of economic models.

When I began my research I was inspired by a desire to understand the potential to utilise the semi-parametric approaches to address spatial data problems and hoped to contrast these approaches to the alternative spatial econometric models developed in the spatial econometrics literature (Anselin, 1988, 2004). I began by examining three semi-parametric approaches: matching, difference in differences (DID) and spatial smoothing estimation (SSE).

The principle of matching is to group observations according to a set of explanatory variables for which no parametric form is specified (Rubin, 1973, Gibbons and Machin, 2003, 2006). The first application of matching was provided by Belson (1956) who employed the method to derive the treatment effect of exposure to a number of television broadcasts. Since then matching has been applied to a variety of issues, for example, Rubin (1973) examines automobile accident severity with and without seatbelts, Rosenbaum (1986) explores US high school dropouts, Ichino et al (2006) examine the impact of temporary work agency on future employment in Italy and Brodaty et al (2000) use kernel matching to evaluate youth employment schemes.

This literature is closely associated with analyses that adopt a difference in differences (DID) approach, which considers the difference in outcomes between an otherwise equivalent treatment and control group. Examples of the DID estimation approach are most frequently presented by treatment effect studies in health economics and policy evaluation, one of the most well-known studies is Card and Krueger's (1994) evaluation of the effect of New Jersey's minimum wage. More recently, Gibbons and Machin (2003) examine DID methodologies in the valuing of English primary schools, Smith and Todd (2005) adopt a DID approach to assess the National Supported Worker demonstration, Autor and Hausman (2006) use it to examine the impact on earnings of temporary agency placements and Wagstaff et al (2005) compares a pure DID approach to a DID with matching approach in an assessment of the World Bank's Health VIII Project.

Finally, this led to my consideration of spatial smoothing estimators which have been applied, for example, by Gibbons and Machin (2003) for valuing English primary schools, Kneib Muller and Hothorn (2006) for assessing habitat suitability for birds in Northern Bavarian forests and Day et al (2007) in assessing the impact of noise pollution. This early research allowed me to demonstrate that both the matching and difference in difference approaches are constrained versions of a spatial smoothing estimator. This led to a focusing of the thesis upon a comparison of spatial smoothing estimators (SSEs) with spatial parametric estimators (SPEs).

To compare the performance of these two approaches I sought to explore a range of spatial-DGPs that present challenges to the analysis of spatial data. In doing so, I defined a more general framework which encompasses a broad range of spatial-DGPs including spatial econometric models (such as might result from spatial interactions, peer group effects etc.), spatial misspecification, spatial measurement error and omitted spatial variables.

Driven by a desire to enable SSEs to be employed with confidence, I directed my research towards developing a deeper understanding of the conditions under which identification is achievable and present this in a clear, concise and accessible manner for applied researchers. In the context of a partial linear model (Robinson, 1988) the first chapter develops a set of sufficiency conditions for identification of the linear parameters of the model in the presence of a variety of spatial-DGPs. This led to the development of the two chapters presented in Part I.

The first chapter demonstrates how a broad class of spatial-DGPs can be treated as an unknown additive element in a linear model. I derive a set of conditions under which a SSE returns unbiased parameter estimates in the presence of that broad class of spatial-DGPs. I interpret these conditions in the context of a local polynomial regression based SSE. The performance of this estimator in small samples is demonstrated through a series of Monte Carlo simulations.

Through this work it became clear that the smoothing parameters play a crucial role in determining the bias elements associated with spatial smoothing and, as a result, are central to understanding the conditions for identification. This motivated me to explore more recent work on selecting local smoothing parameters (Fan and Gijbels, 1995). In chapter 2, I adapt the methodology derived by Fan and Gijbels to develop a Local Adaptive SSE (LASSE). I examine the properties of a LASSE in comparison to a Fixed SSE and demonstrate that the revised estimator has greater efficiency properties and moreover, it presents additional opportunities for identification. I derive a set of identification conditions for the LASSE with a local polynomial kernel density estimator. Finally, I explore the properties of the LASSE estimator and compare it to a Fixed SSE using a series of Monte Carlo simulations.

CHAPTER 1

Addressing spatial dependence in linear models using semi-parametric estimation

1.0 Introduction

Economic activities are fundamentally influenced by their location in space. Location determines the physical and natural environment in which those activities take place. Location defines the social context of economic activity prescribing the particular laws, regulations and social norms to which it should conform. Location defines proximity, shaping the costs of accessing factor inputs, product markets and other economic and social institutions, as well as the apparatus of law and government. In fact, spatial location mediates most forms of interaction, intended and unintended, that may arise from communication and connections between economic agents.

In empirical analyses of economic data, the multitude of paths through which location can influence economic activity can be problematic. In particular, parameter bias is likely to arise should an analyst's econometric specification fail to fully account for the array of spatial processes generating the data.

The possibility of unaccounted-for spatial processes in economic data has long been recognised. The extensive literature on spatial econometrics (Anselin, 2010), for example, focuses on one possible form for such spatial processes. In spatial econometric specifications, outcomes in one location are assumed to be partly determined by variables at nearby locations, where 'nearby' is identified by some analyst-defined spatial weights matrix and the spatially-lagged regressors are any combination of dependent variables (spatial lag models), explanatory variables (spatial cross-regressive models) and error terms (spatial error models).

Since the 1990s spatial econometric models (SEMs) have seen widespread application in the analysis of spatially organised economic data (see Anselin et al (2004) for a review and Anselin (2010) for a discussion of the evolution of the field). While the field has continued to evolve (Brady and Irwin, 2011) , recent literature has struck a rather more critical tone (McMillen, 2010, Pinkse and Slade, 2010, Gibbons and Overman, 2012, Partridge et al., 2012). At the heart of that criticism is the concern that spatial lag processes are a rather specific form of spatial data generating process. The routine application of SEMs to deal with spatial dependence in economic data may not be appropriate, particularly when there is no strong theoretical justification for suspecting a spatial lag process to be at work in the data. Moreover, even if the assumption of a spatial lag process is correct, the reliance on an analyst-defined spatial weights matrix and a parametric functional form for the spatial lag process opens up a significant possibility of misspecification.

The first contribution of this paper is to demonstrate that the spatial processes envisaged by a large variety of SEMs can be recast as an unknown additive element in a linear model. In addition, the paper demonstrates that numerous other spatial processes that might otherwise confound parameter identification in linear models also share this same basic form. More importantly, the paper demonstrates that this broad class of spatial data-generating processes is amenable to estimation using a semi-parametric estimator that I term a Spatial Smoothing Estimator (SSE). Accordingly, the paper contends that, in many circumstances, there are good reasons to prefer application of the SSE over the application of SEMs.

The SSE is an application of Robinson's (1988) partial linear model, which uses the data itself to determine the nature of any unaccounted-for spatial processes. The SSE has seen previous application in the economic literature, for example, in the valuation of English primary schools (Gibbons and Machin, 2003, 2006) and noise avoidance (Day, 2005, Day et al., 2007). Building on the work of Paciorek (2010), a second contribution of this paper is to provide a systematic investigation of the SSE, formally identifying the set of conditions under which the estimator is able to identify parameters of interest in the presence of unaccounted-for spatial processes of different forms. The performance of the estimator is illustrated through a series of Monte Carlo experiments.

2.0 Unaccounted-for Spatial Processes

Consider a standard estimation problem in which an analyst has spatially organised data relating measures of a dependent variable, denoted by the vector Y, and a set of independent variables given by the $N \times k$ data matrix $\boldsymbol{X}$, where the variables in $\boldsymbol{X}$

are a function of spatial location, $\boldsymbol{S}$, and possibly other non-spatial factors, $\boldsymbol{\eta}$**.** The focus in this paper is on the standard linear model which, given the available data, the analyst may specify as,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + \epsilon \tag{1.1}$$

Where ϵ is an error term. The objective of the analysis is assumed to be the identification of the parameters $\boldsymbol{\beta}$.

Without loss of generality $\boldsymbol{X}$ can be represented by its orthogonal decomposition. Hence, $\boldsymbol{X}$ can be expressed as additively separable functions of location and non-spatial elements, $\boldsymbol{\eta}$, which is orthogonal to space, $\boldsymbol{S}$, such that,

$$\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta}) = \boldsymbol{a}(\boldsymbol{S}) + \boldsymbol{\eta} \tag{1.2}$$

Where $\boldsymbol{a}(\boldsymbol{S})$ is a vector function mapping $\boldsymbol{S}$ into the k dimensions of $\boldsymbol{X}$. By definition,

$$E[\boldsymbol{\eta}'\boldsymbol{S}] = 0 \tag{1.3}$$

Without loss of generality it is also assumed that,

$$E[\boldsymbol{\eta}] = 0 \tag{1.4}$$

Such that the means of $\boldsymbol{X}$ are treated as constants in the spatial functions, $\boldsymbol{a}(\boldsymbol{S})$.

2.1 Spatial Econometric Models (SEMs)

While (1.1) is amenable to estimation using OLS, the analyst may be concerned that the application of that estimator may be confounded by the presence of some unaccounted-for spatial data-generating process. Such a diagnosis may be prompted by the application of one of a raft of post-estimation specification tests suggested by the spatial econometrics literature (Anselin, 1988). Following the recommendations of that literature, the analyst may proceed by fitting some particular form of SEM.

While there are very many different SEMs, here I focus on two of the most popular specifications: the spatial lag and spatial error models (see, for example, Pace and Giley, (1997), Kelejian and Prucha, (1998)).[1] These SEMs have the general form,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\beta + \lambda\boldsymbol{W}Y + \varphi(\boldsymbol{M}) + \epsilon \tag{1.5}$$

In a model with N observations, $\boldsymbol{W}$ and $\boldsymbol{M}$ are $N \times N$ spatial weights matrices and λ is a scalar known as a spatial multiplier. In a spatial weights matrix the ij^{th} element, being the element in the i^{th} row and the j^{th} column, describes the relationship between observations i and j. For example, in a binary spatial weight matrix the ij^{th} element would equal one when observations i and j are in the same neighbourhood. Alternatively, the spatial weights matrix may record the distance between observations. The spatial weights matrix $\boldsymbol{W}$ constitutes the spatial autoregressive lag (SAR) component of the model and $\varphi(\boldsymbol{M})$ is a spatial error dependence component defined by the spatial weights matrix $\boldsymbol{M}$.[2]

2.2 Spatial Misspecification

Frequently, the spatial elements of (1.1) require analysts to specify some particular form for the spatial process generating the data. An obvious example concerns the choice of spatial weights matrices, $\boldsymbol{W}$ and $\boldsymbol{M}$, in (1.5). Should the elements of the weights matrix be binary or should they somehow increase with greater proximity? Should that proximity be measured by straight line distance, travel time or perhaps travel cost? Faced with such a diversity of possibilities, misspecification raises a serious concern.

[1] Here I focus on spatial lag and spatial error models. One may also wish to consider a model with a lagged subset of independent variables $\lambda\boldsymbol{W}\boldsymbol{X_s}$, where $\boldsymbol{X_s} \in \boldsymbol{X}$.

[2] The models can be estimated using Maximum Likelihood or Generalized Method of Moments (Kelejian and Prucha, 1998; Fingleton, 2008).

Consider, for example, a true underlying spatial data generating process (DGP) which is characterised by equation (1.5). Should the spatial weights matrices be incorrectly specified as $\widehat{\boldsymbol{W}}$ and $\widehat{\boldsymbol{M}}$, then that equation can be rewritten as,

$$Y = \boldsymbol{X}\beta + \lambda\widehat{\boldsymbol{W}}Y + \varphi(\widehat{\boldsymbol{M}}) + \lambda\left(\boldsymbol{W} - \widehat{\boldsymbol{W}}\right)Y + (\varphi(\boldsymbol{M}) - \varphi(\widehat{\boldsymbol{M}})) + \epsilon \tag{1.6}$$

The two latter elements $\left(\boldsymbol{W} - \widehat{\boldsymbol{W}}\right)$ and $(\varphi(\boldsymbol{M}) - \varphi(\widehat{\boldsymbol{M}}))$ are not observed.

2.3 Spatial Measurement Error

Another potential issue is that spatial elements of the model may be measured with error. Of particular concern here is when that measurement error is itself a function of space, for example, when estimates of a regressor are measured more accurately in densely populated regions than in sparsely populated ones.

Consider the case where the true variable of interest, $\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})$, is measured with spatially correlated error as $\widehat{\boldsymbol{X}}(\boldsymbol{S},\boldsymbol{\eta})$. The spatial data generating process can then be rewritten as,

$$Y = \widehat{\boldsymbol{X}}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + \left(\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta}) - \widehat{\boldsymbol{X}}(\boldsymbol{S},\boldsymbol{\eta})\right)\boldsymbol{\beta} + \epsilon \tag{1.7}$$

Where in this case $\left(\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta}) - \widehat{\boldsymbol{X}}(\boldsymbol{S},\boldsymbol{\eta})\right)$ is unobserved.

2.4 Omitted Spatial Variables

Given the myriad spatial processes that influence economic activity, a final concern with specifying a model for spatially organised data is the omission of relevant spatial variables. That omission may occur either because those variables are: i) overlooked, ii) assumed to be irrelevant or iii) impractical or impossible to measure.

In this case the true spatial data generating process is,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + Z(\boldsymbol{S},\nu) + \epsilon \tag{1.8}$$

Where $Z(\boldsymbol{S},\nu)$ represents the confounding spatial variable or variables that have been omitted from the econometric specification of the model. Notice that as well as being dependent on location the values of $\boldsymbol{X}$ and Z may also be functions of non-spatial elements.

2.5 A General Form for Spatial Processes

It is straightforward to show that the SEM specifications in (1.5), the spatial misspecification illustrated in (1.6) and the spatial measurement error shown in (1.7) are all simply particular cases of the omitted spatial variable model in equation (1.8).

Table 1 frames each of the four spatial-DGPs in the context of equation (1.8).

DGP	$\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta}$	$Z(\boldsymbol{S},\nu)$
Spatial Parametric Models: SAR SARAR	$\boldsymbol{X\beta}$	$Z = \Sigma_{j=1}^{\infty}\lambda^j\boldsymbol{W}^j\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + \Sigma_{j=1}^{\infty}\lambda^j\boldsymbol{W}^j\epsilon$ $Z = \Sigma_{j=1}^{\infty}\lambda^j\boldsymbol{W}^j\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + (\Sigma_{j=0}^{\infty}\Sigma_{l=1}^{\infty}\lambda^j\boldsymbol{W}^j\varphi^l\boldsymbol{M}^l + \Sigma_{l=1}^{\infty}\varphi^l\boldsymbol{M}^l)\epsilon$
Spatial Misspecification	$\boldsymbol{X\beta} + \widehat{\boldsymbol{W}}\boldsymbol{Y}$	$Z = \lambda(\boldsymbol{W} - \widehat{\boldsymbol{W}})Y + (\varphi(\boldsymbol{M}) - \varphi(\widehat{\boldsymbol{M}}))$
Spatial Measurement Error	$\widehat{\boldsymbol{X}}\boldsymbol{\beta}$	$Z = \left(\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta}) - \widehat{\boldsymbol{X}}(\boldsymbol{S},\boldsymbol{\eta})\right)\boldsymbol{\beta}$
Omitted Spatial Variable	$\boldsymbol{X\beta}$	$Z = b(\boldsymbol{S}) + \nu$

Table 1.1: Spatial Data Generating Processes

In the same manner as for X, Z can be represented its orthogonal decomposition without loss of generality. Hence, Z can be expressed as additively separable functions of location and non-spatial elements, ν, that are orthogonal to space, $\boldsymbol{S}$, such that,

$$Z(\boldsymbol{S}, v) = b(\boldsymbol{S}) + v \tag{1.9}$$

By definition,

$$E[v'\boldsymbol{S}] = 0 \tag{1.10}$$

Such that it must also be the case that,

$$\begin{aligned} E[\boldsymbol{\eta}' b(\boldsymbol{S})] &= E[\boldsymbol{\eta}] = 0 \\ E[v' \boldsymbol{a}(\boldsymbol{S})] &= E[v] = 0 \end{aligned} \tag{1.11}$$

Without loss of generality it is also assumed that,

$$E[v] = 0 \tag{1.12}$$

Such that the mean of Z is treated as a constant in the spatial component $b(\boldsymbol{S})$.

For simplicity and to ensure that the model is identifiable, I make the following assumption,

- **Assumption 1:** The non-spatial variables in $\boldsymbol{X}$ are orthogonal to the non-spatial variables in Z,[3]

$$E[\boldsymbol{\eta}' v] = 0 \tag{1.13}$$

This prevents omitted variable bias of a non-spatial nature.

3.0 Estimation with Omitted Spatial Variables

[3] Note that this assumption does not imply that each vector within $\boldsymbol{\eta}$ is independent across observed variables, simply that they are not correlated with ν.

Let us now turn our attention to the issue of estimating parameters when the data are generated by equation (1.8). In this section I consider the estimation of models with a broad range of omitted spatial variables using OLS, SEMs and SSE. In each case the objective is to identify the parameters, $\boldsymbol{\beta}$. For each estimator, the conditions under which these parameters can be identified are explored.

For simplicity let us assume that the error, ϵ, conforms to the following standard assumptions,

- **Assumption 2**: The independent identically distributed innovations, ϵ, are mean zero,

$$E[\epsilon] = 0 \tag{1.14}$$

- **Assumption 3:** The independent innovations, ϵ, are homoskedastic and non auto-correlated,

$$E[\epsilon\epsilon'] = \sigma_\epsilon^2 I \tag{1.15}$$

- **Assumption 4:** The observed data, $\boldsymbol{X}$**,** and the confounding spatial process**,** Z**,** are independent of the innovations, ϵ,

$$\begin{aligned} E[\boldsymbol{X}'\epsilon] &= 0 \\ E[Z'\epsilon] &= 0 \end{aligned} \tag{1.16}$$

3.1 Ordinary Least Squares (OLS)

As is well known, if Y, $\boldsymbol{X}$ and Z are perfectly observed, then unbiased and efficient estimates of the parameters, $\boldsymbol{\beta}$, can be recovered through the application of the OLS estimator. Unfortunately, when, for some reason, Z is omitted from the estimating equation, bias in the parameter estimates may result.

Consider the parameter estimates obtained using OLS,

$$\begin{aligned}\widehat{\boldsymbol{\beta}} \quad &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\,\boldsymbol{X}'Y \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + Z + \epsilon)\end{aligned} \tag{1.17}$$

Which under assumption 4 becomes,

$$\widehat{\boldsymbol{\beta}} \;\; = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'Z \tag{1.18}$$

The expected bias in the OLS parameter estimate is therefore given by,

$$\begin{aligned}bias(\widehat{\boldsymbol{\beta}}) \quad &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\,\boldsymbol{X}'Z \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{a}(\boldsymbol{S}) + \boldsymbol{\eta})'(b(\boldsymbol{S}) + \nu)\end{aligned}$$

(1.19)

Which, from equation (1.11) becomes,

$$= (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{a}(\boldsymbol{S})'b(\boldsymbol{S}))$$

The bias is non-zero when the spatial components of observed data $\boldsymbol{X}$ are correlated with the confounding spatial data generating process Z, $E[\boldsymbol{a}(\boldsymbol{S})'b(\boldsymbol{S})] \neq 0$.

When more than one observed variable is included in the analysis, equation (1.19) consists of a number of multiple regression coefficients and the direction of the bias is complex to deduce. Note that bias may be introduced into each of the parameter estimates, $\widehat{\boldsymbol{\beta}}$, not just those relating to the observed variables that are correlated with relevant confounding spatial process, Z (Greene, 2003). As such, omitted spatial variables can cause bias in the estimates of parameters relating to non-spatial variables as well as spatial ones, undermining all aspects of the analysis. As a result, the findings of empirical investigations of spatially organised economic data that neglect to consider the potential influence of confounding spatial processes should be interpreted with caution.

3.2 Estimating Spatial Econometric Models

An alternative to OLS is to estimate a SEM. In this paper I consider the two SEMs outlined in section (2.1),

1. The Spatial Auto-Regressive Lag Estimator (SAR),

$$SAR: Y = \boldsymbol{X\beta} + \lambda \boldsymbol{W} Y + \epsilon \tag{1.20}$$

2. The Spatial Auto-Regressive Lag with Auto-Regressive Error Estimator (SARAR),

$$SARAR: Y = \boldsymbol{X\beta} + \lambda \boldsymbol{W} Y + \varphi \boldsymbol{M} u + \epsilon \tag{1.21}$$

Where ϵ and u are independent, identically distributed innovations. [4]

The parameters of these models can be estimated using either maximum likelihood or the general method of moments (GMM). The latter has been shown to be more robust to misspecification error (Kelejian and Prucha, 1998, 1999) and is the procedure used in the Monte Carlo analyses reported subsequently.

One of the most significant challenges for identification of $\boldsymbol{\beta}$ when estimating SEMs comes from the potential for misspecification. In many applications there is often little guidance for selecting an appropriate SEM (Bell and Bockstael, 2000). At best, analysts can combine theoretical justifications and insights with extensions of the LM-test to guide a choice between potential models (a discussion of these tests is provided in Anselin (1988), and the performance of classical testing approaches against a number of SEMs is provided in Florax and Folmer (1992)). In many existing applications the motivation for adopting a particular specification over others is not discussed (Leggett and Bockstael, 2000, Choumert and Salanié, 2008) and, as a result, the chosen spatial weights matrices may be a poor approximation to the true underlying spatial dependence structure (Pinkse and Slade, 2010).

---

[4] Comparable results were also obtained for the AR, MA and SARMA models and estimators. These results are also available from the author.

This is problematic because the asymptotic properties of SEMs rely on the assumption that the models are correctly specified (McMillen, 2010). When this assumption is violated, ambiguity surrounds the interpretation of the results. As was shown in the previous section, in the context of addressing spatial data problems, misspecification simply transforms the problem rather than overcoming it (Lee, 2008). Kelejian and Prucha's (1998, 1999) general methods of moments (GMM) estimation routine has been shown to produce parameter estimates that are more robust to misspecification error (Bell and Bockstael, 2000). Similarly, an alternative approach is to allow the spatial weights to be non-parametric (Pinkse and Slade, 2010), for example by estimating them through series expansion. This more flexible approach allows the data to determine the restrictions placed upon the functional form of spatial dependence. However, unlike the spatial smoothing estimator, which I will develop in the following section, the nonparametric estimation of spatial weights relies upon the correct specification of the structure of the model. In particular, the analyst must determine whether the model is a spatial auto-regressive lag or spatial error component model. The spatial weights of this specified model are then estimated nonparametrically and depend upon the standard conditional independence assumption on the error term. Misspecification of the model structure or the presence of further omitted spatial variables violates this assumption. As a consequence the resulting parameter estimates continue to be at risk of bias.

3.3 Spatial Smoothing Estimator (SSE)

While the analyst may have little information on the exact form of omitted spatial processes generating the data, one thing that is known is that those spatial processes will have a similar impact on each of the observations in a particular region in space. The SSE exploits this commonality. In essence, the SSE proceeds by estimating, using a nonparametric estimator, that part of the data which an observation holds in common with other observations in its environs. That nonparametric estimate is then subtracted from the data itself so as to sweep out the effects of spatial processes held in common across observations. The intention is that this spatial differencing will remove from the data any confounding omitted spatial processes.

4.0 Parameter Identification with Spatial Smoothing

The mechanics of the SSE can be seen by taking expectations of equation (1.8) with respect to space and subtracting this from the original equation. Subtraction produces the transformed relationship,

$$\begin{aligned} Y - E\,[Y|\boldsymbol{S}] = {} & (\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta}) - E\,[\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})|\boldsymbol{S}])\boldsymbol{\beta} \\ & + \left(Z(\boldsymbol{S},\nu) - E\left[\left(Z(\boldsymbol{S},\nu)\right)\middle|\boldsymbol{S}\right]\right) + \epsilon - E\,[\epsilon|\boldsymbol{S}] \end{aligned} \tag{1.22}$$

Where $E[.\,|\boldsymbol{S}]$ is hereafter referred to as a conditional spatial expectation. Replacing $\boldsymbol{X}$ and Z with their orthogonal decompositions I arrive at,

$$\begin{aligned} Y - E[Y|\boldsymbol{S}] = {} & (\boldsymbol{a}(\boldsymbol{S}) + \boldsymbol{\eta} - E[\boldsymbol{a}(\boldsymbol{S}) + \boldsymbol{\eta}|\boldsymbol{S}])\boldsymbol{\beta} \\ & + (b(\boldsymbol{S}) + \nu - E[(b(\boldsymbol{S}) + \nu)|\boldsymbol{S}]) + \epsilon - E[\epsilon|\boldsymbol{S}] \end{aligned} \tag{1.23}$$

Since $\boldsymbol{a}(\boldsymbol{S})$ and $b(\boldsymbol{S})$ are purely functions of space and $\boldsymbol{\eta}$ and ν are orthogonal to space $\boldsymbol{a}(\boldsymbol{S}) + \boldsymbol{\eta} - E[\boldsymbol{a}(\boldsymbol{S})|\boldsymbol{S}] = \boldsymbol{\eta}$ and $b(\boldsymbol{S}) + \nu - E[b(\boldsymbol{S})|\boldsymbol{S}] = \nu$. Then from assumption one, and assuming the conditional spatial expectations are known perfectly, the model reduces to,

$$Y - E[Y|\boldsymbol{S}] = \boldsymbol{\eta\beta} + \nu + \epsilon \tag{1.24}$$

Note that the process of spatial differencing reduces the observed regressors, $\boldsymbol{X}$, to their non-spatial components $\boldsymbol{\eta}$ and the omitted spatial process, Z, to its non-spatial component ν. Since by assumption 1 $E[\boldsymbol{\eta}'\nu] = 0$, spatially differencing the data removes the correlation between $\boldsymbol{X}$ and Z thereby removing the source of bias. It follows that, provided $\boldsymbol{\eta} \neq 0$, unbiased estimates of $\boldsymbol{\beta}$ can be recovered through simple OLS regression of the spatially differenced data.

The development provided above does not, however, tell the full story of the SSE. For a start, it assumes that conditional spatial expectations are known, when in reality these must be estimated from data. It also suggests that the best strategy is to seek an unbiased estimate of those conditional spatial expectations. As I shall show subsequently, however, under certain circumstances using biased estimates of those expectations may actually improve the efficiency of the SSE. It may also allow for identification of $\boldsymbol{\beta}$ even when $\boldsymbol{\eta} = 0$.

4.1 Empirical Estimation of Conditional Spatial Expectations

In practice, conditional spatial expectations of the dependent and independent variables, Y and $\boldsymbol{X}$, with respect to space are not known and instead must be estimated empirically. A number of different nonparametric estimation techniques (for example, local polynomial regression or smoothing splines), can be used to construct an estimate of these expected values. Each of these methodologies depends upon some set of smoothing parameters, which determine how the estimate is constructed. Let us represent a non-parametric empirical estimate of the conditional expectation with respect to space as,

$$\hat{E}_{\phi}[.\,|\boldsymbol{S}] \tag{1.25}$$

Where ϕ denotes the smoothing parameters. Subtracting this estimated conditional expectation from equation (1.8) produces the transformed relationship,

$$Y - \hat{E}_{\phi}[Y|\boldsymbol{S}] = (\boldsymbol{X} - \hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}])\boldsymbol{\beta} + (Z - \hat{E}_{\phi}[Z|\boldsymbol{S}]) + \epsilon \tag{1.26}$$

Where $(Z - \hat{E}_{\phi}[Z|\boldsymbol{S}])$ can be interpreted as a transformed unobserved variable. Under some conditions, equation (1.26) can be estimated using OLS to identify $\boldsymbol{\beta}$ parameters. Let us now consider the set of conditions under which this is possible.

4.2 Sufficient Conditions for Parameter Identification

As was observed by Robinson (1988), unbiased estimates of the parameters of the model can be identified using an SSE if and only if,

$$E[(\boldsymbol{X} - \hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}])'(Z - \hat{E}_{\phi}[Z|\boldsymbol{S}])] = 0 \tag{1.27}$$

AND

$$E\left[\left(\boldsymbol{X}-\ \hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}]\right)'\left(\boldsymbol{X}-\ \hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}]\right)\right] \mathit{is\ positive\ definite} \tag{1.28}$$

Robinsons first condition, (1.27), indicates that spatially smoothing the data must remove the correlation between $\boldsymbol{X}$ and Z thereby removing the source of omitted variable bias. The second condition indicates that the spatially differenced data must retain sufficient variation across the sample to allow for unbiased identification of the $\boldsymbol{\beta}$ parameters.

To understand better these conditions, it is useful to consider the relationship between the spatial variables and their conditional spatial expectations. To do this let us rewrite each variable as a linear combination of: i) its non-spatial components, ii) an empirical estimate of the conditional spatial expectation $\hat{E}_{\phi}[.|\boldsymbol{S}]$ and iii) a component, $f(.|\boldsymbol{S},\phi)$, capturing the bias between the true conditional spatial expectation, $E[.|\boldsymbol{S}]$, and the estimate provided by $\hat{E}_{\phi}[.|\boldsymbol{S}]$,

$$\begin{gathered}\boldsymbol{X}=\boldsymbol{\eta}+\hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}]+\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)\\ Z=\nu+\hat{E}_{\phi}[Z|\boldsymbol{S}]+\boldsymbol{f}(Z|\boldsymbol{S},\phi)\end{gathered} \tag{1.29}$$

Taking the conditional spatial expectations to the left hand side,

$$\begin{gathered}\boldsymbol{X}-\hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}]=\boldsymbol{\eta}+\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)\\ Z-\hat{E}_{\phi}[Z|\boldsymbol{S}]=\nu+\boldsymbol{f}(Z|\boldsymbol{S},\phi)\end{gathered} \tag{1.30}$$

and substituting these into equation (1.28),

$$E[(\boldsymbol{\eta}+\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi))'(\nu+\boldsymbol{f}(Z|\boldsymbol{S},\phi)]=0 \tag{1.31}$$

And by expansion,

$$E[\boldsymbol{\eta}'\nu]+E\left[\boldsymbol{\eta}'\ \boldsymbol{f}(Z|\boldsymbol{S},\phi)\right]+\boldsymbol{E}\left[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'\ \nu\right] \tag{1.32}$$

$$+E\left[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'\,\boldsymbol{f}(Z|\boldsymbol{S},\phi)\right] = 0$$

To remove the correlation between the observed data and the confounding spatial process it is necessary that there exist some smoothing parameters, ϕ, such that the smoothed observed data, which consists of the bias component from the estimated expectation and the non-spatial elements of $\boldsymbol{X}$, and the smoothed confounding spatial data process, which also consists of a bias component and non-spatial elements of Z, are no longer correlated. Under assumption 1 and by definition the first three terms in equation (1.32) are equal to zero, which leaves the following condition for identification,

$$E[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'\boldsymbol{f}(Z|\boldsymbol{S},\phi)] = 0 \tag{1.33}$$

Notice that this condition is expressed purely in terms of the bias in the empirical estimate of the conditional spatial expectations. As was discussed in section 4.0, when the conditional spatial expectation of $\boldsymbol{X}$ is known accurately the spatial component of $\boldsymbol{X}$ can be smoothed out. In this case equation (1.26) can be estimated by OLS without bias and equation (1.33) is met. Likewise, if unbiased empirical estimates of the conditional spatial expectations are available then identification is possible.

Moreover, it is clear from equation (1.33) that there are actually three possible ways to satisfy the expression,

IC1. There exists some ϕ whereby the empirical estimate of the conditional spatial expectation of $\boldsymbol{X}$ is unbiased,

$$\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi) = 0 \tag{1.34}$$

In this case the only variation remaining in $\boldsymbol{X}$ once the data has been spatially smoothed is provided by $\boldsymbol{\eta}$.

IC2. There exists some ϕ whereby the empirical estimate of the conditional spatial expectation of Z is unbiased,

$$\boldsymbol{f}(Z|\boldsymbol{S},\phi) = 0 \tag{1.35}$$

In some cases it may be possible to remove the spatial component of the confounding spatial data without removing all of the spatial variation in $\boldsymbol{X}$, thus leaving additional information which increases the efficiency with which the parameter estimates can be made.

IC3. There exists some ϕ whereby the empirical estimate of the conditional spatial expectations of $\boldsymbol{X}$ and Z are biased but the biases are uncorrelated,

$$\begin{gathered} E[(\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'(\boldsymbol{f}(Z|\boldsymbol{S},\phi)] = 0 \\ \boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi) \neq 0 \\ \boldsymbol{f}(Z|\boldsymbol{S},\phi) \neq 0 \end{gathered} \tag{1.36}$$

Under this final condition, it may be possible to remove the correlated spatial components of $\boldsymbol{X}$ and Z whilst leaving even further variation in $\boldsymbol{X}$, further increasing efficiency.

5.0 Identification using Local Polynomial Regression

One method for estimating the conditional spatial expectations is by means of local polynomial regression (LPR). LPR provides an estimate of the conditional spatial expectation at any particular location by plotting a local polynomial of order ρ through neighbouring observations. What constitutes 'neighbouring' is determined by the bandwidth parameter, h, in conjunction with the kernel function, K, which weights each observations contribution to the estimation of the expectation. Accordingly, for LPR,

$$\hat{E}_{\phi}[\boldsymbol{X}|\boldsymbol{S}] = \hat{E}_{K_{\{h,\rho\}}}[\boldsymbol{X}|\boldsymbol{S}] \tag{1.37}$$

The LPR estimator is defined by the kernel, K, the smoothing bandwidth, h, and the order of local polynomial regression, ρ.

The order of local polynomial regression, ρ, determines the order of the polynomial plotted through the observations. Local constant ($\rho = 0$) estimation refers to the case where the average value is taken, local linear ($\rho = 1$) refers to the value obtained by plotting a straight line through the values and so on. The choice of the order of local polynomial estimation affects identification in conflicting ways. Choosing a higher order reduces the bias in the estimation of the confounding spatial data process, Z, and thus reduces the bias in the parameter estimates. On the other hand it also reduces the amount of variation left in the smoothed observed data, increasing the variance of the parameter estimate.

5.1 Bias in Kernel Density Estimation

The accuracy with which LPR can recover conditional spatial expectations depends primarily on how well the bandwidth, h, and order of local polynomial, ρ, allow the LPR to capture the curvature of the spatial function at each location in space. Intuitively, contracting the bandwidth puts greater weight on more proximate observations in estimating the expectation at some location. As a result smaller bandwidths tend to reduce bias. At the same time, however, in reducing the quantity of data upon which the estimate is made, smaller bandwidths also increase the variance of that estimate (Ruppert et al., 1995, Boente and Rodriguez, 2008). Similarly, increasing the order of the local polynomial allows the local regression to approximate the curvature of the spatial function more accurately and, as such, reduces bias in the estimate. That increased flexibility comes at the cost of estimating more parameters from limited data that, in turn, increases the variance of the estimate.

Accordingly, in LPR the analyst must trade off bias and variance when selecting the bandwidth parameter or the order of the local polynomial fitted to the data. For this purpose, a number of selection criteria have been developed (See Hardle and Marron (1985), Mallows (1973), Akaike (1974), Moran (1950), Lee (2003), Lee and Solo (1999) and Hall et al.(1992)). These selection criteria include leave one out cross validation, partitioned cross validation, Speckman's rule of thumb, exact plug-in, asymptotic plug-in and bootstrapping (Francisco-Fernández and Vilar-Fernández, 2005). In practical applications the analyst with fix either the bandwidth or the order of polynomial regression and use one of the selection criteria above. These selection criteria are derived from various measures of best fit and do not explicitly consider the issue of parameter bias. As an alternative, it is possible to utilise Taylor series approximation to construct polynomial approximations of **X** and the residuals, $\hat{\epsilon}$, from OLS regression of Y on **X**. Under the two identification conditions derived in this paper it is then possible to construct a set of hypotheses that place restrictions on the coefficients on the higher order terms of these polynomial approximations which can be tested using the delta method. The results of these tests determine the order of local polynomial regression that is required to attain unbiased parameter estimates. In practice, this test can be combined with a selection criterion for choosing the bandwidth to minimise the mean squared error (Binner and Day, 2010).

Our primary concern with regards to understanding the identification conditions in (1.34). (1.35) and (1.36) in the context of LPR is the issue of bias. Ruppert and Wand (1994) derive general expressions for that bias. Consider, for illustration, the orthogonal decomposition of one of the variables in $\boldsymbol{X}$,

$$X = a(\boldsymbol{S}) + \eta \tag{1.38}$$

For a model of this type the bias in conditional spatial expectations is given by,

For even $\boldsymbol{\rho}$:

$$\begin{aligned} f(a(\boldsymbol{S})|\boldsymbol{S}, \rho, h) &= a(\boldsymbol{S}) - E_{K_{\{h,\rho\}}}[a(\boldsymbol{S})|\boldsymbol{S}] \\ &= \int u^{\rho+2} K_\rho(u) du \left[\frac{a^{\rho+1}(\boldsymbol{S}) g'(\boldsymbol{S})}{g(\boldsymbol{S})(\rho+1)!} + \frac{a^{\rho+2}(\boldsymbol{S})}{(\rho+2)!}\right] h^{\rho+2} + o_\rho(h^{\rho+2}) \end{aligned} \tag{1.39}$$

For odd $\boldsymbol{\rho}$:

$$
\begin{aligned}
f(a(\boldsymbol{S})|\boldsymbol{S},\rho,h) &= a(\boldsymbol{S}) - E_{K_{\{h,\rho\}}}[a(\boldsymbol{S})|\boldsymbol{S}] \\
&= \int u^{\rho+1} K_\rho(u)du \left[\frac{a^{\rho+1}(\boldsymbol{S})}{(\rho+1)!}\right] h^{\rho+1} + o_\rho(h^{\rho+1})
\end{aligned}
\tag{1.40}
$$

Where $a^\rho(\boldsymbol{S})$ denotes the ρ^{th} derivative of $a(\boldsymbol{S})$ and $g(\boldsymbol{S})$ is the probability distribution of $\boldsymbol{S}$. Using equations (1.39) and (1.40), Table 1.2 presents the approximate bias for local constant, local linear (Racine, 2001) and local quadratic estimation. These expressions show how the magnitude of the bias depends on the derivatives of the function $a(\boldsymbol{S})$ (the source of the curvature-based bias), the probability distribution of $\boldsymbol{S}$, $g(\boldsymbol{S})$ (the source of the boundary-based bias), and the order of local polynomial estimation, ρ. The choice of bandwidth also affects the magnitude of the bias, scaling the bias introduced as a result of selecting too small an order of local polynomial regression.

Estimator	Bias
Local Constant (Nadaraya-Watson)	$h_n^2\left(\frac{1}{2}a''(\boldsymbol{S})+\frac{a'(\boldsymbol{S})g'(\boldsymbol{S})}{g(\boldsymbol{S})}\right)\int u^2k(u)du$
Local linear	$h_n^2\left(\frac{1}{2}a''(\boldsymbol{S})\right)\int u^2k(u)du$
Local Quadratic	$h_n^4\left(\frac{1}{2}a''''(\boldsymbol{S})+\frac{a'''(\boldsymbol{S})g'(\boldsymbol{S})}{3!\,g(\boldsymbol{S})}\right)\int u^4k(u)du$

Table 1.2: Pointwise bias in kernel density estimation with second order kernels

Two sources of bias, boundary-based and curvature-based bias, are evident in equations (1.39) and (1.40) (Ruppert and Wand, 1994). Boundary-based bias is driven by changes in the density and marginal density of $\boldsymbol{S}$ at the boundaries of the data range, seen through the term $g'(\boldsymbol{S})$, and is only associated with even order local polynomial estimation, i.e. local constant, quadratic estimation etc. When the density of observations becomes small, the expression becomes large and introduces significant bias in the estimate. Odd order local polynomial regressions automatically induce a boundary bias correction.

Curvature-based bias, represented by the term in square brackets in equation (1.40), is driven by a difference between the local order of polynomial regression and the local curvature of the function. When the order of polynomial regression is lower than the local order of the function it is impossible to plot the polynomial accurately through the function. In particular, the estimated function will differ more markedly at points where the curvature of the function is high. As an illustration, consider plotting a straight line through a quadratic function.

5.2 Sufficient Conditions for Parameter Identification with LPKD estimation

The identification conditions (1.34), (1.35) and (1.36) show us that increasing the bandwidth or order of LPR in order to reduce bias in the conditional spatial expectations affects identification in conflicting ways. Less bias in the spatial expectation of the confounding spatial data process, Z, reduces the bias in the parameter estimates. On the other hand, less bias in the spatial expectation of $\boldsymbol{X}$, reduces the variation left over in the smoothed observed data, increasing the variance of the estimates of the parameter estimates, $\boldsymbol{\beta}$.

While bias can be reduced by either increasing the order of local polynomial regression or reducing the bandwidth, for simplicity of exposition I focus just on the case where the analyst selects a fixed bandwidth, $\bar{h}$,[5] but adjusts ρ .[6] Accordingly, let us define three orders of local polynomial regression $\rho^{*}, \rho^{**}$and ρ^{***} as the lowest orders that satisfy the identification conditions relating to IC1, IC2 and IC3 in section 4.4. As such,

IC SSE 1: There exists some ρ^{*} whereby the LPR estimate of the conditional spatial expectation of $\boldsymbol{X}$ is unbiased,

$$\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S}, \bar{h}, \rho^{*}) = 0 \tag{1.41}$$

In this case all of the spatial variation in $\boldsymbol{X}$ is removed.

IC SSE 2: There exists some ρ^{**} whereby the LPR estimate of the conditional spatial expectation of Z is unbiased,

$$\boldsymbol{f}(Z|\boldsymbol{S}, \bar{h}, \rho^{**}) = 0 \tag{1.42}$$

If $\rho^{**} < \rho^{*}$ it is possible to remove all of the spatial variation in Z whilst leaving

---

[5] This can be chosen using an automatic bandwidth selection criterion.

[6] A smaller bandwidth can be adopted in local linear estimation to take into account only closer neighbours where a linear fit is more accurate and in local quadratic estimation a larger bandwidth will be adopted to utilise the available data. As a result, in absolute terms the difference in bias between local linear and local quadratic estimation will be smaller when an automatic bandwidth selection criteria is used as opposed to a fixed bandwidth.

some spatial variation in $\boldsymbol{X}$. Remaining spatial variation in $\boldsymbol{X}$ aids in the identification of parameters and provides more efficient parameter estimates than is obtained using ρ^{*}.

IC SSE 3: There exists some ρ^{***} whereby the LPR estimate of the conditional spatial expectations of $\boldsymbol{X}$ and Z are biased but the biases are uncorrelated,

$$\begin{gathered} E[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\bar{h},\rho^{***})'\boldsymbol{f}(Z|\boldsymbol{S},\bar{h},\rho^{***})] = 0 \\ \boldsymbol{f}\big(\boldsymbol{X}\big|\boldsymbol{S},\bar{h},\rho^{***}\big) \neq 0 \\ \boldsymbol{f}\big(Z\big|\boldsymbol{S},\bar{h},\rho^{***}\big) \neq 0 \end{gathered} \tag{1.43}$$

If $\rho^{***} \leq \rho^{*}, \rho^{**}$ then it is possible to remove the source of bias in the parameter estimates whilst further increasing the efficiency of the estimator by leaving additional spatial variation in $\boldsymbol{X}$.

These results translate into conditions regarding the nature of $\boldsymbol{X}$ and Z. If it is possible to represent $\boldsymbol{X}$ and Z by their polynomial expansions then parameter identification is possible using a LPR SSE of order ρ^{*} when either one of two sufficient conditions holds,

CONDITION 1 (C1): Z is a ρ^{*} times differentiable function AND $\boldsymbol{X}$ are NOT purely spatial variables, such that $\boldsymbol{\eta} \neq 0$.

CONDITION 2 (C2): $\boldsymbol{X}$ is $\rho^{*}+2$ times differentiable[7] AND the bias in the conditional spatial expectations of $\boldsymbol{X}$ and Z are uncorrelated.

IC SSE 1 and IC SSE 2 map into C1, whilst IC SSE 3 is the equivalent of C2.

---

[7] Here the (+2) is required so that the ρ^{*}th derivative of X is not constant and variation remains in the smoothed data for identification. In the multivariate case it is necessary that $n+2$ is the lowest order of $X \in \boldsymbol{X}$.

In the context of hedonic house price estimation, Condition 1 will be met by observed property characteristics that are not purely spatial. Using the same example, road noise is exacerbated by factors such as double-glazing, the aspect of a property and the presence of trees and other buildings that cause echoing. Contrastingly, aircraft noise is determined purely by the location of a property with respect to the flight path.

Condition 2 implies that observed property characteristics are more variable over space than the components of the spatial data process with which they are correlated. As an illustration, consider the estimation of a model where house prices are a function of transport noise and some omitted amenity effect. The omitted amenity variable is likely to vary from neighbourhood to neighbourhood. Now consider some different possible sources of transport noise: i) aircraft noise exhibits little variation (under the flight path), ii) railway noise varies from neighbourhood to neighbourhood and iii) road noise is highly variable, differing from street to street. Condition 1 is likely to be met in the case of road noise but not for aircraft or railway noise in this example.

Unfortunately ρ^* is unknown and, at present, little guidance on selecting the order of local polynomial estimation is available from the literature on semi-parametric estimation. This paper adopts both local linear and local quadratic estimation in its simulations. These are fairly low orders that challenge the spatial smoothing estimator. Considering both orders enables us to investigate the importance of the chosen order of polynomial regression and the relevance of the identification conditions derived above.

6.0 Monte Carlo Simulations

The performances of OLS, SEMs and SSEs, in dealing with various spatial data-generating processes are investigated through a series of Monte Carlo experiments. In those simulations, the true data-generating process is given by the linear model,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + Z(\boldsymbol{S},\nu) + \epsilon \tag{1.5}$$

where $\boldsymbol{S} = \{S_1, S_2\}$ are two-dimensional locations drawn from a plane in sample sizes of 500. The innovations, ϵ, are generated as normally distributed random variable. $\boldsymbol{X} = \{X_1, X_2\}$ is a matrix of two regressors where X_1 is a non-spatial variable simulated as draws from a standard normal,

$$X_1 = \eta_1 \sim N(0,1) \tag{1.44}$$

While X_2 is a spatial variable simulated as the sum of a spatial polynomial of order n and some non-spatial element, η_2,

$$X_2 = \sum_{k=0}^{n} \sum_{j=0}^{n} a_{k,j} S_1^k S_2^j + \eta_2 \tag{1.45}$$

Here the $a_{k,j}$ are constants and the exact value of n and the definition of η_2 differs across simulations.

As I will demonstrate through simulations, the relative scale at which the observed data and confounding spatial data generating process vary (often referred to as their spatial scales) is an important determinant of the magnitude of bias when spatial models are estimated (Paciorek, 2010). As such I consider two different orders of X_2 (order 5 and order 1). Each simulation was completed both for $\eta_2 = 0$ and $\eta_2 \sim N(0,1)$.

$\boldsymbol{\beta}$ is a vector of parameters of interest relating to the observed characteristics, $\boldsymbol{X}$. In each of the Monte Carlo simulations these are defined as,

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.46}$$

The Monte Carlo experiment proceeds by simulating data imitating each of the four different processes (discussed in Section 2) by which Z, the confounding spatial process, might arise in data.

6.1 Spatial Data Generating Processes (DGPs)

Table 1.3 summarises the data generating processes used in the Monte Carlo simulations and their corresponding parameter values.

DGP 1. A Spatial Autoregressive Lag (SAR)

$$Z(\boldsymbol{S},\nu) = \Sigma_{j=1}^{\infty}\lambda^{j}\boldsymbol{W}^{j}\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + \Sigma_{j=1}^{\infty}\lambda^{j}\boldsymbol{W}^{j}\epsilon \tag{1.47}$$

DGP 2. A Spatial Autoregressive Lag with Autoregressive Spatial Error Dependence (SARAR)

$$\begin{aligned} Z(\boldsymbol{S},\nu) = {} & \Sigma_{j=1}^{\infty}\lambda^{j}\boldsymbol{W}^{j}\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + (\Sigma_{\mathrm{j=0}}^{\infty}\Sigma_{l=1}^{\infty}\lambda^{j}\boldsymbol{W}^{j}\varphi^{l}\boldsymbol{M}^{l} \\ & + \Sigma_{l=1}^{\infty}\varphi^{l}\boldsymbol{M}^{l})\epsilon \end{aligned} \tag{1.48}$$

In both cases $\lambda = 0.6$ and W was generated as an Epanechnikov spatial weights matrix with limits of 3. In the SARAR model $\varphi = -0.6$ and M was generated as an Epanechnikov spatial weights matrix with limits of 0.5.

DGP 3. To examine spatial misspecification I consider the estimation of a spatial lag process. The estimation model is,

$$Y = \boldsymbol{X\beta} + \lambda\widehat{\boldsymbol{W}}Y + \epsilon \tag{1.49}$$

When the true data generating process is,

$$Y = \boldsymbol{X\beta} + \lambda\boldsymbol{W}Y + \epsilon \tag{1.50}$$

As was discussed in section 2.2 (see Table 1.1) the confounding spatial DGP, Z, in this case is,

$$Z = \lambda(\boldsymbol{W} - \widehat{\boldsymbol{W}})Y \quad (1.51)$$

Two versions of this model were generated for the Monte Carlo simulations. First $\boldsymbol{W}$ and $\widehat{\boldsymbol{W}}$ were specified with broad limits of 3, such that the confounding spatial data varies less than the observed variables over a localised area and, thus, Condition 2 is likely to be met. Second, $\boldsymbol{W}$ and $\widehat{\boldsymbol{W}}$ were specified with narrow limits of 0.4 such that the confounding spatial data process is more variable over local areas and Condition 2 is less likely to be met. In both cases, $\lambda = -0.6$ and $\widehat{\boldsymbol{W}}$ is a misspecified version of $\boldsymbol{W}$ in that the former specifies binary weights when the true weights in $\boldsymbol{W}$ are Epanechnikov. The parameter values are presented in Table 1.3.

DGP 4. In the spatial measurement error model, the measured values of X_2 were generated as,

$$X_2^* = X_2 + Z(\boldsymbol{S}) \quad (1.52)$$

Where Z is generated as a spatial polynomial of order 1,

$$Z = b_{1,0}S_1 + b_{0,1}S_2 + b_{1,1}S_1S_2 \quad (1.53)$$

Recall that two versions of X_2 are considered in the Monte Carlo simulations. When X_2 is a polynomial of order 5, Condition 2 is met and it is possible to smooth out the measurement error, Z, and identify the parameters of the model even when X_2 is purely spatial (Condition 1 is violated). In contrast, when X_2 is a polynomial of order 1, Condition 2 is not met and spatial smoothing may not be able to identify unbiased parameter estimates.

DGP 5. For the spatial omitted variable, a spatial polynomial of order 2 was generated,

$$Z = \sum_{k=0}^{2} \sum_{j=0}^{2} b_{k,j} S_1^k S_2^j \tag{1.54}$$

Where $b_{k,j}$ are constant parameters.

Table 1.3 summarises the parameter values used in the simulations for each DGP.

DGP		$X(\boldsymbol{S}, \eta)$	$Z(\boldsymbol{S}, \nu)$
Spatial Parametric Models:	SAR	Polynomial of order 5	$\lambda = 0.6, W = 3$ (Epanechnikov)
	SARAR	Polynomial of order 5	$\lambda = 0.6, W = 3$ (Epanechnikov) $\varphi = 0.6, M = 0.5$ (Epanechnikov)
Spatial Misspecification:	Condition 2 violated	Polynomial of order 1 +SAR: $\lambda = 0.6, \widehat{W} = 0.4$ (Binary)	$\lambda = 0.6, W = 0.4$ (Epanechnikov)
	Condition 2 met	Polynomial of order 1 +SAR: $\lambda = 0.6, \widehat{W} = 3$ (Binary)	$\lambda = 0.6, W = 3$ (Epanechnikov)
Spatial Measurement Error:	Condition 2 violated	Polynomial of order 1	Polynomial of order 1
	Condition 2 met	Polynomial of order 5	Polynomial of order 1
Spatial Omitted Variables:	Condition 2 violated	Polynomial of order 1	Polynomial of order 2
	Condition 2 met	Polynomial of order 5	Polynomial of order 2

Table 1.3: Parameter Values in the Monte Carlo Simulations (N=500)

7.0 Results

7.1 OLS Estimation

Table 1.4 reports the results using OLS estimation when the estimating equation fails to include the confounding spatial data generating process. The left column presents results when X_2 is a purely spatial regressor ($\eta = 0$) and the right presents results when there is some non-spatial variation in X_2, ($\eta \neq 0$). Note that if X_2 is a purely spatial regressor ($\eta = 0$) it is anticipated that the parameters of the model cannot be identified under identification condition 1.

Model	Condition 1 violated $\eta = 0$		Condition 1 met $\eta \neq 0$	
	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)
Spatial Parametric Models				
SAR	1.2256 (0.4395)	-1.4011 (0.0105)	1.2355 (0.4409)	-1.4009 (0.0105)
SARAR	1.2256 (0.4395)	-1.4011 (0.0105)	1.2356 (0.4410)	-1.4009 (0.0105)
Spatial Measurement Error Models				
Condition 2 violated (Order(**X**)<Order(Z))	0.9990 (0.0223)	-0.0327 (1.0421)	0.9809 (0.0497)	-0.0414 (0.0031)
Condition 2 met (Order(**X**)>Order(Z))	0.9927 (0.1040)	-0.9173 (0.0023)	0.9909 (0.1041)	-0.9175 (0.0023)
Spatial Omitted Variables				
Condition 2 violated (Order(**X**)<Order(Z))	1.0279 (0.1305)	30.8516 (0.2665)	1.4068 (0.6158)	6.1906 (0.5574)
Condition 2 met (Order(**X**)>Order(Z))	0.9399 (0.2950)	-0.6562 (0.0070)	0.9325 (0.2957)	-0.6572 (0.0070)

Table 1.4: OLS Estimation

The OLS estimator neglects to account for the spatial endogeneity (SAR) and auto-correlated autoregressive error (AR) components of the spatial interaction DGPs. As is shown in rows 1 and 2, for both the SAR and SARAR processes there is significant bias in the OLS parameter estimates of β_2 both when X_2 is purely spatial (Condition 1 is violated) and when it is not. $\beta_1$is also biased with very large standard errors. The marginal effect of X_2, β_2, is overstated by roughly 40 percent in each case.

Again as anticipated by (1.19), spatial measurement error leads to bias in the OLS parameter estimate of β_2. As with the omitted spatial covariates model, the magnitude of the bias differs when Condition 2 does and does not hold. When Condition 2 holds, that is when X_2 is a polynomial of order 5 and Z is a polynomial of order 1, the bias in the parameter estimate is small. In contrast, when Condition 2 does not hold (that is X_2 and Z are both polynomials of order 1) the OLS parameter estimates suggest that X_2 has a negligible marginal effect.

When the data is generated with an omitted spatial variable, taking the form of a polynomial over space, in the case where Condition 2 does not hold the bias in β_2 is severe. However, the bias is reduced when there is some non-spatial variation in X_2 (column 4). When Condition 2 holds β_2 is biased (by 35%) and there is also a small bias in β_1.

7.2 Spatial Econometric Estimation

Table 1.5 reports the results of SAR and SARAR estimation for each of the DGPs. Columns 1-2 and 3-4 relate to SAR and SARAR estimation respectively.[8]

---

[8] Confounding spatial processes often leave a tell-tale indicator in the form of spatially dependent residuals. This enables their presence to be detected through post-estimation tests. A number of tests have been developed; Moran's I provides a test for general spatial dependence (Cliff and Ord 1972), whilst LM-tests can be used to test the null hypothesis of no spatial auto-correlation against the alternative hypotheses of specific parametric models of dependence (Anselin 1988). Likewise, Kelejian and Robinson (1992) develop specification robust tests that allow an analyst to test, for example, for the presence of a spatial lag when spatial error dependence may also be present and vice versa.

Consider first the two spatial interaction models. Rows 1 and 2 present the results when the spatial weights matrices are correctly specified. In both cases the endogenous component of the model is correctly accounted for. This removes the source of bias and allows both models to recover unbiased estimates. However, whilst the SARAR estimator is sufficient for estimating the SAR model, the SAR estimator lacks the auto-regressive error component of the SARAR model. As a result, the standard errors obtained by SAR estimation of the SARAR model are biased and inflated.

It is well known that spatial parametric estimators perform most effectively when they are correctly specified. In contrast, rows 3 and 4 of Table 1.5 present the results obtained when the spatial weights matrices are incorrectly specified as binary when the true spatial lag is generated by a matrix with Epanechnikov weights. Row 3 presents the results when the spatial weights matrices have narrow limits such that Condition 2 is not met. The simple misspecification of the weightings as binary causes the SAR estimator to return biased parameter estimates. Row 4 presents results for a further SAR model when the lag process is broad. Again the weights matrix is misspecified as binary rather than Epanechnikov. As before, the parameter estimates of β_2 obtained using a misspecified SAR estimator are significantly biased in all cases. The simulation results in rows 3 and 4 confirm that SEMs are sensitive to specification errors.

These general spatial data problems are indicative of the sort of real world problems that plague econometric analysis, such as hedonic analysis. Our simulation results

---

For completeness I employed these robust–LM tests against each DGP. The results uniformly demonstrate the ability of the tests to detect the presence of a spatial data generating process, however in almost all cases there is not enough information to guide a choice between specifications, even when the true DGP is a SPM and the spatial weights matrices are correctly specified in the tests. This is consistent with the findings of Anselin and Griffith (1988).

clearly illustrate the continuing challenges that they pose for spatial econometric estimators (Brady & Irwin 2010, McMillen 2010, Pinske & Slade 2010).

Model	SAR				SARAR			
	Condition 1 violated $\eta = 0$		Condition 1 met $\eta \neq 0$		Condition 1 violated $\eta = 0$		Condition 1 met $\eta \neq 0$	
	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)	$\hat{\beta}_1$ ($\beta_1 = 1$)	$\hat{\beta}_2$ ($\beta_2 = -1$)
Spatial Parametric Models								
SAR	0.9999 (0.0219)	-0.9999 (0.0010)	0.9997 (0.0219)	-1.0001 (0.0010)	0.9999 (0.0219)	-1.0000 (0.0010)	0.9998 (0.0219)	1.0001 (0.0010)
SARAR	0.9997 (0.0229)	-1.0000 (0.0011)	0.9998 (0.0229)	-1.0001 (0.0011)	0.9999 (0.0222)	-1.0000 (0.0007)	0.9998 (0.0222)	-1.0001 (0.0007)
Spatial Misspecification Models								
Condition 2 violated (i.e. narrow lag process in Z)	1.0416 (0.0097)	-0.8230 (0.0193)	1.0402 (0.0098)	-0.8378 (0.0193)	-	-	-	-
Condition 2 met (i.e. broad lag process in Z)	0.9863 (0.0185)	-0.9314 (0.0422)	0.9861 (0.0186)	-0.9341 (0.0420)	-	-	-	-
Spatial Measurement Error Models								
Condition 2 violated (Order(X)<Order(Z))	1.0035 (0.0232)	-0.0253 (0.0073)	0.9279 (0.1001)	-0.1882 (0.0207)	0.9992 (0.0239)	-0.0256 (0.0072)	0.5863 (0.0552)	-0.1974 (0.0168)
Condition 2 met (Order(X)>Order(Z))	0.9999 (0.1165)	-1.0763 (0.0318)	1.1417 (0.0885)	-1.0584 (0.0050)	1.0187 (0.0243)	-0.9727 (0.0072)	1.0173 (0.0243)	-0.9756 (0.0070)
Spatial Omitted Variables								
Condition 2 violated (Order(X)<Order(Z))	1.0317 (0.1045)	24.1248 (5.2237)	1.0528 (0.0441)	-1.1007 (0.0477)	0.9884 (0.0306)	30.0920 (1.8795)	1.1424 (0.0403)	-1.1942 (0.0410)
Condition 2 met (Order(X)>Order(Z))	0.9824 (0.2063)	-0.3579 (0.0401)	1.0025 (0.1142)	-1.0720 (0.0309)	1.0330 (0.0292)	-0.8351 (0.0071)	1.0173 (0.0243)	-0.9756 (0.0070)

Table 1.5: SPE Estimation

The results for the spatial measurement error model in rows 5 and 6 highlight how parametric spatial weights matrices fail to correct the bias when X_2 does not have any non-spatial variation and the measurement error is highly localised (Condition 2 is not met). This results from the poor approximation of the measurement error made by the spatial weights matrices.

Rows 7 and 8 present parameter estimates from SAR and SARAR estimating models when the data are generated by an omitted spatial process that takes the form of a polynomial over space and not a spatial lag process. The most striking result is the degree of bias that remains in the parameter estimates when X_2 is a purely spatial regressor. In these cases the spatial weights matrices poorly approximate the structure of the omitted spatial covariates and do little to remove bias. In addition, when the unaccounted-for spatial process, Z, exhibits highly localized spatial dependence (i.e. Condition 2 is not met) an overly broad spatial weights matrix provides only a poor approximation to the true spatial data-generating process and parameter estimates continue to be biased even when there is non-spatial variation in X_2.

7.3 Spatial Smoothing Estimation

Thus far I have demonstrated the poor performance of OLS when the econometric model does not account for a variety of confounding spatial data generating processes. Likewise, the results in section 7.2 demonstrate that SEMs also fail to identify unbiased parameter estimates when the model is misspecified. The final set of results, presented in this section, explore the performance of local linear and local quadratic SSEs against these DGPs. In these simulations the bandwidth is selected using leave one out cross validation (Hardle and Marron, 1985).

Table 1.6 presents the parameter estimates obtained using local linear (columns 1 and 2) and local quadratic (columns 3 and 4) SSEs for each of the models. For the spatial misspecification model I combine spatial smoothing and a SAR component to estimate a model with spatial interactions where the structure of the interaction is anticipated but not known with certainty. In each case I pay particular attention to whether either of the sufficient conditions for identification are met.

In the spatial interaction models, the results of which are presented in rows 1 and 2, a local linear SSE has clear advantages over OLS even when there is no independent variation X_2 but does not return completely unbiased estimates. This bias is dispelled when there is some independent variation in X_2 (such that Condition 1 is satisfied). The local quadratic SSE is able to capture and extract the spatial interaction processes more accurately, and presents unbiased parameter estimates.

Rows 3 and 4 present results for the spatial misspecification model. In this case the estimation first smoothes the data and then estimates the SAR model using the smoothed data. When Condition 2 holds, it is possible to smooth out Z and leave some variation in $\boldsymbol{X}$, allowing the SSEs to return unbiased parameter estimates both with and without additional non-spatial variation (Condition 1). The addition of non-spatial variation reduces the standard errors. When neither of the identification conditions is met, both spatial smoothing estimates are biased. When only Condition 1 holds the order of local polynomial estimation is lower than the order of the spatial data problem. As a result, the bias in the spatial smoothing estimates is reduced but not completely removed[9]. Again, this is in line with the expectations that can be drawn from theory.

Rows 5 and 6 of Table 1.6 present the spatial smoothing estimates for the spatial measurement error model and rows 7 and 8 present those for the omitted spatial covariates models. In line with the theoretical results, when neither Condition 1 or 2 hold these estimators cannot correctly identify the parameters of the model. It is noticeable that when the identification conditions are not met, the parameter estimate for $\beta_2$is large, with large standard errors. The reason for this is that under these conditions a SSE removes almost all of the variation in X_2.[10]

---

[9] Further simulations revealed that the bias is removed using a local cubic smooth.

[10] Although in these cases the parameters of the model are not identified, the large estimated values and standard errors signal that the parameters are unidentifiable.

In contrast, when Condition 2 holds the SSEs perform well. Although some bias remains in the parameter estimates obtained using a local linear smooth, this is expected since the omitted data is of order 2 and consequently there is bias in the local linear regression estimate of Z. The bias is correlated with the smoothed observed data and results in a small bias in the parameter estimate. Consistent with the derived identification conditions, the local quadratic estimator removes this problem and obtains unbiased parameter estimates. Columns 2 and 4 confirm that non-spatial variation in X (condition 1) facilitates the identification of the parameters enabling both SSEs to return unbiased parameter estimates.

8.0 Discussion

As more and better spatially organised micro-datasets become available to economists, the problem of confounding spatial processes is likely to become of more prominent concern to empirical analysts. This paper has explored issues of parameter identification and estimation of linear models in the presence of a broad range of underlying spatial DGPs. Four types of problems were considered: spatial lag processes, omitted spatial covariates, spatial measurement error and spatial misspecification. These problems are shown to undermine OLS estimation, potentially introducing bias into the parameter estimates

Whilst some progress has been made to address the issues associated with spatial data through the development and implementation of spatial econometric estimators there are some serious limitations to these approaches. In this paper I have considered and contrasted the ability of commonly employed spatial econometric estimators (SAR and SARAR) and alternative local linear and local quadratic SSEs in handling a variety of confounding spatial data generating processes.

	Local Linear SSE				**Local Quadratic SSE**			
Model	Condition 1 violated $\eta = 0$		Condition 1 met $\eta \neq 0$		Condition 1 violated $\eta = 0$		Condition 1 met $\eta \neq 0$	
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
Spatial Parametric Models								
SAR	0.9971 (0.0226)	-1.2993 (0.0186)	0.9976 (0.0225)	-1.0160 (0.0147)	0.9984 (0.0226)	-1.0172 (0.0146)	0.9944 (0.0222)	-0.9951 (0.0032)
SARAR	0.9975 (0.0247)	-1.0300 (0.0168)	0.9981 (0.0247)	-0.9994 (0.0142)	0.9985 (0.0231)	-1.0175 (0.0148)	0.9943 (0.0239)	-0.9969 (0.0233)
Spatial Misspecification Models								
Condition 2 violated	1.1274 (0.0081)	-0.8292 (0.0941)	1.1310 (0.0080)	-0.9166 (0.0951)	1.1353 (0.0077)	-5.0898 (2.7766)	1.1476 (0.0073)	-1.4138 (0.3586)
Condition 2 met	1.0032 (0.0023)	-0.9565 (0.0150)	1.0031 (0.0023)	-0.9589 (0.0147)	1.0036 (0.0022)	-0.9794 (0.0745)	1.0036 (0.0022)	-0.9906 (0.0539)
Spatial Measurement Error Models								
Condition 2 violated	0.9992 (0.0219)	-0.0009 (0.0224)	0.9986 (0.0224)	-0.9984 (0.0233)	0.9994 (0.0219)	-1.32x10^{10} (1.39x10^{11})	0.9993 (0.0220)	-0.9998 (0.0230)
Condition 2 met	0.9993 (0.0220)	-1.0001 (0.0014)	0.9992 (0.0220)	-0.9999 (0.0045)	1.0004 (0.0219)	-1.0028 (0.0461)	1.0004 (0.0219)	-1.0001 (0.0201)
Spatial Omitted Variables								
Condition 2 violated	1.1274 (0.0081)	-0.8292 (0.0941)	1.1310 (0.0080)	-0.9166 (0.0951)	1.1353 (0.0077)	-5.0898 (2.7766)	1.1476 (0.0073)	-1.4138 (0.3586)
Condition 2 met	0.9972 (0.0223)	-0.9240 (0.0122)	0.9957 (0.0224)	-0.9412 (0.0112)	0.9997 (0.0219)	-1.0000 (0.0471)	1.0001 (0.0219)	-0.9999 (0.0230)

Table 1.6: SSE Estimation

The primary contribution of this paper is to define three identification conditions which establish when parameter identification is possible using a general nonparametric SSE. Furthermore, I show the specific form of these identification conditions for a SSE using local polynomial regression and derive from them two sufficient identification conditions in terms of the structure of the data.

A second contribution is to present a series of Monte Carlo simulations, which confirm that those Conditions are sufficient for identification of the parameters. Under either Condition, a SSE provides a simple and robust approach to dealing with spatial data generating processes. The results also highlight the importance of selecting an appropriate order of local polynomial estimation.

In the Monte Carlo experiment, SEMs are shown to be more efficient than SSEs when the SEM correctly anticipates the form of spatial lag process and correctly specifies the spatial weights matrices. However, SEMs are shown to not be robust to misspecification and perform poorly when the spatial data generating process is not a spatial lag.

Accordingly, when there is little guidance on the nature of unaccounted-for spatial processes, the SSEs offers an attractive alternative to SEMs even when it is suspected that that spatial process may be in the form of a spatial lag. The SSE does not impose restrictive assumptions on the form of the spatial data-generating process and has been shown to produce unbiased parameter estimates when either of two identification conditions is met. Moreover, by adopting a spatially smoothed SAR estimator in the misspecified model, I demonstrate that spatial smoothing can be combined with other estimation methods, such as two-stage least squares, in the estimation of spatial models[11]. Again, when Conditions

[11] Although this is not the primary objective of this chapter, additional simulations were conducted to examine the ability of spatially smoothed SAR estimators to recover unbiased estimates of the spatial multiplier when the true underlying DGP is a SAR with an omitted spatial variable. The results confirm that when the spatial weights matrix is correctly specified the spatial smoothing estimator returns unbiased parameter estimates, including an unbiased estimate of the spatial multiplier. In contrast, a SAR estimator

1 or 2 are met, the spatial smoothing estimator can be adopted to overcome spatial misspecification problems, combining the advantages of the two approaches.

In conclusion, this paper has sought to demonstrate the flexibility and simplicity of SSEs and to provide a clear and intuitive understanding of the conditions under which this estimator can be used to identify the parameters of the model. This paper has demonstrated how it is possible to facilitate identification and further improve the efficiency properties of estimators through better matching the demands of economic analysis with developments in non-parametric econometric techniques. It is my hope that this work, in conjunction with the increasing availability of non-parametric code in statistical packages, will support the increased utilisation of SSEs in economics. Future work should extend this principle and further explore the potential for econometric theory to contribute to achieving the objectives of applied analysis. In particular, in the next chapter I will explore the potential for further identification and efficiency gains through the implementation of local adaptive spatial smoothing estimators.

returns biased parameter estimates even when the spatial weights matrix is correctly specified. Results are available from the author upon request.

CHAPTER 2

A Locally Adaptive Spatial Smoothing Estimator

1.0 Introduction

In the previous chapter, this thesis focused its attention upon estimating partial linear models using a spatial smoothing estimator (SSE) with fixed smoothing parameters, ϕ, hereafter referred to as a Fixed SSE. In this chapter, I focus on a refinement of that estimator which improves its efficiency and potentially expands the range of spatial data-generating processes from which it can successfully identify unbiased parameter estimates.

As per the previous chapter, the general form for the data-generating process is given by equation (1.5), which is repeated below,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + Z(\boldsymbol{S},\nu) + \epsilon \tag{1.8}$$

Recall that the dependent variable, Y, is an additively separable function of an $N \times k$ vector of observed data, $\boldsymbol{X}(\boldsymbol{S},\eta)$, a confounding spatial process, $Z(\boldsymbol{S},\nu)$, and independent, identically distributed innovations, ϵ. Also, $\boldsymbol{S}$, represents spatial location as defined by a vector of Cartesian co-ordinates, $\boldsymbol{S} = \{S_1, S_2\}$. Finally, $\boldsymbol{\beta}$ is a k × 1 vector of parameters relating to the observed data. The objective of the analysis is to recover unbiased estimates of these parameters.

In the previous chapter I demonstrated that the parameters of the model can be identified when equation (1.33) is satisfied, that is,

$$E[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'\boldsymbol{f}(Z|\boldsymbol{S},\phi)] = 0 \tag{1.33}$$

Where $f(.|\boldsymbol{S},\phi)$ is the bias component associated with the empirical estimate of the conditional spatial expectation. From this expression I derived three identification conditions relating to the estimation of models within this framework using a Fixed SSE. These conditions are as follows,

IC1. There exists some ϕ whereby empirical estimate of the conditional spatial expectation of $\boldsymbol{X}$ is unbiased,

$$\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi) = 0 \tag{1.34}$$

In this case spatial smoothing works so as to remove all spatial variation in $\boldsymbol{X}$, a process that addresses the omitted variable bias problem but limiting identification of the parameters to variation provided by η, the non-spatial elements of $\boldsymbol{X}$.

IC2. There exists some ϕ whereby the empirical estimate of the conditional spatial expectation of Z is unbiased,

$$\boldsymbol{f}(Z|\boldsymbol{S},\phi) = 0 \tag{1.35}$$

In this case spatial smoothing works so as to remove the spatial component of the confounding spatial process without removing all of the spatial variation in X.

IC3. There exists some ϕ whereby the empirical estimate of the conditional spatial expectations of $\boldsymbol{X}$ and Z are biased but the biases are uncorrelated,

$$\begin{aligned} E[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi)'\boldsymbol{f}(Z|\boldsymbol{S},\phi)] &= 0 \\ \boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\phi) &\neq 0 \\ \boldsymbol{f}(Z|\boldsymbol{S},\phi) &\neq 0 \end{aligned} \tag{1.36}$$

In this case, smoothing works so as to remove just the correlated spatial components of $\boldsymbol{X}$ and Z.

2.0 Spatially Inhomogeneous Data

When the functions that underlie the model are spatially homogeneous (i.e. their variability is similar across different regions) and one of the identification conditions is met, the Fixed SSE provides unbiased estimates of the parameters of interest. In this chapter I consider situations in which the observed data, $\boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})$ and confounding spatial process, $Z(\boldsymbol{S},\nu)$ are spatially inhomogeneous.

First it is prudent to define what is meant by spatially inhomogeneous. A simple definition is to say that a function is spatially inhomogeneous if it is highly variable over some regions of space and less variable over others. More formally, consider a function that can be expressed as a function of space using polynomial approximation. This function is spatially homogeneous when the order of the approximated polynomial is the same when approximated using any local subset of the function. An example of a spatially homogeneous function is provided in Figure 2.1. Note that a polynomial approximation would have the same order whether evaluated over region 1, region 2 or both regions simultaneously.

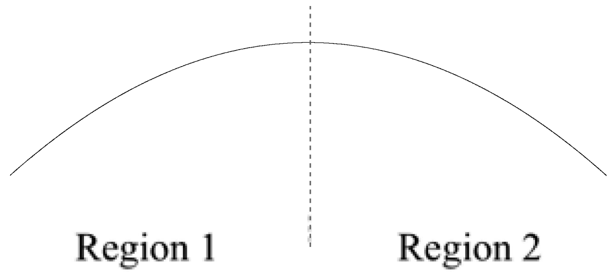

Figure 2.1: An example of a spatially homogeneous function

In contrast, a spatially inhomogeneous function exhibits different orders across different local regions. An example of a spatially inhomogeneous function is presented in Figure 2.2. Note that this function can be approximated by a high order polynomial over region 1 and a low order polynomial over region 2.

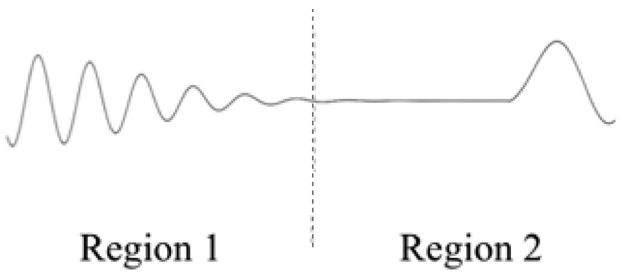

Figure 2.2: An example of a spatially inhomogeneous function

I demonstrate that in these cases a Fixed SSE is inefficient and may produce misleading results. I show that the implementation of a Local Adaptive SSE, which applies locally varying smoothing parameters, can increase efficiency and permit identification in cases where a Fixed SSE cannot.

3.0 Parameter Identification with Locally Adaptive Spatial Smoothing Estimation

In essence the innovation of the Locally Adaptive Spatial Smoothing Estimator (LASSE) is to allow the degree of smoothing to vary from locality to locality across space. The degree of smoothing at any particular location is selected so as to generate efficiency in the estimation of the parameters. Taking Figure 2, for example, a locally adaptive smooth might adopt one set of smoothing parameters over region 1 and another over region 2.

3.1 Improving efficiency through LASSE

Building on the presentation in the previous chapter, let us now consider how allowing for locally adaptive smoothing might impact on the three identification conditions previously derived for SSE. Again, the focus will be on smoothing implemented using local polynomial regression (LPR). In particular, I examine the case in which the analyst adopts some fixed bandwidth, $\bar{h}$, but allows the order of LPR to vary across locations[12]. I denote the N-vector of location-specific orders of polynomial regression by $\boldsymbol{\rho}_{\boldsymbol{local}}$. Of course, it is possible to set each local order of polynomial regression to the fixed order that is known to achieve each of the identification conditions. As such, this optimal fixed order defines the highest order that will be used at any location such that $\boldsymbol{\rho}^{*}_{\boldsymbol{local}}(n) \leq \rho^{*}$ for any n.

[12] An alternative would be to hold the order of polynomial regression constant while varying the bandwidth locally.

In the LASSE, the identification conditions becomes,

IC LASSE 1: There exists some $\boldsymbol{\rho}^{*}_{\boldsymbol{local}}$ whereby LPR estimate of the conditional spatial expectation of $\boldsymbol{X}$ is unbiased,

$$\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S}, \bar{h}, \boldsymbol{\rho}^{*}_{\boldsymbol{local}}) = 0 \tag{2.1}$$

As with a Fixed SSE, all of the spatial variation in $\boldsymbol{X}$ is removed even if $\boldsymbol{\rho}^{*}_{\boldsymbol{local}}(n) < \rho^{*}$ for some n. As a result, moving to a LASSE does not provide any efficiency gains over the Fixed SSE.

IC LASSE 2: There exists some $\boldsymbol{\rho}^{**}_{\boldsymbol{local}}$ whereby LPR estimate of the spatial expectation of Z is unbiased,

$$\boldsymbol{f}(Z|\boldsymbol{S}, \bar{h}, \boldsymbol{\rho}^{**}_{\boldsymbol{local}}) = 0 \tag{2.2}$$

In this case it is possible to remove all of the spatial variation in Z whilst leaving some spatial variation in $\boldsymbol{X}$. Moreover, recall that ρ^{**} is the lowest fixed order of polynomial regression that satisfies IC SSE 2. If $\boldsymbol{\rho}^{**}_{\boldsymbol{local}}(n) < \rho^{**}$ for some n, then the locally adaptive estimator results in a greater bias in the estimate of the conditional spatial expectation of $\boldsymbol{X}$. Of course, greater bias in that estimate results in more variation in the smoothed $\boldsymbol{X}$ and this variation provides additional data with which to identify the parameters of the model. As a result, under this condition, LASSE parameter estimates are likely to be more efficient than the Fixed SSE.

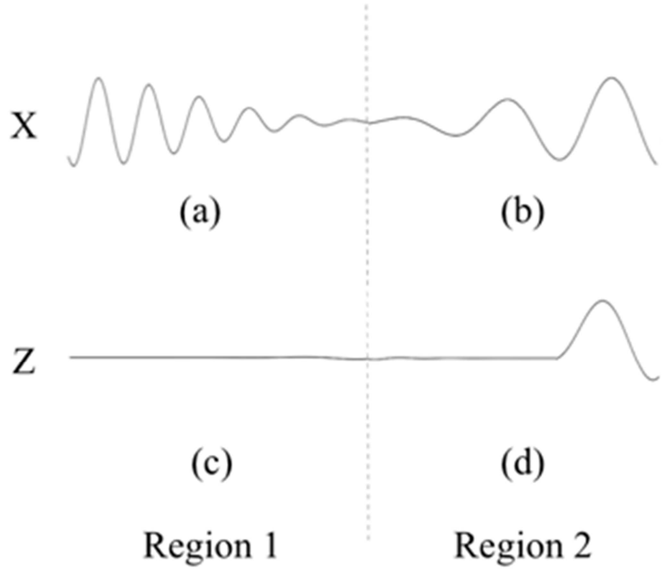

Figure 2.3: IC 2 Spatially inhomogeneous data where $b > d$
(Local order in paretheses)

Figure 2.3 presents an example of spatially inhomogeneous data for which a LASSE would provide efficiency gains over Fixed SSE. For clarity, this example presents a simple two-dimensional representation where the variability of two functions, X and Z, differs across two distinct regions, 1 and 2. The order of X is high across region 1 but lower across region 2 and the opposite is true of function Z. The order of the local polynomial approximations of X and Z are shown in parentheses in the Figure.

If the estimating model assumes the data-generating process to be $Y = X\beta + \epsilon$, and is estimated using the Fixed Order SSE then both X and Z would be smoothed at the same level across both regions 1 and 2. If $a > d$ then the second identification condition is met and β is identifiable using Fixed SSE with order b. However, a more efficient estimate could be obtained by smoothing using a local order of polynomial regression of c over region 1 and d over region 2. This locally adaptive approach would still achieve the objective of removing the omitted variable and eliminating the bias in the parameter estimates but would also leave greater variation in the smoothed observed variable over region 1, which enhances the efficiency of the estimator.

IC LASSE 3: There exists some $\boldsymbol{\rho}_{\boldsymbol{local}}^{***}$ whereby the LPR estimate of the conditional spatial expectations of $\boldsymbol{X}$ and Z are biased but the biases are uncorrelated,

$$\begin{gathered} E[\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S},\bar{h},\boldsymbol{\rho}_{\boldsymbol{local}}^{***})'\boldsymbol{f}(Z|\boldsymbol{S},\bar{h},\boldsymbol{\rho}_{\boldsymbol{local}}^{***})] = 0 \\ \boldsymbol{f}\big(\boldsymbol{X}\big|\boldsymbol{S},\bar{h},\boldsymbol{\rho}_{\boldsymbol{local}}^{***}\big) \neq 0 \\ \boldsymbol{f}\big(Z\big|\boldsymbol{S},\bar{h},\boldsymbol{\rho}_{\boldsymbol{local}}^{***}\big) \neq 0 \end{gathered} \tag{2.3}$$

Recall that ρ^{***} is the lowest fixed order of polynomial regression that satisfies IC SSE 3. If $\boldsymbol{\rho}_{\boldsymbol{local}}^{***}(n) < \rho^{***}$ for some n then it is possible to remove the source of bias in the parameter estimates whilst further increasing the efficiency of the estimator by leaving additional spatial variation in $\boldsymbol{X}$. Consider Figure 2.3, while it is known from identification condition 2 (IC 2) that it would be possible to smooth out Z using orders c and d over regions 1 and 2 respectively, identification condition IC LASSE 3 shows that it may also be possible to remove the correlation between the bias components and increase the efficiency of the estimates by smoothing at orders less than c and d over region 1 and region 2 respectively.

3.2 Relaxing the Identification Conditions through LASSE

If $\boldsymbol{X}$ are purely spatial functions, and by using a fixed bandwidth it is impossible to smooth out the correlation between $\boldsymbol{X}$ and Z without perfectly predicting $\boldsymbol{X}$, then identification using a Fixed SSE is impossible. However, it is possible to identify the parameters of the model using a LASSE if the local order of $\boldsymbol{X}$ is greater than the local order of Z over some localised regions. In this case, a lower order of polynomial regression is adopted in some regions. Across those regions, smoothing no longer perfectly predicts X such that variation is left in the smoothed data from which the $\boldsymbol{\beta}$ parameters can be estimated.

Figure 2.4 presents a second example of spatially inhomogeneous data for which a LASSE would provide efficiency gains over Fixed SSE. If $a < d$ and $\rho^{***} = \rho^{*} = a$, it is impossible to remove the correlation between the smoothed observed data and the smoothed confounding data without removing all of the spatial variation in $\boldsymbol{X}$. In this case, the parameters of the model cannot be identified using a Fixed SSE if $\boldsymbol{X}$ is a purely spatial function. However, identification is possible using LASSE with a local order of polynomial regression of c over region 1 to remove the spatial component of Z, and b over region 2 to remove the spatial component of $\boldsymbol{X}$, leaving the smoothed data uncorrelated. This locally adaptive approach would achieve the objective of eliminating the bias in the parameter estimates, providing additional identification over the Fixed SSE.

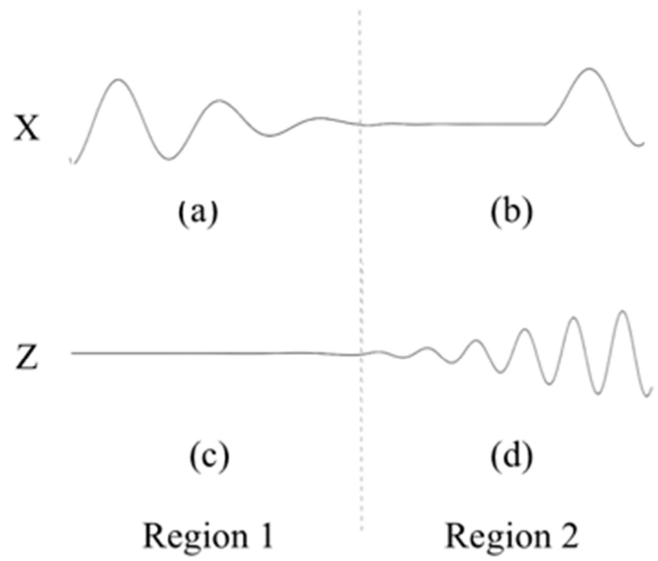

Figure 2.4: Identification under IC 3 with Spatially Inhomogeneous Functions

(Local order in parentheses)

This intuition underpins the adaptive order approach developed by Fan and Gijbels (1995) for non-parametric kernel density estimation of functions. The methodology can be combined with the properties of the compound error, $\hat{e} = Y - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$, to develop an Local Adaptive SSE which not only achieves efficiency savings over the Fixed SSE but also enables identification for a broader range of data.

In chapter 1 (section 5.2) I derived two sufficient conditions for identification using a Fixed SSE,

CONDITION 1 (C1): Z is a ρ^* times differentiable function AND $\boldsymbol{X}$ are NOT purely spatial variables, such that $\boldsymbol{\eta} \neq 0$.

CONDITION 2 (C2): $\boldsymbol{X}$ is $\rho^* + 2$ times differentiable[13] AND the bias in the conditional spatial expectations of $\boldsymbol{X}$ and Z are uncorrelated.[14]

For the LASSE, IC LASSE 3 also translates into a third sufficient condition for identification,

CONDITION 3 (C3): There exists some set of local adaptive smoothing orders, $\boldsymbol{\rho_{local}}$, such that $\boldsymbol{f}(\boldsymbol{X}|\boldsymbol{S}, \bar{h}, \boldsymbol{\rho_{local}})$ and $\boldsymbol{f}(Z|\boldsymbol{S}, \bar{h}, \boldsymbol{\rho_{local}})$ are uncorrelated AND variation remains in $\boldsymbol{f}\left(\boldsymbol{X}\middle|\boldsymbol{S}, \bar{h}, \boldsymbol{\rho_{local}}\right) + \boldsymbol{\eta}$**.**

This condition is automatically met when either CONDITION 1 or CONDITION 2 are met and may hold independently.

4.0 Methodology

Fan and Gijbels (1995) develop a method for non-parametrically estimating a function using an order that adapts locally. Consider the model,

---

[13] Here the (+2) is required so that the ρ^*th derivative of X is not constant and variation remains in the smoothed data for identification. In the multivariate case it is necessary that $n + 2$ is the lowest order of $X \in \boldsymbol{X}$.

[14] A smaller bandwidth can be adopted in local linear estimation to take into account only closer neighbours where a linear fit is more accurate and in local quadratic estimation a larger bandwidth will be adopted to utilise the available data. As a result, in absolute terms the difference in bias between local linear and local quadratic estimation will be smaller when an automatic bandwidth selection criteria is used as opposed to a fixed bandwidth.

$$X = a(\boldsymbol{S}) + \eta \tag{2.4}$$

Where X is the dependent variable, $\boldsymbol{S}$ is a two dimensional explanatory variable denoting location such that $\boldsymbol{S} = \{S_1, S_2\}$, $E[\eta] = 0$ and $E[\boldsymbol{S}'\eta] = 0$.

Since $a(\boldsymbol{S})$ can be approximated by the Taylor series expansion for $a(\boldsymbol{S})$ in the neighbourhood of X_0, the objective is to estimate $a(\boldsymbol{S})$ by polynomial regression with a fixed bandwidth, $\bar{h}$,

$$\begin{aligned} a(\boldsymbol{S}) \approx\ & a(\boldsymbol{S_0}) + a'(\boldsymbol{S_0})\,(\boldsymbol{S} - \boldsymbol{S_0}) + \left\{\frac{a''(\boldsymbol{S_0})}{2!}\right\}(\boldsymbol{S} - \boldsymbol{S_0})^2 + \cdots \\ & + \left\{\frac{a^p(\boldsymbol{S_0})}{p!}\right\}(\boldsymbol{S} - \boldsymbol{S_0})^p \end{aligned} \tag{2.5}$$

It is convenient to rewrite this using matrix notation as,

$$\min_{\gamma}\ (X - \boldsymbol{S_p\gamma})'\boldsymbol{W}(X - \boldsymbol{S_p\gamma}) \tag{2.6}$$

Where $\boldsymbol{W}$ is a diagonal matrix of kernel density weights, $\boldsymbol{\gamma}$ is a vector of terms relating to the coefficients on the terms in the polynomial regression,

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \; and\ \boldsymbol{W} = diag\ K_h(\boldsymbol{S} - \boldsymbol{S_0}) \tag{2.7}$$

And $\boldsymbol{S_P}$ is an $\left(N \times \sum_{k=1}^{\rho} k\right)$ matrix where the $(kp + j)^{th}$ column is given by,

$$S_P(kp + j) = \left(S_1 - S_{1,0}\right)^k \left(S_2 - S_{2,0}\right)^j \tag{2.8}$$

For $0 \leq k \leq P$ and $0 \leq j \leq P$.

The solution vector $\hat{\boldsymbol{\gamma}}$ is estimated by weighted least squares regression and is given by,

$$\hat{\boldsymbol{\gamma}} = \left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{X} \tag{2.9}$$

The first term in $\hat{\boldsymbol{\gamma}}$ relates to the parameter on a vector of ones in the regression. This provides a local estimate of $a(\boldsymbol{S})$ at $\boldsymbol{S_0}$.

The objective of the adaptive order method is to construct the best estimate of $a(\boldsymbol{S})$, defined as the estimate that minimises the mean square error (MSE). The MSE is calculated as the sum of the squared bias and variance of the estimate. The optimal set of local smoothing parameters is therefore chosen using a minimum MSE criterion.

From equation (2.9) Fan and Gijbels (1995) derive their bias and variance expressions. The sum of the bias squared and the variance is then minimised with respect to the order of local polynomial regression to obtain an estimate of $a(\boldsymbol{S})$ which trades off the bias and variance at each individual point,

$$\begin{aligned} E\left[\hat{\boldsymbol{\gamma}}\middle|\boldsymbol{S_p}\right] &= \left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{a} \\ &= \boldsymbol{\gamma} + \left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{r} \end{aligned} \tag{2.10}$$

$$Var\left(\hat{\boldsymbol{\gamma}}\middle|\boldsymbol{S_p}\right) = \left(\boldsymbol{S_p}\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\boldsymbol{S_p}'\boldsymbol{\Sigma}\boldsymbol{S_p}\left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1} \tag{2.11}$$

Where $\boldsymbol{a} = \{a(S_1), \dots, a(S_P)\}$, $\boldsymbol{r}$ is a vector of polynomial regression residuals, $\boldsymbol{r} = \boldsymbol{a} - \boldsymbol{S_p}\boldsymbol{\gamma}$, and $\boldsymbol{\Sigma}$ is the diagonal matrix $diag\{K_h^2(\boldsymbol{S}-\boldsymbol{S_0})\sigma^2(\boldsymbol{S})\}$ both of which are unknown quantities. To construct an estimate of the mean squared error, Fan and Gijbels replace these unknown quantities with approximations constructed using the data.

$\boldsymbol{r}$ is estimated by $(p+a)^{th}$ order polynomial regression where p is the order of local polynomial regression adopted and a is chosen such that the $(p+a)^{th}$ order polynomial regression provides an unbiased approximation of $a(\boldsymbol{S})$. An estimate of the conditional variance is obtained first by assuming local homoscedasticity of ϵ such that,

$$Var(\boldsymbol{\gamma}|\boldsymbol{S_p}) = \left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\boldsymbol{S_p}'\mathbf{W^2}\boldsymbol{S_p}\left(\boldsymbol{S_p}'\boldsymbol{W}\boldsymbol{S_p}\right)^{-1}\sigma^2(S_0) \tag{2.12}$$

An estimate of $\sigma(S_0)$ is then constructed from the normalised weighted residual sum of squares from the $(p+a)^{th}$ polynomial fit of the data with a fixed bandwidth, h,

$$\hat{\sigma}^2(S_0) = \frac{\Sigma\left(X-\hat{X}\right)^2 K_h(\boldsymbol{S}-\boldsymbol{S_0})}{tr\{\boldsymbol{W}^* - \boldsymbol{W}^*\boldsymbol{S_p}^*\left(\boldsymbol{S_p}^{*\prime}\boldsymbol{W}^*\boldsymbol{S_p}^*\right)^{-\mathbf{1}}\boldsymbol{S_p}\boldsymbol{W}^*\}} \tag{2.13}$$

Where ***** denotes a design matrix, similar to $\boldsymbol{S_P}$ and $\boldsymbol{W}$, with elements up to the $(p+a)^{th}$ order.

These approximations of the conditional bias and variance can then be combined to estimate the mean squared error (MSE) of the fit of equation (2.4) at each location for local polynomial regression of various orders. The optimal local order, or adaptive order, can then be chosen by finding the order that minimises the MSE at that point.

5.0 The Adaptive Order Procedure

The procedure detailed by Fan and Gijbels (1995) for adaptive order selection for estimating the density of a function is as follows,

Step 1: For each order ρ $([0] < \rho \leq R)$ and for each location obtain $\widehat{MSE}_{[0],\rho}(S, \bar{h})$.

Step 2: For each order ρ, and for each location calculate the smoothed estimated MSE by taking the weighted local average of the estimated MSE in the neighbouring $2(\bar{h}/\Delta) + 1$ locations.

Step 3: For each location $\boldsymbol{S}$ choose the order ρ_j which has the smallest smoothed estimated MSE and use a ρ_j order polynomial approximation to estimate $a(\boldsymbol{S})$.

Where the analyst is interested only in approximations up to order R. Fan and Gijbels note that this adaptive order approach reduces the sensitivity of the estimate to the choice of bandwidth allowing the adoption of a plug-in bandwidth.

6.0 Extension to Spatial Smoothing

The insights and general methodology proposed by Fan and Gijbels can be transferred to the method of spatial smoothing to obtain a good estimate of the confounding spatial data generating process, Z. In this instance the analyst is faced with the additional complication of Z being unobserved. In this section, I propose a LASSE, which incorporates Fan and Gijbel's adaptive order procedure in a SSE.

The objective, as before, is to identify the $\boldsymbol{\beta}$ parameters in the model,

$$Y = \boldsymbol{X}(\boldsymbol{S},\boldsymbol{\eta})\boldsymbol{\beta} + Z(\boldsymbol{S},\nu) + \epsilon \qquad (1.8)$$

One approach would be to use Robinson's (1988) approach and adopt Fan and Gijbels' MSE criterion to select locally adaptive smoothing parameters in estimation. However, this approach is problematic since the MSE criterion is attempting to find a best fit for Y, which can be achieved by selecting local smoothing parameters that remove all of the spatial variation from both $\boldsymbol{X}$ and Z. To achieve greater identification and efficiency an alternative approach is needed. To address this problem I develop an alternative procedure that resembles a back-fitting algorithm (Müller, 2001) to augment the methodology of Fan and Gijbels.

The LASSE procedure begins by constructing an initial estimate of the confounding spatial data, Z. Consider an initial choice of fixed bandwidth, $\bar{h}$[15] and set of orders of local polynomial regression (LPR) for each point which are initially set to zero, $\boldsymbol{\rho_0}$. An initial estimate of Z is then provided by taking the conditional spatial expectation of the residuals by LPR with local orders, $\boldsymbol{\rho_0}$,

$$\begin{aligned} \hat{Z}_{\boldsymbol{\rho_0}} &= E_{\boldsymbol{\rho_0},\bar{h}}[Y - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}}|\boldsymbol{S}] \\ &= E_{\boldsymbol{\rho_0},\bar{h}}[(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{0}})\boldsymbol{X} + Z + e|\boldsymbol{S}] \end{aligned} \tag{2.14}$$

Where,

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} = \boldsymbol{X}'Y(\boldsymbol{X}'\boldsymbol{X})^{-1} \tag{2.15}$$

The initial estimate of Z is then used to obtain a revised estimate of $\widehat{\boldsymbol{\beta}}$ by subtracting $\hat{Z}$ from $Y$and re-estimating, again using LPR, such that,

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}} = \boldsymbol{X}'(Y - \hat{Z}_{\boldsymbol{\rho_0}})(\boldsymbol{X}'\boldsymbol{X})^{-1} \tag{2.16}$$

At this point, the procedure then uses the mean squared error criteria developed by Fan and Gijbels (1995) to determine whether to increase (by one) the local order of smoothing at each point. The new set of orders, $\boldsymbol{\rho_1}$, is then used to revise $\hat{Z}$. This process is repeated until the adaptive order converges. The intention is to hone in on an unbiased estimate of $\boldsymbol{\beta}$ before removing all of the spatial variation in $\boldsymbol{X}$.

If the estimate of $\boldsymbol{\beta}$ at iteration t, $\widehat{\boldsymbol{\beta}}_{\boldsymbol{\rho_t}}$, is biased then $\hat{Z}_{\boldsymbol{\rho_t}}$ will contain both Z itself and a multiple of $\boldsymbol{X}$. In these cases it must be true that at some points the order of Z or $\boldsymbol{X}$ exceed the current local order of spatial smoothing. Hence, the mean squared error of the estimate $\hat{Z}_{\boldsymbol{\rho_t}}$ at these local points will be smaller under a larger order of spatial smoothing. This process will continue until the point where either i) the bias in $\widehat{\boldsymbol{\beta}}_{\boldsymbol{\rho_t}}$ is removed and the method selects the appropriate

[15] This could be selected, for example, by a plug in bandwidth, an asymptotic plug in, cross validation etc (Hardle and Marron (1985).

local orders required to smooth out Z or ii) $\boldsymbol{X}$ is completely smoothed and the estimate of $\boldsymbol{\beta}$ will be distinct in that it will be mean zero with large standard errors.

The iterative approach is necessary in order to avoid both $\boldsymbol{X}$ and Z being entirely smoothed out in the first stage. The adaptive method has the potential to provide improvements in efficiency and also enables identification under a third condition: where, for a fixed bandwidth, $\bar{h}$ and a set of local adaptive orders, $\boldsymbol{\rho}^*$, the global correlation between the smoothed $\boldsymbol{X}$ and Z is zero but variation remains in the smoothed $\boldsymbol{X}$.

This condition guarantees that the iterative procedure will achieve an unbiased estimate of Z from which the optimal adaptive order of smoothing can be determined and an efficient estimate of $\boldsymbol{\beta}$ can be obtained.

7.0 Monte Carlo Simulations

To investigate and compare the performance of Fixed SSE and LASSE, I undertake a number of Monte Carlo simulations. Each simulation comes from the spatial data-generating process,

$$Y = X(S,\eta)\beta + Z(\boldsymbol{S},\nu) + \epsilon \qquad (2.17)$$

Where X and Z are generated as polynomial functions of space. The observed and unobserved variables were generated as polynomial functions of two-dimensional location; this is the same approach as was adopted in the Fixed Order SSE simulations presented in Chapter 1 (see section 6.0). Two types of data generating processes were adopted: i) spatially homogeneous polynomials and ii) spatially inhomogeneous polynomials[16]

[16] Joining two polynomials creates a point where the resulting joint function is not continuous and differentiable with finite moments, violating Robinson's (1988)

In each Monte Carlo simulation I set,

$$\beta = -1 \tag{2.18}$$
$$\eta = 0$$
$$\nu = 0$$

Three versions of the model were constructed by varying the maximum local orders of $\boldsymbol{X}$ and $\boldsymbol{Z}$ to create data relating to conditions C1, C2 and C3 (page 62).

		Fixed Order SSE	Locally Adaptive Order SSE
Identification Condition	Maximum Local Order	$\hat{\beta}_{FSSE}$	$\hat{\beta}_{LASSE}$
Spatially Homogeneous Functions			
Condition 2	X=5 Z=3	-1.039 (0.345)	-1.015 (0.111)
Spatially Inhomogeneous Functions			
Condition 2	X=5 Z=2	-0.928 (0.273)	-0.955 (0.071)
Condition 3	X=5 Z=6	16.190 (57.357)	-0.902 (0.169)

Table 2.1: Simulation results for Fixed Order SSE and Local Adaptive Order SSE

Table 2.1 presents the maximum local orders of X and Z in each set of simulations alongside the results from the Monte Carlo simulations. The results in Table 2.1 demonstrate the scope of the potential efficiency gains that can be achieved by employing the adaptive order estimator when the second

---

assumptions. A gap in the data was therefore inserted between the join of the polynomials to ensure that the functions maintained an analytic quality.

identification condition is met[17]. Even in the example where both functions are spatially homogeneous the Adaptive Order SSE achieves a reduction in the standard errors of 68 per cent by adopting a lower order of smoothing at local points where Z is less variable. Similar results have also been obtained for differing polynomial orders for X and Z, although the magnitude of the efficiency savings varies depending on the degree of inhomogeneity in the omitted spatial variable. Row 3 of Table 2.1 illustrates estimation under the third identification condition. Here, the Fixed SSE fails to identify the parameters as it is forced to smooth out all of the variation in the observed variable, X. In contrast, the LASSE is able to identify β through adopting a lower order of smoothing in some locations, thus leaving variation in X to facilitate identification of the parameter.

8.0 Discussion

The Fixed Order SSE can be implemented to identify β when either of the two identification conditions, CONDITION 1 and CONDITION 2, derived in Chapter 1 of the thesis, are met. In these cases, fixed order SSE overcomes many of the limitations of the traditional spatial econometric estimators. However, it then becomes important to be able to identify whether these two conditions are met, since if they are not, the parameter estimate, $\hat{\beta}$, may not be a reliable estimate. An alternative approach is to employ a Locally Adaptive SSE (LASSE). In this chapter I have developed a LASSE in which adaptation is achieved through choosing a location-specific order of local polynomial regression. That choice is based on the method suggested by Fan and Gijbels (1995). The LASSE has the potential to provide efficiency gains over the Fixed Order SSE and enables identification under a relaxed set of conditions.

---

[17] I do not discuss the second identification condition here but would like to note that the conclusions remain unchanged. Further results are available from the author upon request.

PART II

INTRODUCTION TO PART II

Part I of the thesis was concerned with an examination of how the spatial nature of economic problems can affect analysis and inference in the context of estimating parameters from a reduced form equation. The adoption of semi-parametric estimators, which address potential omitted variable bias, is beneficial when the primary objective lies in recovering a particular parameter value, for example, in estimating willingness to pay (WTP) for a marginal change in a non-market good from a hedonic price function. In the context of Part I, the objective was to eliminate confounding spatial variables and identify unbiased and efficient parameter values. Part II moves on to consider situations that are explicitly concerned with understanding the spatial processes, which determine the distribution of households. In particular, Part II considers how to model such spatial processes in order to evaluate the impact of policy changes on the provision of endogenous public goods and, ultimately, on household welfare. To facilitate this Part II progresses from the reduced form representation of equilibrium to a class of Equilibrium Sorting Models (ESMs), which provide an agent based modelling approach that considers the interaction of households, landlords and governments across the urban landscape. Part II extends the framework of ESMs to include endogenous tenure choice and uses this new model to examine the magnitude and distribution of welfare changes resulting from policy changes.

These models were originally developed in attempts to explain observed patterns of socioeconomic and racial segregation, however the flexibility of the modelling framework lends them to the analysis of many other economic problems including addressing questions about environmental justice, deriving estimates of willingness to pay for non-market goods and predicting general equilibrium adjustments to policy changes.

In previous applications of ESMs the tenure choice of households has not been of primary concern. As a result the models have developed under the assumption that all households rent properties from absentee landlords. In a static context this simplifying assumption does not appear remarkably unwarranted, however consider the distinction between renters and owners in the face of a policy change that improves the quality of a neighbourhood. For simplicity let us assume that property prices would rise in response to increased competition. For renters this is a burden that reduces the gains in utility from the improved quality and may even force them to relocate to an area with lower quality. In contrast, an owner would experience a rise in the value of their asset. This could enable them to move to an area with even greater quality. It is clear from this simple example that the tenure status of a household will alter how it is affected by a policy change.

Building upon the existing literature, I develop a spatially explicit microeconomic model that unites household location decisions, landlord supply decisions and government policy decisions, whilst simultaneously accounting for the endogenous provision of local public goods. The model developed in this thesis extends beyond the current literature by incorporating an endogenous tenure choice and exploring the disparate impact of policy changes on renters and owners. Introducing homeownership leads to path dependency through capital gains. This is complemented by the introduction of housing stock constraints that determine how the stock of properties can adjust in response to policy changes.

The work presented in the following sections is the product of research that was undertaken with the involvement of two institutions: Arizona State University (ASU) and The Department for Transport. My work on ESMs began with an interest in the work of Nesheim (2002), which presents an ESM exploring the sorting of heterogeneous households across locations based on differences in school quality. In the same paper, Nesheim addresses the hedonic estimation problem in a single market, using restrictions developed from his ESM to introduce non-linearity into the system, which supports the identification of parameters from the hedonic price function. However, my interest in ESMs developed with a greater focus upon their usefulness for exploring feedbacks and endogeneity within the system and the implications of these processes for the magnitude and distribution of impacts upon households. This led to an interest in understanding the role of the property market in the mediation and redistribution of welfare gains. Of course, the property market influence occurs through adjustments in prices. In understanding how the property market can redistribute across households, it is therefore natural to explore how these price adjustments affect individual households. A key determinant of this is tenure choice. Indeed, tenure choice plays two central roles: First, initial tenure status determines how a household is affected by a change in price that alters the cost of its current consumption bundle. More specifically, renters face a higher cost when rental prices rise whereas homeowners are shielded from such burdens and may instead benefit from capital gains when prices rise. Second, tenure choice presents households with the flexibility to switch between tenure options in response to changes in prices.

My work on ESMs began with an exploration of the existing literature with the intention of considering tenure choice within this context. The work presented in this thesis was part funded by the Department for Transport, as such my initial objectives were driven by an interest in transport policy and the environment. However, at the onset of my research into ESMs I was fortunate in successfully obtaining an Overseas Institutional Visit award from the ESRC. The award enabled me to attend a series of advanced lectures on environmental economics and work on developing an ESM paper at Arizona State University. This was an incredible opportunity to work with academics at the forefront of the current research into ESMs including Professors Kerry Smith and Nicolai Kuminoff. During my visit I made significant headway in developing my ESM with an endogenous tenure choice and was inspired to undertake an application of this model to the issue of Mortgage Interest Deduction (MID) in the United States. This led to the third paper presented in this thesis.

Upon my return to the UK I returned to working on applying the ESM with endogenous tenure choice to problems involving localised changes in environmental quality. This included collaborating with the Department for

Transport to explore the usefulness of ESMs for analysing policy changes over medium to long-term time horizons. To provide a more fluid transition from part I to part II of the thesis the ESM papers are not presented chronologically. Instead, the work developed with the Department for Transport is presented first and the application of the model to the more general economics problem of reforming MID policy follows.

The first paper develops an ESM with endogenous tenure choice and uses it to explore the divergence between conventional welfare analysis based on a static partial equilibrium approach and a general equilibrium analysis which accounts for adjustments in the property market as households relocate in response to policy changes. To explore policy questions using this framework it is necessary to either estimate or calibrate the model. This involves making specific assumptions about the structure of the economy, for example, the number of neighbourhoods, the distribution of income and preferences, the definition of environmental quality and local public goods etc. To illustrate the mechanics of the framework, I calibrate a stylized two-neighbourhood version of the ESM model. Although this stylized model could be calibrated to any location, I calibrate it to the town of Polegate, East Sussex in 2001 using data from the Census, Expenditure and Food Survey, house price databases and Ordnance Survey GIS data layers. The choice of Polegate was motivated partially arbitrarily and partially as a result of the construction of the A27 Polegate Bypass which provides a suitable policy motivation for the analysis. The two neighbourhoods are labelled “Town Centre” and “Suburbs” and households have preferences defined over proximity from the centre of the town and exposure to road noise from traffic. I consider the impact of constructing a bypass, which runs around the outside of the suburbs and diverts traffic from the main road that runs through the town centre.

This policy application serves to demonstrate the ability of the model to capture the magnitude and distribution of welfare changes as the initial policy change induces behavioural responses that propagate through the property market affecting capital gains, housing supply and income from rental properties. The simulations illustrate the complex nature of these adjustments and indicate that a partial equilibrium analysis provides an incomplete picture of the resulting welfare effects. Furthermore, these complex adjustments distort targeted policies and redistribute benefits from renters towards owners. This suggests that environmental policies may be a poor tool for redistribution.

The second paper presented in Part II uses a calibrated version of the model to investigate the impact of reforming Mortgage Interest Deduction in the US. The model is extended further to allow homeownership to affect the provision of local public goods, leading to a model in which the provision of local public goods is endogenously determined. I simulate and compare the distributional impacts of four policy reforms: a Cap on MID, a Flat Rate MID, a Tax Rebate and a New Owner Payment. The results highlight a number of interesting issues: first, they highlight the disparity between renters and owners. Second, they illustrate the prevalence of path dependency caused through homeownership and housing stock constraints. Finally they demonstrate how the features of the model can interact with policies and create a discord between policy design and policy outcomes.

CHAPTER 3

Why welfare analysis in environmental economics simply cannot afford to ignore tenure

1.0 Introduction

Many projects and policies result in environmental impacts that differ across space. Consider as examples, the construction of a new park, the closure of a landfill site or the project that motivates this paper, the building of a bypass that directs road traffic around rather than through a town, In evaluation such projects, policymakers often wish to quantify the benefits of the change in local environmental quality in such a way as to enable a Hicks-Kaldor type comparison of the costs and benefits to be made. In addition, the importance of distributional impacts of environmental changes has become the subject of increasing attention and concern (Liu, 2000). Indeed, the remit of policymakers is increasingly driven by the environmental justice agenda (Walker, 1998, Walton and Shaw, 2003, Poustie, 2004) and has begun to consider how environmental planning and policy can be utilised as a tool for redistribution. In the US this process has been underway for the last two decades following the publication of the Commission for Racial Justice's 1987 Toxic Wastes and Race in the United States, with the incorporation of environmental justice considerations into the working directives of the Environmental Protection Agency and other environmental planning agencies (Walker). In the UK, environmental justice concerns have developed more slowly. Following the publication of the Friends of the Earth (FoE) (1999) report on pollution injustice, there has been a growing body of evidence linking environmental problems and social injustices in the UK (ESRC Global Environmental Change Programme, 2001, Walker, 2003). In addition, increased attention to environmental justice in EU and UN[18] directives and initiatives like the Aarhus Convention (1998)[19] have

---

[18] Declaration of the UN Conference on Environment and Development, http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm

World Charter for Nature UNGA

[19] Aarhus Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters (1998).

Convention on Biological Diversity 31 ILM (1992)

Convention on Civil Liability for Damage Resulting from Activities Dangerous to the Environment, 32 ILM (1993) raised the profile of these issues and encouraged the UK government to incorporate environmental justice into its policy and decision making (Walker and Bickerstaff, 2000).

To illustrate the challenges involved in assessing the magnitude and distribution of welfare effects, consider the example of the construction of a bypass road that seeks to reduce the flows of through traffic from the centre of a town. Implementation of that policy would result in a variety of spatially defined impacts. Those living closest to the town centre, for example, might enjoy the greatest improvements in quality while those living closest to the new bypass might experience the largest losses. Conventional project appraisal would seek to establish the benefits of building the bypass by estimating how much residents in those different locations would be willing to pay, (or potentially willing to accept in compensation) for the changes in environmental quality that they experience. I shall refer to this standard practice as the *static approach* since it assumes that the residents make no behavioural changes in response to the changing conditions brought about by the construction of the bypass. The suitability of this static approach for evaluating policy changes that precipitate wider behavioural responses is the subject of an on-going debate (Just et al., 2004, Klaiber and Smith, 2010) with many arguing that the approach is insufficient and inaccurately represents the underlying welfare effects (Goulder and Roberton C. Williams, 2003, Smith and Carbone, 2007, Carbone and Smith, 2008, Bayer et al., 2009, Klaiber and Phaneuf, 2010). In particular, as a result of Tiebout's (1956) hypothesis stating that households purchase environmental quality through their location, environmental changes may alter household behaviour through, for example, relocation in the property market, adjustments in the labour market and changes in recreation site choices (Kerry Smith et al., 2004, Sieg et al., 2004, Tra, 2010).

Residential location, for example, is mediated through a set of adjustments in the property market. These adjustments are extremely complex and can have a substantial impact upon both the magnitude and distribution of welfare effects resulting from an environmental policy. Environmental economists are beginning to get a handle on the mechanisms at work through the estimation and simulation of equilibrium sorting models (ESMs). These models are particularly suited to this area as they explicitly model and characterise the equilibria resulting from the interaction of households, landlords, government and property suppliers in a system of local neighbourhoods (Epple and Platt, 1998, Bayer et al., 2004, Ferreyra, 2007, Kuminoff et al., 2010). Over the last decade, the complexity of ESMs and the range of applications has expanded very rapidly (see (Kuminoff et al., 2010) for a thorough review of the development of this literature). Smith et al. (2004) , Sieg et al (2004) and Tra (2010), for example, use the ESM framework to evaluate policies relating to air quality in the Los Angeles area. Ferreyra (2007) explores the impact of large-scale private school voucher programs in the Chicago metropolitan area using an ESM approach. Likewise, Walsh (2007) uses the ESM framework to examine policies aimed at preventing urban sprawl in North Carolina. In each of these cases, the authors have demonstrated how feedback effects have major policy implications. For example, Walsh’s policy simulation shows how increasing the amount of land in public preserves can actually decrease the total amount of land in open space in the metropolitan area. The structure of ESMs permits them to be used not only to estimate preference parameters, but also to use these estimates (or calibrated values in their place) to solve the model under counterfactual policy scenarios, taking into account the general equilibrium impact of policy change through spill-overs, feedbacks and endogeneity.

ESMs attempt to reflect the true complexity of the underlying problem being studied. The purpose of this paper is to contribute to that endeavour. In particular, while much progress has been made in terms of developing the ability of the ESMs to capture the discrete nature of the supply of environmental amenities and reflect a wide range of household heterogeneity, other important aspects remain undeveloped. Specifically, this paper seeks to extend the ESM framework to incorporate two important elements of the property market: housing supply and tenure choice.

In an ESM of a property market, an equilibrium of the model is defined as a set of set of prices that equate the quantities of housing supplied and quantities of housing demanded. Whilst the specification of the demand side of the model has been well-studied (Epple and Platt, 1998, Bayer et al., 2004, Ferreyra, 2007), the supply side has often been specified in a highly simplified manner. In most empirical applications of ESMs the usual assumption is that either the total stock of housing remains fixed (Bayer et al., 2004, Kerry Smith et al., 2004, Ferreyra, 2007, Tra, 2010), or there exists some constant elasticity of supply function for the construction of new housing (Sieg et al., 2004, Walsh, 2007, Epple and Ferreyra, 2008, Epple et al., 2010). The implications of simplifying the housing supply function depend upon the way in which housing is defined within the model. Different approaches have been adopted in different ESMs, in this paper I focus upon the model developed by Epple and Platt (1998). In this model housing is defined as a continuous good using an indexing procedure to convert expenditure on housing into a quantity measure defined in terms of homogeneous quality units of housing (the mechanics of this conversion is detailed in Sieg et al (2004)). In this context, fixing the total stock of housing implies restricting the total number of homogeneous quality blocks. In many of the empirical applications these blocks can be repackaged at no cost into any of the other possible combinations. This allows blocks to be transferred between households but with a fixed housing supply new development is not permitted. The first objective of this paper is to develop a more realistic medium to long run housing supply specification.

Secondly, the majority of empirical ESMs assume that all households are renters who pay rents to an absentee landlord (or landlords). While this approach reduces the computational complexity of the model to be solved, there are a number of good reasons to believe that renters and owners might be affected differently by environmental policy changes. To illustrate, consider an environmental improvement that leads to an increase in demand for housing in a particular neighbourhood. As demand increases, prices also increase so as to clear the property market. Renters in the neighbourhood would face increases in rents, which, in the medium term, may force them out of the area. In contrast, for homeowners in the neighbourhood a price rise represents an additional option, not only can homeowners choose to stay in the area and enjoy the improvement in environmental quality, they may instead choose to sell their housing, thus benefitting from capital gains, and relocate to another neighbourhood. In some cases, this process may induce the gentrification of a neighbourhood in light of a local improvement in quality (Banzhaf and Walsh, 2004, Sieg et al., 2004). Clearly, these mechanisms have important implications for the distribution of welfare effects across renters and owners. Moreover, changes in prices lead to changes in rental revenues. In an all-renter model these benefits flow directly to absentee landlords. However, in this paper a more realistic design is adopted in which rental revenue is recycled to households. Since, empirically, wealthier households are more likely to own a share of the stock of rental housing and are also more likely to own their own home, the model developed here provides a second channel through which the welfare effects of a policy may diverge between owners and renters.

Our discussion thus far has sought to establish the importance of understanding the role of tenure choice in order to capture some of the feedback mechanisms and model the distribution of welfare effects following a policy change. Furthermore, our illustration demonstrates how wealthier households might capture the benefits of an environmental improvement even when the policy itself is initially spatially targeted towards poorer households. Conversely, a reduction in environmental quality may lead to benefits for poorer households through reductions in rental costs that offset the environmental loss whereas homeowners face either lower environmental quality or capital losses if they wish to relocate to another neighbourhood. In practice, these capital losses are often covered by compensation. For example, in the UK by Part 1 of the 1973 Land Compensation Act, which reimburses homeowners for reductions in the value of their properties resulting from changes in environmental quality resulting from public works. Since compensation is paid to homeowners only, there is good reason to believe that current policies work to protect owners but not renters from changes in environmental quality.

In this paper I develop an ESM with an endogenous tenure choice, rental revenue recycling, a more realistic housing supply function and a compensation mechanism. That model provides a means of examining the magnitude and distribution of the welfare changes caused by a policy change. The model is illustrated through a simplified simulation exercise calibrated to reflect the features of a real environmental policy change: the construction of the Polegate A27 Bypass, in the county of East Sussex in the UK. The mechanisms at work in the model are complex. To intuitively demonstrate the key features/capacities the calibrated model presented in this paper is a highly simplified version of the model with just two-neighbourhoods. The simulation exercises reveals significant differences between the welfare effects of the construction of the Polegate bypass as identified by a static analysis and those identified in the general equilibrium analysis. Both the magnitude and distribution of the welfare effects are altered once medium to long-term adjustments are accounted for. The model suggests that those adjustments tend to channel the benefits of environmental improvements to owners and away from renters. Rental revenue recycling and compensation for homeowners further exacerbate that process. The paper concludes with a discussion of the implications of our findings and a commentary on the further developments that are needed to enable ESMs to be adopted by policymakers as a tool for accurately assessing the magnitude and distribution of welfare effects of policies or projects that result in localised environmental change.

2.0 Housing Supply and Tenure Choice

Following Binner and Day (submitted to the Journal of Public Economics), this paper uses the theoretical framework of Epple and Platt (1998) as a point of departure. Two major innovations are introduced. First, a more realistic medium to long run housing supply specification is developed. Second, an explicit tenure choice is introduced and households are permitted to alter their tenure status in response to changes in conditions.

2.1 Housing Supply

The nature of the housing supply function depends critically on the time horizon under consideration. In the short to medium term housing is likely to be fixed. In the medium to long term there is scope for new construction to take place and for the existing housing stock to be modified.

In ESMs of the property market, the standard assumption is that housing is supplied in homogeneous quality/quantity blocks that can be purchased at some market price. Within the model households' decide how many of these blocks they wish to purchase. Buying more blocks imitates a household purchasing a better quality/larger sized property.

With this building block representation of housing supply responses to changes in market conditions are typically handled by allowing for changes in the number of blocks available for purchase. For example, a number of authors specify a constant elasticity of supply (CES) supply function for housing blocks (Epple, Romano and Sieg, 2010, Fernandez and Rogerson, 1998, Hallstrom et al, 2003). Of course, such a specification implicitly assumes that the supply of housing in any neighbourhood can be expanded without limit, an assumption that is clearly unrealistic. Another popular specification, which avoids that critique, is that the number of housing blocks is fixed (Bayer et al., 2004, Kerry Smith et al., 2004, Ferreyra, 2007, Tra, 2010). Both these specifications, however, make the untenable assumption that housing blocks can be costlessly repackaged from one particular configuration of properties into another: a neighbourhood of 50 two block properties, for example, can be transformed without cost into a neighbourhood with 100 one block properties.[20]

In this paper two housing supply specifications are considered and compared.

---

20 The housing stock is comprised of supply from a number of submarkets including new builds and conversions. In the UK conversions accounted for 5,240 (3.9 per cent) of the 134,900 increase in the number of properties in 2011-12 (Housing Statistical Release, 2012) Although repackaging, also referred to as conversion, has been discussed in a number of papers which point to costs arising from bargaining, regulatory requirements and construction costs (Capozza and Helsley, 1989, Pogharian, 1990, Rothenberg et al., 1991, Montgomery, 1992, Maddison, 2000), the housing supply literature has focused on new construction and, as a result, it is difficult to find any empirical work quantifying the extent of these costs.

First, the conventional fixed supply of total housing blocks. Second, a new specification that allows flexibility in the supply of housing blocks within a neighbourhood but also limits the number of individual properties constructed from those blocks that can be accommodated within that neighbourhood .

2.2 Tenure Choice

The vast majority of empirical applications of ESMs treat all households as renters (Kuminoff et al., 2010) , the exceptions are Epple and Platt (1998) who consider a fixed proportion of the population to be owners with the proportion of owners increasing with income and Bayer et al (2005) who treat tenure as a characteristic of the housing stock. In reality, there are a number of tenure types from which a household can choose with a key distinction being between ownership and renting21. To integrate an endogenous tenure choice into an ESM, households are provided with a choice between renting housing or purchasing it. The appeal of owning as opposed to renting one's home may differ across the population. In reality the tenure decision is dynamic and multifaceted, concerning issues such as borrowing constraints, attitudes to risk, wealth, transaction costs, the utility derived from achieving social status through ownership, preferences for property maintenance and so on. In this model I condense these dimensions into a single metric, a preference for homeownership, which captures the additional utility that households derive from housing units when they own them. This preference parameter is defined such that, for a fixed price for housing blocks, some households may prefer renting while others may prefer owning. In addition, purchasing housing requires a household to take out a mortgage, leading to the payment of a mortgage interest payment that is not incurred by renters. The size of mortgage needed by a household is determined in part by the loan-to-value ratio of their mortgage. In the model, loan-to-value

[21] Although this point has been raised in a number of theoretical papers (Epple & Platt 1998, Bayer et al. 2004) it has received little formal attention in the development of ESMs.

ratios differ across households partly in response to differences in a household's income.

Tenure choice introduces a second discrete component into the residential decision: not only do households have to choose between neighbourhoods, but they also have to choose whether to rent or purchase their home. As a result, richer patterns of substitution are allowed by the model. Since, tenure status has important ramifications for how a household is affected by policy changes, how it can respond to them and thus how its welfare is affected, the choice of tenure is an integral part of a household's residential choice. Moreover, by explicitly accounting for tenure status it is possible for potential capital gains (and losses) to be modelled when owners consider relocating. As a result, the model provides a means of examining the differential impact of policy reform on renters and owners. As far as the author is aware, this is the first time that the welfare implications of these tenure based property market features have been explored in the context of an ESM. More importantly, in the simulations reported in this paper, tenure is shown to have a very significant influence on the welfare effects of policy changes. As such, I contend that the distinction in tenure status is of crucial import in policy analysis.

3.0 The Model

Consider a closed spatial economy consisting of households, $i = 1... I$. The model is "closed" in as much as households are not allowed to migrate into or out of the economy and "spatial" in as much as households make a choice as to the location in which they reside within the economy.[22]

---

22 The closed economy model discussed here makes three assumptions: i) there is no immigration in, ii) there is no emigration – though with the Cobb Douglas utility is bounded by zero since households can opt not to consume any housing which is equivalent to a zero utility outside option and iii) there is no population growth.

In the event of deterioration in quality, the availability of an outside option providing some positive level of utility could lead to more pronounced reductions in price to

Households differ in their incomes, y, preferences for quantity of housing, β, and preferences for homeownership, θ. The preference for homeownership parameter, θ, represents the private returns to homeownership anticipated by a household. The motivation for including this parameter stems from i) the freedom homeowners have to modify their housing to their personal tastes, ii) the satisfaction of achieving homeownership status and iii) the financial returns from capital gains. These private returns may differ across households for a number of reasons: i) households may differ both in their ability to modify housing and in the costs that they face to do so and ii) they may differ in the satisfaction that they derive from the status of owning their home. The distribution of household types across the population is defined by the joint multivariate density function, $f(y, \beta, \theta)$.

The economy is defined by a geographical region which is divided up into a set of spatially discrete neighbourhoods, $j = 1, \dots, J$. Each neighbourhood provides a vector of K local public goods[23], $\boldsymbol{g_j} = \{g_{j,1}, g_{j,2}, \dots, g_{j,K}\}$. The provision of

---

maintain demand as households emigrate. Likewise, quality improvements would draw in households from the outside population (though this would be mediated through moving and transaction costs) and push up prices, potentially leading to the displacement of incumbent residents by incomers.

It would be possible, though not trivial, to extend the model to incorporate these additional complexities. The particular difficulties lie in characterising the income and preference distribution of the outside population, characterising the outside option and how it evolves in response to policy changes and in defining welfare in these settings, for example extended in the model to include migration and an "outside" population raises questions whether a local authority is interested in welfare increases accruing to new residents and those affecting households who leave the area? These questions have ethical and political dimensions as well as being methodological.

[23] In this paper local public goods are considered to be exogenously determined, however, more generally, local public goods may be comprised of *exogenous* elements,

these public goods is assumed to be homogeneous within a given neighbourhood. Households decide in which of the *J* neighbourhoods they will reside.

3.1 The Demand Side

To reside in neighbourhood *j* household i must purchase or rent housing in that neighbourhood. The decision to rent, R, or own, O, housing is referred to as the *tenure choice*. The set of tenure options is $T = \{R, O\}$. Accordingly, the model is characterised by households' residential choices, defined by a neighbourhood and tenure bundle, $\{j, t\}$. The set of residential choices is given by $C = J \times T$.

Households must also choose how much of their income to spend on housing[24]. Following previous treatments (Epple and Romer, 1991, Epple and Platt, 1998, Epple et al., 2001, Bayer et al., 2004, Ferreyra, 2007), housing is defined as a homogeneous good that can be purchased at a constant per unit price within a neighbourhood[25]. Housing is supplied in neighbourhood *j* at a per unit property price, p_j. The quantity of housing demanded by household i in neighbourhood *j* for tenure type t is denoted,

$$h_{i,j,t} = h_{j,t}(\boldsymbol{p}, \boldsymbol{g}; y, \beta, \theta) \tag{3.1}$$

$z_{j,k}$, such as environmental quality, transport infrastructure or proximity to a central business district, and *endogenous* elements, $q_{j,k}$, the level of provision of which is dependent upon the composition of the set of households that reside within the neighbourhood, such that $g_j = \{z_{j,1}, \dots, z_{j,k}, q_{j,k+1}, \dots, q_{j,K}\}$. See Nesheim (2002), Nechyba (2003b) and Ferreyra (2007).

[24] In the model, households choose the number of housing units that they wish to consume, a decision approximating real life choices over the size and type of house to buy or rent.

[25] In reality housing is not homogeneous, however, as Sieg et al (2002) illustrated, if housing enters the utility function through a sub-function that is homogeneous degree one, it is possible to construct a ``housing quantity" index tantamount to an empirical analogue to the homogeneous housing unit, *h*.

Where $\boldsymbol{p}$ is a vector of per unit property prices in each of the neighbourhoods and $\boldsymbol{g}$ is a vector containing the local public good index for each neighbourhood.

To become a homeowner a household must take out a mortgage[26] and pay mortgage interest, m_i, to the lender. Differences in the mortgage rate across households can be interpreted as representing the differing abilities of households to secure a mortgage and bargain for cheaper interest rates. Mortgage interest is paid only on the amount borrowed, which is equal to the product of the loan-to-value ratio, δ_i, and the value of the housing purchased, $p_j h_{i,j,t}$. Differences in the loan-to-value ratio represent differences in the ability of households to make a down-payment on housing.

Aggregate demand for housing units conditional on residential choice, $\{j,t\}$, is calculated by integrating over all households,

$$H_{j,t}^{D} = \int\int\int h_{j,t} \,.\, f(y,\beta,\theta)\,dy\,d\beta\,d\theta \tag{3.2}$$

3.2 The Supply Side

Housing is produced using land and non-land inputs. The supply of land available for supplying housing, L_j, may differ between neighbourhoods. As a result, the supply of housing may also vary between neighbourhoods. The housing supply function for a particular neighbourhood is denoted,

$$H_j^s = f_2(L_j, p_j, D_j, n_j) \tag{3.3}$$

L_j is the share of total land available in neighbourhood j and p_j is the price of housing in neighbourhood j. The population size is denoted by n_j. The largest population that can be supported by a neighbourhood is capped by introducing a capacity, D_j. As the population increases towards the capacity the housing

[26] For simplicity, the model assumes that all households initially take out a mortgage.

supplied at a given price declines as the marginal cost increases. This specification enables density restrictions to be included in the model so as to capture constraints on the number of properties that can be built in a neighbourhood more realistically.

Traditionally ESMs assume that absentee landlords own the stock of rental housing. This approach leaves the model open with some of the costs and/or benefits of policy change flowing out of the economy. It is possible to alter the model set up to allow rental revenues to be returned to households, as if some of those households were themselves landlords. In the model, that process is approximated by treating all or some of the stock of rental housing as being communally owned and specifying the shares that households have in this stock. Households each receive a share, π_i, of the total rental revenues, $\Pi = \sum_j p_j H_{j,R}^S$, in accordance with their shares in the stock. Accounting for such rental revenue recycling enables us to consider how changes in rental revenues feed back through the economy.

3.3 Government

In this paper the role of government is limited to decision making about the location of a bypass that is known to impact environmental quality and the administration of compensation[27]. The payment of compensation is modelled on Part 1 of the Land Compensation Act 1973, whereby homeowners are compensated for reductions in the value of their housing resulting from the environmental damage generated by the public works.

3.4 Local Public Goods

[27] This is not a restriction of ESMs in general. The models are capable of encompassing a variety of behaviour, for example governments may levy income and property taxes, fund expenditure on public goods, set production standards or provide subsidies.

Households derive utility from the combined provision of the local public goods, which is represented by the index of local public goods provision,

$$g_j = \sum_{k=1}^{K} \gamma_k g_{j,k} \tag{3.4}$$

Where γ_k is the weight placed on the k^{th} element in g_j[28]. For simplicity, the calibrated simulation exercise developed in the following sections assumes that the index combines levels of two exogenous public goods,

$$g_j = g_{j,1} + \gamma g_{j,2} \tag{3.5}$$

γ is the weight that households place on $g_{j,2}$ relative to $g_{j,1}$. It is assumed that γ is the same for all households and across neighbourhoods, such that households agree on the ranking of neighbourhoods in terms of their local public good provision.

3.5 Household Optimisation

Households derive utility from their access to local public goods, g_j, their consumption of housing, h, and other consumption, c. Tenure status affects the way in which households enjoy the flow of services provided by housing. All else equal, household i derives the same level of utility from owning h units of housing as from renting $\theta(h)$ units. For simplicity and clarity, households are assumed to have the same preference for local public goods, α. Household utility is defined by the function,

$$U_{i,j,t} = U(g_j, h_{i,j,t}, t, c; y_i, \alpha, \beta_i, \theta_i) \tag{3.6}$$

[28] As g_j is an index, these weights can be normalised such that $\gamma_1 = 1$.

The optimisation problem of household i can be decomposed into two stages. First, households calculate their optimal housing and consumption choices for each neighbourhood and tenure bundle. The conditional maximisation problem is,

$$\max_{(h,c|j,t)} U(g_j, h_{i,j,t}, t, c; y_i, \alpha, \beta_i, \theta_i)$$

$$s.t.$$

$$y_i + \pi_i = \begin{cases} p_j h_{i,j,t} + c_i, & t = R \\ (1{+}m_i\delta_i) p_j h_{i,j,t} + c_i, & t = O \end{cases} \tag{3.7}$$

Where p_j is the price of a unit of housing in neighbourhood j. The owner's budget constraint also includes the term $(1 + \delta_i m_i)$ indicating that a homeowner must pay for each unit of housing and also for mortgage interest.

The budget constraints in equation (3.7) assume that households do not initially own any housing. When we consider adjustments to new equilibrium it is important to consider the potential for capital gains. For an existing homeowner the model accounts for three ways in which capital gains can accrue, i) when a homeowner sells some but not all of their housing units and stays in the same neighbourhood and ii) when a homeowner sells their housing in one neighbourhood and becomes an owner in a different neighbourhood, and iii) when a homeowner sells their housing in one neighbourhood and becomes a renter. The budget constraint for household i can be expressed formulaically as,

$$y_i + \pi_i + p_j{}^1\left[h^0_{i,j^0,O} - h^1_{i,j^0,O}\right] = (1 + \delta_i m_i) p_j{}^1 h^0_{i,j^0,O} + c^1, \qquad \textit{staying as an owner}$$

$$y_i + \pi_i + \left[p_j{}^1 - (1 + \delta_i m_i) p_j{}^0\right] h^0_{i,j^0,O} = (1 + \delta_i m_i) p_j{}^1 h^1_{i,j^1,O} + c^1, \qquad \textit{moving as an owner}$$

$$y_i + \pi_i + \left[p_j{}^1 - (1 + \delta_i m_i) p_j{}^0\right] h^0_{i,j^0,O} = p_j{}^1 h^1_{i,j^1,O} + c^1, \qquad \textit{moving and now renting}$$

Where superscript 0 denotes a baseline variable and 1 denotes a new variable choice. In each equation the third expression on the right hand side denotes the capital gains made on units of housing sold the first expression on the left hand side represents the new expenditure on housing, including mortgage interest.

Maximisation of the direct utility function subject to the constraints yields the following conditional indirect utility functions,

$$V_{i,j,t} = \begin{cases} V(p, g; y, \alpha, \beta), & t = R \\ V(p, g; y, \alpha, \beta, \theta), & t = O \end{cases} \tag{3.9}$$

Second, households select the neighbourhood and tenure choice that provides the greatest level of utility.

3.6 Equilibrium

An equilibrium in this model is defined by a set of neighbourhoods, J, a one to one correspondence of households to neighbourhoods and an associated set of property prices, $\boldsymbol{p} = \{p_1, \dots, p_J\}$ for each neighbourhood, such that,

1. Each household resides in the neighbourhood that maximises its utility given the equilibrium vector of prices and endogenous public good provision.
2. All housing markets clear, $H^S_{j,t} = H^D_{j,t}\ \forall\ j, t$.

The introduction of an endogenous tenure choice and preferences for homeownership generalises the pure characteristics ESMs developed by Epple and Platt (1998) and does not alter the underlying properties that support the existence of equilibria. Namely, the single crossing, boundary indifference, ordered bundles and stratification properties continue to hold[29] (Epple and Platt, 1998, Epple et al., 2010).

---

29 Epple and Romer (1991) demonstrated the existence and properties of a pure characteristics equilibrium sorting model. These properties are: i) stratification - each neighbourhood is occupied by households within a certain set of income and preferences, ii) boundary indifference - ranking neighbourhoods by price, there exists a locus of households defined by their income and preferences who are indifferent between any two consecutive neighbourhoods and iii) ordered bundles - the price ranking of neighbourhoods is the same as the ranking of neighbourhoods by their public

4.0 Response Pathways to Exogenous Policy Change

At this point it is instructive to contemplate the types of responses that can be accommodated by the model and the distinction between renters and owners. Imagine an environmental improvement that leads to a rise in house prices in a neighbourhood. For renters that means higher rents, for owners it means they can sell their property at a higher price.

Consider first what those price increases mean for renters. Increases in the cost of renting mean that they are no longer able to afford their current housing and consumption bundle. A number of options are available. First, the renter could remain renting in the same neighbourhood but consume fewer units of housing. Alternatively, the rise in rental price may make homeownership relatively cheap and thus more desirable in which case the renter may wish to become an owner in the same neighbourhood. In addition, the renter may consider relocating to a different neighbourhood and may continue to rent or begin owning.

Now consider the consequences of a similar rise in prices on homeowners. The rise in prices has rather different implications for this group. Homeowners own their properties outright. As such, homeowners can always afford their current housing and consumption bundles, even when property prices rise. In this way, homeowners are shielded against price fluctuations. A rise in prices cannot make an existing owner worse off, instead it presents an opportunity to sell units of housing and use the capital gains to increase consumption or to relocate to a neighbourhood that provides more desirable public goods, in which case they might continue to own or may become renters. The choices made by each household will be determined by its preferences, characteristics and existing ownership of housing.

---

goods index. These properties hold under the assumption that indifference curves exhibit the single crossing property and utility is monotonically increasing in its attributes.

In addition to enabling us to distinguish between owners and renters, the ESM model developed in this paper also considers the impact of policy change on rental revenues that are returned to households. The total value of rental revenues is dependent on the distribution of households resulting from the sorting process. Moreover, the rental revenues received by each household may alter in response to policy changes as households adjust their tenure choice and housing consumption and property prices rise or fall. Changes in rental revenues will have a more significant impact on households with a larger share of the rental stock; these are likely to be wealthier households. In turn, changes in rental revenues will feedback into the economy as changes in consumption through an income effect. The model easily extends to more general endogenous public goods.[30]

A notable limitation of this model is that is it does not incorporate labour market decisions. There is growing evidence to suggest that transport and infrastructure improvements can impact employment decisions, productivity and wages (Gibbons et al., 2012, Sanchis-Guarner, 2012). In future work these additional response pathways could be incorporated by moving to a dual market model that simultaneously addresses location and labour market decisions (Kuminoff 2009).

5.0 Calibrating the Model

The simulations presented in the following sections draw on work carried out for the Department for Transport (DfT), funders of this PhD studentship. The objective of that work is to examine the usefulness of ESMs for real world project evaluation. As part of that research a multiple-neighbourhood version of

[30] For example, other models have considered the impact of social interaction between the occupants of a neighbourhood, such as the influence of peer group effects, on education quality and human capital accumulation and the influence of political systems of social choice on the level of provision of a local amenities, such as education spending per pupil and expenditure on law enforcement. See Bergstrom and Goodman (1973).

the ESM described above has been developed and calibrated to investigate a real DfT project: construction of the A27 Polegate bypass. The results of this analysis are presented in Appendix A.

While multiple-neighbourhood models are well-suited to the task of evaluating a real world policy they tend to result in very complicated patterns of substitution across the numerous neighbourhoods. In order to draw insights as to the importance of tenure choice as revealed by the ESM, it is significantly more informative to work with a simplified two-neighbourhood version of the model. Accordingly, the analyses in this section present results from a stylised two-neighbourhood simplification of the more general Polegate analysis presented in the Appendix.

The model is calibrated using census information giving homeownership rates in Polegate in 2001, the year before the bypass opened, and using data the Post Opening Project Evaluation (POPE) A27 Polegate Bypass report (2009) which details the population shares that were affected by the construction of the bypass. Baseline neighbourhood prices for a unit of homogeneous housing were derived by taking the neighbourhood specific intercepts from a fixed effects hedonic regression. That regression used property price data from 2000 provided by the UK Land Registry. That price data was matched, using GIS techniques, to further information on property characteristics from OS Mastermap and Edina[31].

5.1 The Economy

In the simplified calibration the economy is divided into two regions: one comprising the town centre, which has a main road running through it, and the second comprising the suburbs. The town centre is closer to a range of amenities (for illustrative purposes these amenities could include parks, shops, a school, a medical centre etc.). However, whilst properties located in the town centre

---

[31] See Kuminoff et al (2010) for a discussion of this methodology.

benefit from greater access to these positive local amenities they are also exposed to greater road noise as a result of traffic on the busy throughway.

5.2 Households

Household utility is represented by a Cobb Douglas utility function, such that,

$$U_{i,j,t} = g_j^{\alpha} \theta_{i,t} h^{\beta_i} c^{1-\alpha-\beta_i} \tag{3.10}$$

Household preferences for public goods, α, and homeownership, θ_i, are assumed to be independent of their income and housing expenditure. With a Cobb Douglas utility function β can be interpreted as the share of income that a household commits to purchasing housing.

A population of 1600 households were drawn from a joint bivariate distribution, $f(\ln(y)\,,\beta)$, of incomes, y, and expenditure on housing, β. The sample size was chosen to replicate the actual number of households affected by the bypass, as reported in the POPE. The parameters of the joint distribution were estimated using data from the Expenditure and Food Survey 2001-2, which provides a breakdown of gross weekly income and expenditures on housing including mortgage costs, maintenance and depreciation (2001 GB Pounds). A joint lognormal bivariate distribution was fitted to the data by maximum likelihood estimation. This process reveals a negative correlation between the two variables indicating that lower income households spend a larger proportion of their income on housing. The resulting parameter estimates were,

$$\left(\mu_{\ln(y)}, \mu_{\beta}\right) = (9.83, 0.17)$$

$$\Sigma_{\ln(y),\beta} = \begin{pmatrix} 2.72 & -0.07 \\ -0.07 & 0.03 \end{pmatrix} \tag{3.11}$$

The loan-to-value ratio was calibrated using data from the FSA Mortgage Product Sales Data Trends Report (2007). This was used to simulate the correct proportion of households with loan-to-value ratios within given intervals (0.00-

0.67, 0.67-0.83, 0.83-0.99, 0.99-1.00). In the absence of detailed information on loan-to-value ratios by income group, loan-to-value ratios (ordered from lowest to highest) were assigned to households (ordered from highest income to lowest) and a mean zero random component was added.

Preferences for homeownership, θ, were drawn from a log normal distribution,

$$\begin{gathered} \ln(\theta) \sim N(\mu_\theta, \sigma_\theta^2) \\ (\mu_\theta, \sigma_\theta^2) = (0.80, 0.03) \end{gathered} \tag{3.12}$$

The mean and variance of the distribution were calibrated using a maximum likelihood procedure designed to minimize the difference between observed and predicted population shares and homeownership rates in the two neighbourhoods at the derived housing prices.

5.3 Local Public Goods

Neighbourhoods are differentiated by their distance to the town centre and road noise levels. The local public goods index is given by,

$$g_j = g_{j,1} + \gamma g_{j,2} + \xi_j \tag{3.13}$$

Where $g_{j,1}$ is distance from Polegate town centre (measured as a negative) and $g_{j,2}$ is reduction in noise levels from a maximum of 100 dB. The proximity to Polegate town centre was calculated using ArcGIS. Noise pollution, is measured using the average 18-hour decibel level and is directly affected by the creation of the bypass. The baseline and post-bypass noise levels are calibrated using information from the A27 Polegate Environmental Statement.

The parameter, α, relating to preferences for local public goods was set using the technique for calibration with non-market goods detailed by Carbone & Smith

(2008).[32]

Using this methodology the parameter can be calibrated using the result,

$$\alpha = \frac{p_{g_1} g_1^0 + p_{g_2} g_2^0}{y + p_{g_1} g_1^0 + p_{g_2} g_2^0} \tag{3.14}$$

Where p_{g_1} and p_{g_2} are implicit prices for distance to Polegate centre and reductions in noise pollution respectively, and subscript 0 denotes a baseline level of the corresponding variable. Using this approach the weighting parameter in the local public goods index, $\boldsymbol{g_j}$, can also be calibrated using implicit prices,

$$\gamma = \frac{p_{g_1}}{p_{g_2}} \tag{3.15}$$

The implicit prices can be derived from any number of studies of willingness to pay for proximity to the town centre and for reductions in road noise. This paper uses the estimates of implicit prices for improvements in noise pollution from the Day et al (2002) hedonic study of the Birmingham area and the estimate of the implicit price for proximity to Polegate centre from a hedonic study[33]. Evaluating the implicit prices at the 2000 mean values for air quality and

32 The Carbone and Smith (2008) methodology employs implicit prices to calibrate the preference for public goods. The procedure seems a little at odds with the general equilibrium nature of the equilibrium sorting model: The calibration technique itself is valid because it assumes that the system is currently in equilibrium and uses implicit prices, which are relevant when evaluating a marginal change to infer preference for the public good using the current hedonic price function. However, the approach is potentially problematic since the process of household sorting may cause endogeneity, which needs to be accounted for when estimating implicit prices to ensure that unbiased estimates are obtained. A preferred method would be to use micro level data to jointly estimate the parameters of the model.

33 In the calibration of this two-neighbourhood model I use data on the sale of properties in Polegate in the year 2000 from Zoopla.co.uk.

proximity to the central business district provides a calibrated value of 0.11 for α and 0.02 for γ.

The model is divided into neighbourhoods with equal shares of the population and homeownership rates of eighty per cent (based on census data for 2000). The unobserved public good, ξ_j, was calibrated using a maximum likelihood procedure designed to minimize the difference between observed and predicted population shares and homeownership rates in the two neighbourhoods for the derived per unit housing prices. The resulting indices show that the baseline index of public good provision is greater in the suburbs, consistent with the higher property price in the suburbs.

5.4 Housing Supply

There are two components to housing supply: baseline housing supply and the medium to long term housing supply function. In the baseline, I assume that the economy is in long run equilibrium where the prevailing housing supply reflects long run housing supply. In that case, the housing supply is equal to, and can therefore be inferred from, the total housing demand in each neighbourhood.
To evaluate policy changes a housing supply function is specified that defines how the supply of housing responds to the market changes precipitated by the construction of the bypass. Ideally this housing supply function would come from detailed information on the true housing supply functions, but that information was not available to us. Accordingly, the housing supply function used in the model is developed using plausible assumptions. That supply function is linear, assumes a positive elasticity of supply and increasing marginal cost associated with the production of additional housing units. In addition, the supply function includes an element reflecting the costs of constructing additional properties out of housing blocks and a population capacity limiting the total number of individual properties that can be supported by a neighbourhood. The resulting housing supply function is defined within the limits $0 \leq POP_j \leq D_j POP_j^0$,

$$H^s(p_j) = \begin{cases} a(p_j - p_j^0) + H_j^0, & POP_j \leq POP_j^0 \\ a(p_j - p_j^0) + H_j^0 - b\dfrac{(POP_j - POP_j^0)}{POP_j^0}, & D_j POP_j^0 \geq POP_j \geq POP_j^0 \end{cases} \tag{3.16}$$

Where POP_j and POP_j^0 are the new and old population size of neighbourhood j, H_j^0 is the housing supply in the baseline, a and b are constants relating to the price elasticity of housing supply and the additional cost associated with constructing a new plot, and D_j is a capacity limit representing the maximum percentage increase in population that can be supported in neighbourhood j. In this example, D is set as a uniform capacity limit that constrains the development of new plots to a maximum of fifteen per cent in each neighbourhood[34]. Whilst I do not assume to be capturing all of the facets of housing supply, this extension provides an insight into the influence of developments in the property market and how these developments propagate and distort outcomes.

5.5 Rental revenue recycling

In the absence of more detailed information, I assume that fifty per cent of rental revenues are recycled to households. The share of rental revenues received by households was calibrated using Expenditure and Food Survey data on gross weekly household income and income from renting out. The corresponding shares are presented in Table 3.1. Data for 2000 reveals that the share of rental revenues being returned to households is increasing in income such that wealthier households are more likely to own a share of the stock of rental housing.

5.6 Government

34 The value of fifteen per cent was supported by the UK Housing Review (2009) which reports that over the period of 1991-2007 the number of owner-occupied new build properties constituted a rise of 15.3 per cent in the housing stock in England.

The government provides compensation to households located in a neighbourhood where road noise levels increase as a result of the construction of the bypass. Compensation is defined as follows,

$$compensation_{i,j,O} = \begin{cases} 0, & mwtp * \Delta g_{j,2} * h_{i,j}^{0} < 50 \\ mwtp * \Delta g_{j,2} * h_{i,j,O}^{0}, & mwtp * \Delta g_{j,2} * h_{i,j}^{0} \geq 50 \end{cases} \quad (3.17)$$

The level of compensation is calculated by defining a marginal willingness to pay value, $mwtp$, (derived from a hedonic regression using baseline property prices and environmental quality) that represents the additional premium a household pays on every unit of housing for a unit increase in environmental quality. The marginal willingness to pay value is multiplied by the size of the change in environmental quality, $\Delta g_{j,2}$, to obtain the premium paid for the total change and, lastly, multiplied by the quantity of housing units initially owned by the household, $h_{i,j,O}^{0}$, to calculate the total compensation. Following the 1973 Compensation Act the level of compensation is paid when this amounts to at least a minimum of £50. The model accounts for both the direct impact of compensation on utility and the interaction between the payment of compensation and the behavioural response of households.

Income Percentile	Share of Rental Revenues
5th	0
10th	0
15th	0.01
20th	0
25th	0
30th	0.01
35th	0
40th	0.01
45th	0
50th	0.01
55th	0.04
60th	0
65th	0.03
70th	0.02
75th	0.04
80th	0.05
85th	0.05
90th	0.25
95th	0.09
100th	0.40

Table 3.1: Share of rental revenues by income percentile

The calibrated model is coded in Matlab[35] and uses simulation and iterative numerical techniques to solve for market clearing prices and endogenous rental revenues (Lagarias et al., 1998).

6.0 Exploring the Model

Table 3.2 summarises the two neighbourhoods in the baseline. Notice that the road noise levels are higher in the town centre (59 dB) than in the suburbs (40 dB). Properties in the town centre are, on average 400m from the centre of the town, whereas properties in the suburbs are an average 1100m from the centre.

[35] The Matlab code is available from the authors upon request.

These two attributes combine to form an index of public good provision in each neighbourhood.

	Town Centre	**Suburbs**
Average distance from centre (m)	400	1100
Road noise level (dB)	59	40
Public goods index	33.0	33.9
Population share	0.5	0.5
Homeownership Rate	0.74	0.82

Table 3.2: Baseline characteristics of the neighbourhoods

The demographics of the two neighbourhoods derived for the calibrated baseline are presented in Table 3.3. A number of patterns of sorting can be seen. First, households with higher incomes tend to locate in the suburbs, driving up the price of properties in the neighbourhood with a greater provision of public goods. Second, poorer households, who spend a larger proportion of their income on housing, are attracted to the town centre by its lower property prices. Lastly, households with relatively high preferences for homeownership (theta) become owners and others become renters.

	Town Centre		Suburbs	
Price (£)	5183		5258	
Population Share	0.50		0.50	
Homeownership Rate	0.74		0.82	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	35,964	53,992	104,075	90,492
Mean $\boldsymbol{\beta}$	0.399	0.36	0.1	0.11
Mean $\boldsymbol{\theta}$	1.00	1.13	0.99	1.12
Mean housing	2.72	3.97	2.07	2.14
Median housing	0.83	1.29	0.41	0.35
Population	205	597	144	654

Table 3.3: Calibrated baseline neighbourhood composition

6.1 Examining Local Environmental Change

In that initial equilibrium, the economy experiences a change in conditions as a result of the construction of a bypass. The immediate impact of the bypass is to change the exposure to traffic noise of properties in the two neighbourhoods. Drawing on data from the real world experience in Polegate, road noise levels in the town centre fall by 2 dB to 57 dB as traffic flows are reduced. In the suburbs road noise rises by 1 dB to 41 dB due to noise from the bypass.

The conventional approach to assessing the welfare impact of constructing the bypass is to look at the direct impact on the households. In this case, the direct impact for households in the town centre is an improvement that equates to a 0.65 per cent rise in utility for all households[36]. For households in the suburbs the rise in road noise translates to a 0.33 per cent decrease in utility. Using the formula for calculating static or partial equilibrium (PE) willingness to pay,

$$WTP_{PE} = y_i^0 + \pi_i^0 - e(p_{j^0}^0, \boldsymbol{g_{j^0}^1}, y_i^0 + \pi_i^0, v^0) \tag{3.18}$$

Which by substitution becomes,

$$WTP_{PE} = \left[1 - \left(\frac{\boldsymbol{g_{j^0}^0}}{\boldsymbol{g_{j^0}^1}}\right)^{\frac{\alpha}{1-\alpha}}\right](y_i^0 + \pi_i^0) \tag{3.19}$$

Table 3.4 summarises the willingness to pay values for renters and owners in each of the two neighbourhoods and the sum of these values, which corresponds to the residents' total willingness to pay for bypass project. Results from the partial equilibrium calculation are presented in the top six rows. The partial analysis suggests that overall the bypass provides a welfare gain for residents with willingness to pay summing to £12,268. Notice that the partial equilibrium

[36] As a result of the Cobb Douglas utility function the percentage change in utility is equal across households, however this translates to different values of willingness to pay.

analysis suggests that utility gains flow predominantly to poorer households in the town centre.

Let us now compare this to the findings from the general equilibrium (GE) analysis of the same project. Results from that analysis are presented in the bottom half of Table 3.4. In the GE model willingness to pay is calculated as[37],

$$WTP_{GE} = y_i^0 + \pi_i^0 - e(p_{j^1}^1, \boldsymbol{g}_{\boldsymbol{j}^1}^{\boldsymbol{1}}, y_i^0 + \pi_i^1 + compensation_i, \nu^0) \tag{3.20}$$

Notice that in the GE calculation it is not only the environmental quality that changes. Households respond to the construction of the bypass by relocating and adjusting their consumption and tenure choice, this in turn has knock on effects on house prices and rental revenues.

Of course, this relies upon the assumption of free mobility and there are a number of additional frictions that we might also want to consider in applications: two important considerations are transaction and moving costs. These costs violate the assumption of free mobility and alter the nature of the choice set faced by households. In this chapter we abstract from these complexities however they have been addressed in the wider literature on equilibrium sorting models. For example Bayer, Keohane and Timmins (2009) show that moving costs related to a household's home town can have a large impact on estimates of willingness to pay, Bayer et al (2011) explore the psychological costs using information on the timing of moves and Ferreira (2010) examines the impact of transaction costs arising through the property tax

---

[37] Again, by substitution, this becomes,

$$WTP_{GE} = \left[1 - \left(\left(\frac{\boldsymbol{g}_{\boldsymbol{j}^0}^{\boldsymbol{0}}}{\boldsymbol{g}_{\boldsymbol{j}^0}^{\boldsymbol{1}}}\right)^{\alpha} \left(\frac{p_{j^1}^1}{p_{j^0}^0}\right)^{\beta} \left(\frac{\theta^0}{\theta^1}\right)\right)^{\frac{1}{1-\alpha}}\right](y_i^0 + \pi_i^0) + (\pi_i^1 - \pi_i^0) \tag{3.1}$$

Where $\theta^T = \begin{cases} \theta, & t^T = O \\ 1, & t^T = R \end{cases}$

regulations in California. Kuminoff (2009) explores the impact of moving costs on the substitution options available to households and demonstrates how failing to account for moving costs can introduce bias into estimates of willingness to pay since it alters the way in which revealed preferences are interpreted. Transaction and moving costs will also affect our prediction of the new equilibrium following a policy change as alter the incentives faced by households and serve to reduce the benefits from re-optimisation.

In the context of our model I anticipate that transaction and moving costs would lead the ESM to predict lower levels of relocation, in a way bringing the partial equilibrium and general equilibrium predictions closer together by preventing the relocation of households. However, prices would also have to adjust in response to moving costs. The overall impact would depend on the distribution of costs. For example, consider an increase in quality in a neighbourhood. In the absence of moving costs we would expect an increase in demand for housing in this neighbourhood from households with a relatively high preference for public goods, which would cause prices to rise. In the presence of moving costs the outcome would be different. If moving costs are very high no relocation takes place, if they are uniform but low then some but not all of the anticipated relocation will take place. In addition, transaction and moving costs could influence the distribution of welfare gains and losses if some households are more mobile than others, for example renters may have lower moving costs than owners, since they do not have to pay realtor's fees or sales taxes when they move, providing them with a greater ability to re-optimise. If moving costs are heterogeneous households with lower moving costs will be more likely to relocate.

	Town Centre		Suburbs	
	Renters	Owners	Renters	Owners
	Partial Equilibrium			
Mean WTP PE	260.60	391.23	-385.49	-335.19
Standard Deviation	702.25	686.99	954.79	985.77
Total WTP PE	12,268			
Average WTP PE	7.67			
Δ Rental revenues	0			
Δ Mortgage payments	0			
	General Equilibrium			
Mean WTP GE	71.76	444.10	300.15	-57.61
Standard Deviation	166.52	991.00	735.89	565.19
Total WTP GE	285,380			
Average WTP GE	178.36			
Δ Rental revenues	0			
Δ Mortgage	480			
Net Benefits	285,860			

Table 3.4: Willingness to pay for the bypass (£)

	Town Centre		Suburbs	
Price (£)	5305		5153	
Population Share	0.65		0.35	
Homeownership Rate	0.85		0.64	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	99,840	82,856	35,313	50,639
Mean $\boldsymbol{\beta}$	0.11	0.25	0.39	0.17
Mean $\boldsymbol{\theta}$	0.99	1.12	1.00	1.13
Mean housing	2.02	2.98	2.71	3.14
Median housing	0.44	0.90	0.82	0.47
	Population Movements			
From Town Renters	9	0	196	0
From Town Owners	0	567	0	30
From Suburb Renters	144	0	0	0
From Suburb Owners	0	313	8	333
Population	153	880	204	363

Table 3.5: Neighbourhood composition after the bypass

To understand more clearly how the difference in the GE WTP results arise, consider Table 3.5 which describes the ESM's prediction of the new equilibrium which arises in the community after the impacts of the new bypass have worked their way through the economy. The first three rows of the table display the new prices, population shares and homeownership rates for the two neighbourhoods, the next five rows presents the new population characteristics and the last five rows detail the population movements and new population totals.

After the construction of the bypass, the town centre supports a greater provision of public goods, this attracts wealthier households. Indeed, as can be seen from the population movement figures, the renters and almost half of the owners from the suburbs migrate to the town centre causing the population share to rise to 65 per cent. The shift in demand causes prices to rise in the town centre and fall in the suburbs. As a result, some lower income households relocate from the town centre to the suburbs as property prices in the town rise. Consequently, the average income of households in the town centre rises, both amongst renters and owners, while the average income of households in the suburbs declines. This process has been described as "environmental gentrification" by Sieg et al (2004) and has been demonstrated in previous applications of ESMs to the study of improvements in air quality in the LA basin (Sieg et al 2004), Toxic Release Inventory (TRI) emissions (Banzhaf and Walsh, 2004) and the provision of open space (Walsh, 2003).

Rising property prices in the town centre lead to quite different welfare effects being experienced by renters and owners originally resident in that neighbourhood. While both initially enjoy the benefits of reduced road noise, those gains are offset for renters by higher rental prices. Indeed, the price differential between the town centre and the suburbs is such that the town centre renters all choose to forgo the improved environmental quality and move to the suburbs. Examining Table 3.4, the ESM analysis shows that the welfare gains for town centre renters are substantially lower than those suggested by the partial analysis, which ignores the price changes. Renters face higher rental costs that offset the environmental gains from lower road noise. As can be seen by comparing the partial and general equilibrium results in Table 3.4, this causes the GE WTP values for renters initially located in the town centre to be lower than their PE equivalents. For owners, on the other hand, price rises present an opportunity to sell up and realise the benefits of capital gains on housing. Accordingly, in the GE analysis, the welfare gains for town centre owners are substantially higher than those suggested by the PE analysis. The combination of these effects leads to a striking difference between the total willingness to pay values calculated in the PE analysis and the GE analysis, with the latter being an order of magnitude greater than the former.

For households in the suburbs, the decrease in utility predicted in the PE analysis is dissipated through a number of mechanisms. First, the ability to relocate allows some households to migrate to the town centre and benefit from a greater provision of public goods. Second, some owners benefit from lower property prices that allow them to expand their housing consumption. Overall, the price effect transforms the welfare impact for renters in the suburbs by more than compensating them for the increase in road noise levels. For owners in the suburbs, the welfare loss can only be partially offset so that, on average, the GE analysis concludes that these households experience a welfare loss but smaller than the one predicted by the partial analysis.

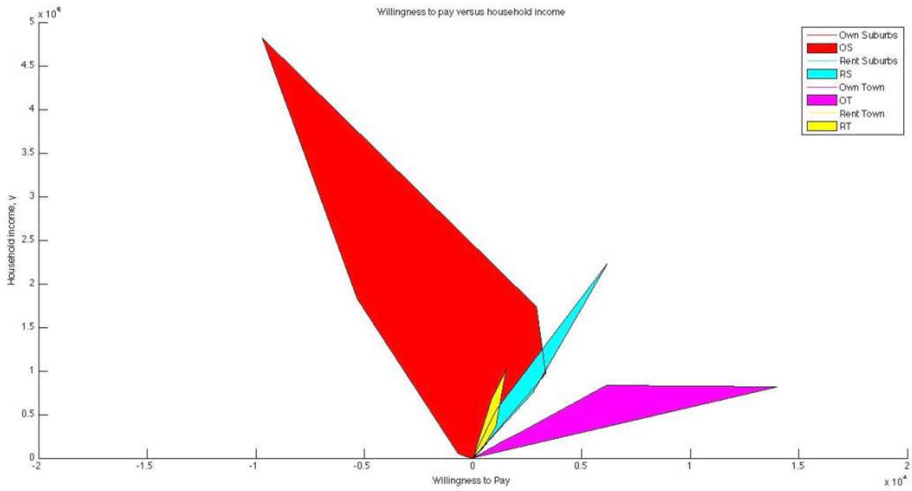

Figure 3.1: The convex hull of willingness to pay values for households grouped by initial residential choice. (Endogenous tenure choice ESM)

Figure 3.1 plots the distribution of GE willingness to pay values by household income. The figure groups together households by their initial residential choices and plots the convex hull of their willingness to pay values. Since the environmental change experienced in each neighbourhood is different, the separation between households in the suburbs and town is anticipated in the PE analysis. However, Figure 3.1 also clearly captures the striking distinction between the willingness to pay of renters and owners, even within a neighbourhood. In particular, at a given income, a renter in the town will benefit less than an equivalent owner, whereas in the suburbs a renter is likely to benefit equally or greater than an equivalent owner in the suburbs.

This simplified ESM demonstrates clearly how the welfare conclusions drawn by a partial equilibrium approach contrast to a general equilibrium approach and the mechanisms through which the two differ. It is clear that the distribution of gains and losses is altered once households are permitted to adjust their behaviour in response to the exogenous policy change. In particular, a large degree of environmental gentrification is observed resulting in the environmental improvements being enjoyed by wealthier households who displace poorer households who are originally resident in the town. Furthermore, it is clear that medium to long term adjustments lead to striking divisions between the welfare effects experienced by renters and owners that are being overlooked in the partial analysis. As a result, the PE analysis provides a misleading picture of welfare changes.

7.0 Model Extensions

In reality of course there are more complexities that need to be included in the analysis. Let us introduce, in turn, three extensions to the model: rental revenue recycling, housing supply constraints and compensation.

7.1 Rental Revenues

In the model considered above rental revenues flow out of the economy to absentee landlords. In the real world, households may also own rental housing stock. In this case, changes in rental revenues also have consequences for household welfare. As described in Section 5.5, the ESM can be extended to allow households in the economy to also be landlords such that some or potentially all of the revenues from rental payments are recycled within the community.

Table 3.5 presents information on the population characteristics for the two neighbourhoods in the calibrated baseline when fifty per cent[38] of the rental

[38] At present there is little information on the proportion of rental revenues that are recycled rather than flowing to government and private organisations. I adopt the

revenues are returned to households. In the baseline total rental revenues amount to £4,552,600. As is clear from a comparison of Tables 3.3 and 3.6, the recycling of these revenues has no impact on the sorting of households between the two neighbourhoods, which is expected since it does not alter the relative desirability of the two areas[39]. Instead, since the amount of money being returned to any individual household is relatively small the response occurs through increases in housing units and other consumption.

	Town Centre		Suburbs	
Price (£)	5183		5258	
Population Share	0.50		0.50	
Homeownership Rate	0.74		0.82	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	35,964	53,992	104,075	90,492
Mean $\boldsymbol{\beta}$	0.40	0.36	0.10	0.11
Mean $\boldsymbol{\theta}$	1.00	1.13	0.99	1.12
Mean housing	2.79	4.07	2.10	2.17
Median housing	0.83	1.31	0.41	0.36
Population	205	597	144	654

Table 3.6: Calibrated baseline with rental revenue recycling

Now let us consider the impact of the construction of the bypass when rental revenues are recycled. The new prices and population characteristics are presented in Table 3.6. The key mechanisms are the same as those discussed above however, the rise in revenues (from £4,552,600 to £4,578,300) leads to an

---

assumption of fifty per cent as a largely uninformed estimate. The successful integration of general equilibrium sorting models would benefit from more informed estimates of the actual share of revenues being recycled.

[39] If the model included minimum housing sizes or other discrete constraints this would likely alter the results such that the distribution of households became sensitive to the recycling of rental revenues.

increase in demand for housing units and other consumption. To clear the market, house prices in both neighbourhoods rise and end up higher than in the analysis when rental revenues are not recycled (Table 3.4).

The corresponding PE and GE welfare impacts are detailed in Table 3.7. Notice that the increase in rental revenues provides additional utility gains for households initially located in the town centre and renters in the suburbs, and partly offsets the losses experienced by owners in the suburbs. As a result, the GE total willingness to pay value rises to £304,160. The distribution of welfare gains and losses between neighbourhoods remains largely the same although the rental revenue mechanism provides additional gains to those at the higher end of the income distribution. Notice by comparing Tables 3.3 and 3.7 that accounting for rental revenue recycling leads to an increase in the standard deviation of the willingness to pay values within each group.

	Town Centre		Suburbs	
Price (£)	5307		5155	
Population Share	0.65		0.35	
Homeownership Rate	0.85		0.64	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	99,840	83,057	35,313	50,061
Mean $\boldsymbol{\beta}$	0.11	0.25	0.39	0.17
Mean $\boldsymbol{\theta}$	0.99	1.12	1.00	1.13
Mean housing	2.05	3.06	2.77	3.18
Median housing	0.44	0.91	0.82	0.47
	Population Movements			
From Town Renters	9	0	196	0
From Town Owners	0	568	0	29
From Suburb Renters	144	0	0	0
From Suburb Owners	0	313	8	333
Population	153	881	204	362

Table 3.7: Neighbourhood composition after the bypass with revenue recycling

	Town Centre		Suburbs	
	Renters	Owners	Renters	Owners
	Partial Equilibrium			
Mean WTP PE	266.57	400.92	-391.72	-341.35
Standard Deviation	714.38	702.85	963.51	993.06
Total WTP PE	14,349			
Average WTP PE	8.97			
Δ Rental revenues	0			
Δ Mortgage payments	0			
	General Equilibrium			
Mean WTP GE	74.53	462.42	311.97	-49.10
Standard Deviation	171.80	1,022.60	750.10	569.50
Total WTP GE	304,160			
Average WTP GE (	190.10			
Δ Rental revenues	300			
Δ Mortgage payments	500			
Net Benefits	304,960			

Table 3.8: Willingness to pay with rental revenue recycling (£)

7.2 Housing supply constraints

The ESMs discussed so far have adopted the conventional approach to housing supply, which is to assume that the number of housing blocks is fixed at the baseline level. In the results that follow the linear housing supply function with capacity limits described in Section 5.4 is adopted to reflect more realistically the medium term adjustments that might occur in housing supply.

Since the housing supply function has no impact on the calibration the baseline for this model is as before (Table 3.6). The prices and population characteristics in the new equilibrium following the introduction of the bypass are presented in the first eight rows of Table 3.9.

The last five rows of Table 3.8 summarise the population movements that took place between the baseline and the new equilibrium. The capacity constraint, which in this case limits the population share to 0.58 in the town, is binding and inflates prices in the town centre as wealthier households relocate from the suburbs to the town centre. As capacity is reached in the town the rising price leads to the displacement of households with relatively high β values from the town to the suburbs: this can be observed through comparing the average β values in each neighbourhood in Table 3.9 to those in Table 3.7. As can be seen in the population movements section of Table 3.9, a greater number of households are displaced in this model than when capacity limits are not accounted for. Consequently, there is an increase in demand for housing units in the suburbs, which pushes up property prices in the neighbourhood despite the small fall in public good provision.

These price effects have important and substantial implications for welfare changes, as summarised in Table 3.10. Price rises in both neighbourhoods lead to welfare losses for renters initially located in both neighbourhoods. In contrast, the price rises confer gains to owners in the town centre and provide a source of compensation for owners in the suburbs. Consequently, the distribution of welfare gains and losses is transformed once the housing supply adjustments are incorporated. Total willingness to pay falls to £159,210. Moreover, the distribution of gains and losses shifts with large losses accruing to lower income renters and large gains to wealthy owners. Figure 3.4 plots the distribution of willingness to pay and household income for each of the initial residential choice groups. Comparing Figure 3.2 to Figure 3.1, it is apparent that the introduction of a medium to long term housing supply specification leads to further adjustments in the property market which redistribute the benefits of environmental changes towards homeowners and away from renters. Moreover, the capacity constraint acts to constrain the increase in the number of households residing in the centre to 120 households in comparison to 232 households in the absence of the capacity constraint. As can be seen in Figure 3.2, these adjustments actually lead to losses for the majority of renters from both the suburbs and the town, despite the direct improvement in environmental quality.

	Town Centre		Suburbs	
Price (£)	5514		5314	
Population Share	0.58		0.42	
Homeownership Rate	0.86		0.67	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	111,352	69,927	37,350	78,579
Mean $\boldsymbol{\beta}$	0.09	0.25	0.37	0.18
Mean $\boldsymbol{\theta}$	0.99	1.12	1.00	1.12
Mean housing	1.98	2.14	2.68	4.60
Median housing	0.38	0.73	0.83	0.64
	Population Movements			
From Town Renters	0	0	205	0
From Town Owners	0	521	0	76
From Suburb Renters	126	0	18	0
From Suburb Owners	0	275	0	379
Population	126	796	223	455

Table 3.9: Neighbourhood composition after the bypass with revenue recycling and linear housing supply

	Town Centre		Suburbs	
	Renters	Owners	Renters	Owners
	Partial Equilibrium			
Mean WTP PE	266.57	400.92	-391.72	-341.35
Standard Deviation	714.38	702.85	963.51	993.06
Total WTP PE	14,349			
Average WTP PE	8.97			
Δ Rental revenues	0			
Δ Mortgage payments	0			
	General Equilibrium			
Mean WTP GE	-374.53	557.42	-117.18	-122.20
Standard Deviation	883.20	1,359.30	623.80	850.20
Total WTP GE	159,210			
Average WTP GE	99.51			
Δ Rental revenues	0			
Δ Mortgage payments	500			
Net Benefits	159,710			

Table 3.10: Willingness to pay with rental revenue recycling and linear housing (£)

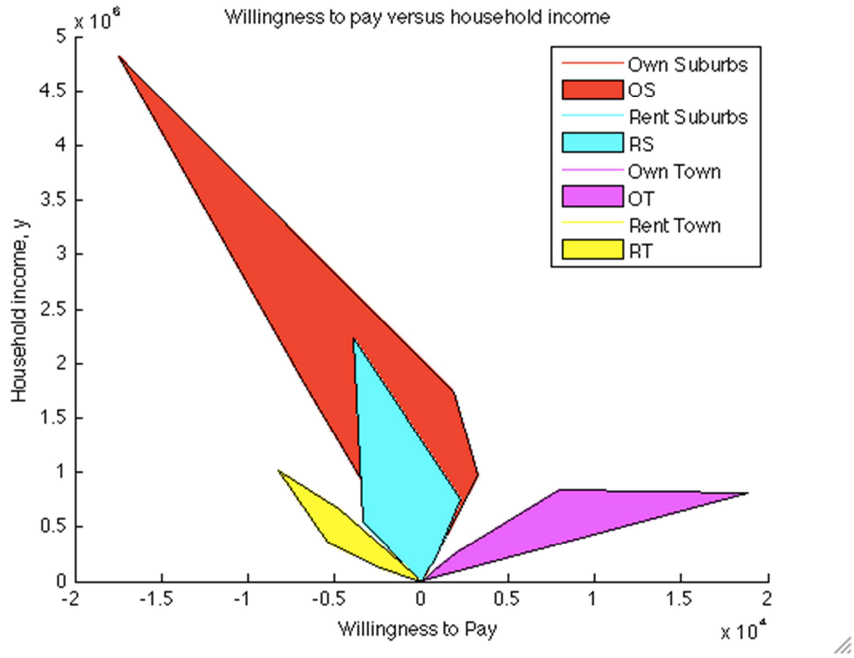

Figure 3.2: The convex hull of willingness to pay values for households grouped by initial residential choice.
(Endogenous tenure choice ESM: Housing Supply)

Comparing these results to the partial equilibrium approach it becomes clear that the partial analysis does not adequately capture the importance of initial tenure status and provides an incomplete picture of the welfare implications of a policy change.

7.3 Compensation

The final extension considered is that of compensating homeowners for loss of value in their properties. For this purpose we incorporate a compensation mechanism into the ESM, which is modelled on Part 1 of the 1973 Compensation Act.

As previously, the relevant baseline is given in Table 3.5. Table 3.10 presents the neighbourhood prices and population characteristics following the construction of the bypass. The largest compensation payment is £29,101. The receipt of compensation has little impact on the behaviour of households: as with the receipt of rental revenues, households receiving compensation use it to increase their housing and other consumption. This leads to a small increase in the average number of housing units consumed by owners initially located in the suburbs.

	Town Centre		Suburbs	
Price (£)	5514		5315	
Population Share	0.58		0.42	
Homeownership Rate	0.86		0.67	
	Population Characteristics			
	Renters	Owners	Renters	Owners
Mean Income (£)	111,352	69,927	37,350	78,579
Mean $\boldsymbol{\beta}$	0.09	0.25	0.37	0.18
Mean $\boldsymbol{\theta}$	0.99	1.12	1.00	1.12
Mean housing	1.98	2.14	2.68	4.61
Median housing	0.38	0.73	0.83	0.64
	Population Movements			
From Town Renters	0	0	205	0
From Town Owners	0	521	0	76
From Suburb Renters	126	0	18	0
From Suburb Owners	0	275	0	379
Population	126	796	223	455

Table 3.11: Neighbourhood composition after the bypass with revenue recycling, linear housing supply and compensation

	Town Centre		Suburbs	
	Renters	Owners	Renters	Owners
	Partial Equilibrium			
Mean WTP PE	266.57	400.92	-391.72	-341.35
Standard Deviation	714.38	702.85	963.51	993.06
Total WTP PE	14,349			
Average WTP PE	8.97			
Δ Rental revenues	0			
Δ Mortgage payments	0			
	General Equilibrium			
Mean WTP GE	-365.80	571.59	-99.38	221.21
Standard Deviation	864.50	1,380.00	613.30	758.70
Total WTP GE	396,610			
Average WTP GE	247.88			
Δ Rental revenues	0			
Δ Mortgage payments	570			
Compensation	221,770			
Net Benefits	175,410			

Table 3.12: Willingness to pay with rental revenue recycling, linear housing and compensation (£)

The influence of compensation is most clearly visible in Table 3.12 and Figure 3.3, which summarise the GE willingness to pay values. Driven by the increased revenues entering the system from government compensation payments, total GE willingness to pay rises to £396,610. Comparing with Table 3.10, the results illustrate that, on average, the payment more than compensates owners initially located in the suburbs. Notice that compensation is paid only to owners initially located in a neighbourhood that experiences environmental depreciation. Accordingly, renters initially located in both the town and the suburbs continue to experience a reduction in utility. Comparing Figure 3.3 to Figure 3.2, the payment of compensation clearly alters the distribution of gains and losses, broadening the division between renters and owners. In both neighbourhoods, the corresponding GE willingness to pay values are negative for the majority of renters, almost uniformly for renters initially in the town centre, and are almost entirely positive for owners.

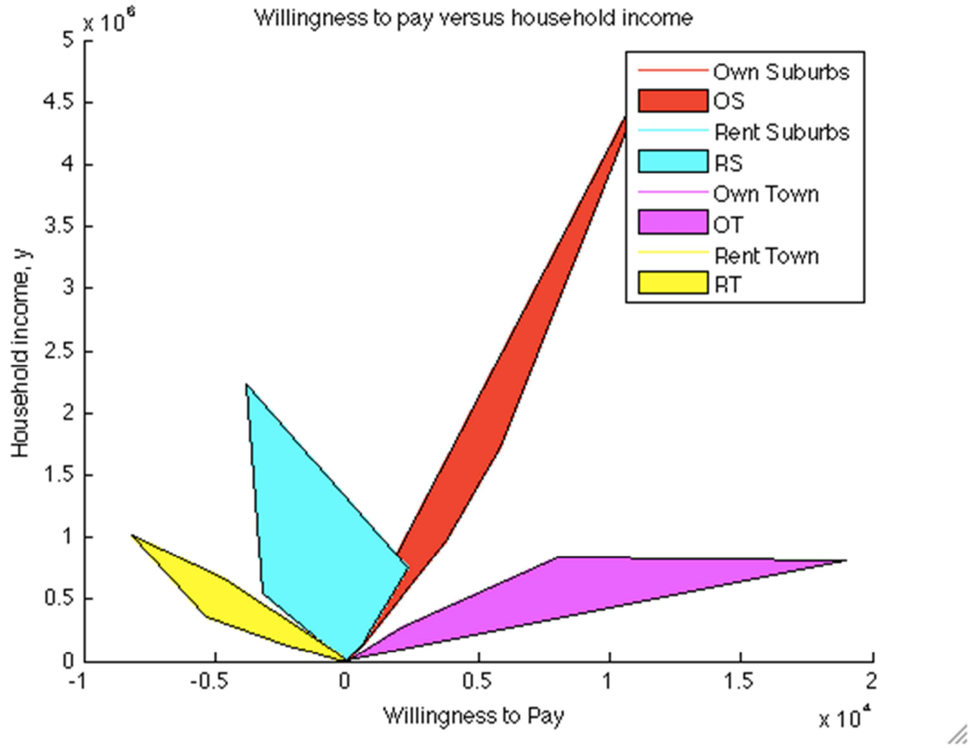

Figure 3.3: The convex hull of willingness to pay values for households grouped by initial residential choice.
(Endogenous tenure choice ESM: Compensation)

8.0 Discussion

A key objective of the work presented above has been to explore the magnitude and distribution of welfare changes precipitated by policies that result in localised environmental change. In terms of magnitude, the results demonstrate a large difference between the welfare changes predicted by a partial equilibrium analysis in comparison to a general equilibrium one. For this particular policy simulation the general equilibrium total willingness to pay value is an order of magnitude greater than the in the partial analysis suggesting that adjustments in the property market serve to amplify the direct benefits of the policy.

To examine the distributional impact of a policy it is useful to develop ways to characterise the distribution of welfare and welfare changes. One approach is to take inspiration from the Lorenz curve (which presents a graphical representation of the Gini-coefficient) and consider the cumulative share of total utility that is enjoyed by a particular share of the total population. Plotting these two measures against each other results in a utility equivalent to the Lorenz curve where the 45-degree line represents the situation where all households in the economy enjoy an equal level of utility. Figure 6 plots this utility-based Lorenz curve for the calibrated baseline with rental revenue recycling. The deviation of the curve from the 45-degree line represents inequality in the distribution of utility across households. More specifically, the line lies below the 45-degree line, which reflects the fact that poorer households receive a less than equal share of utility. The gradient of the Lorenz curve does not reach tangency with the 45-degree line until the horizontal axis reaches roughly eighty per cent of the population. This indicates that the top twenty per cent of the population receive a greater than even share of utility.

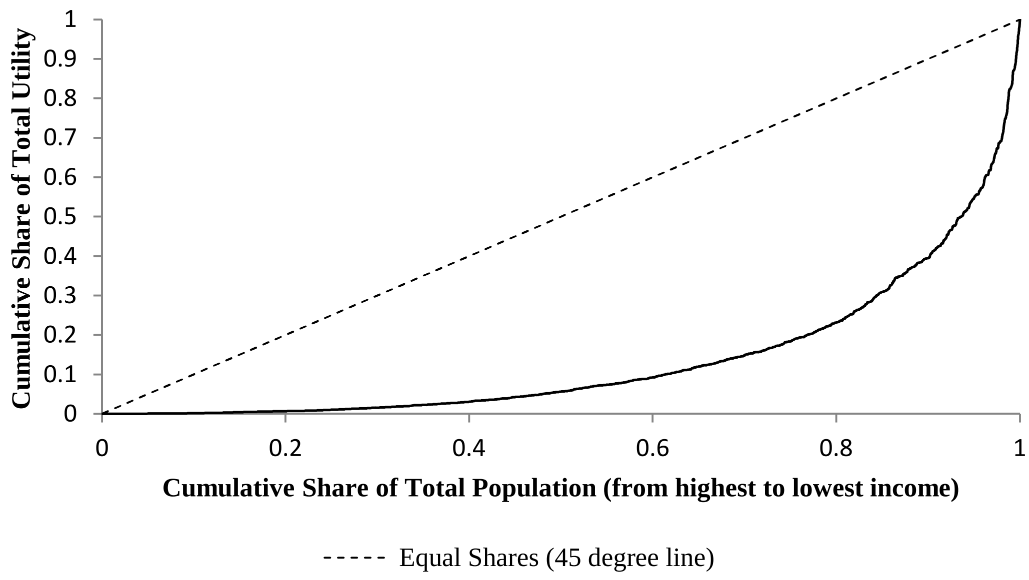

Figure 3.4: The Utility based Lorenz Curve plotting the cumulative share of total utility (in the baseline with rental revenue recycling) against the cumulative share of the total population ranked by lowest to highest income.

Whilst the distribution of welfare described in Figure 3.3 is driven primarily by the assumed distribution of income in the population, our central interest is to understand how policies that affect local environmental quality change this distribution. To provide a perspective on the distributional impact it is more useful to consider changes in the cumulative share of total utility that occur as a result of the policy change.

Figure 3.4 plots the corresponding changes in the cumulative share predicted by the PE and GE analyses against log income. The figure is not intuitive at first glance, the figure in panel 7a plots the PE and GE curves together. Notice that the PE analysis suggests that the policy has large distributional impacts of a sizeable magnitude, whereas the GE analysis suggests that the impacts are much smaller (the plot of these changes is barely distinguishable from the horizontal axis). At log income values where the curves are positive the cumulative share of total utility held by households whose logged incomes are less than or equal to this amount is higher than in the baseline, meaning that households at the lower end of the income distribution have increased their share of utility. When the curve is negative this implies the opposite: that the share of utility has fallen. At any given log-income value, a positive gradient indicates that the share of utility held by that household has increased whereas a negative gradient indicates that their share has fallen.

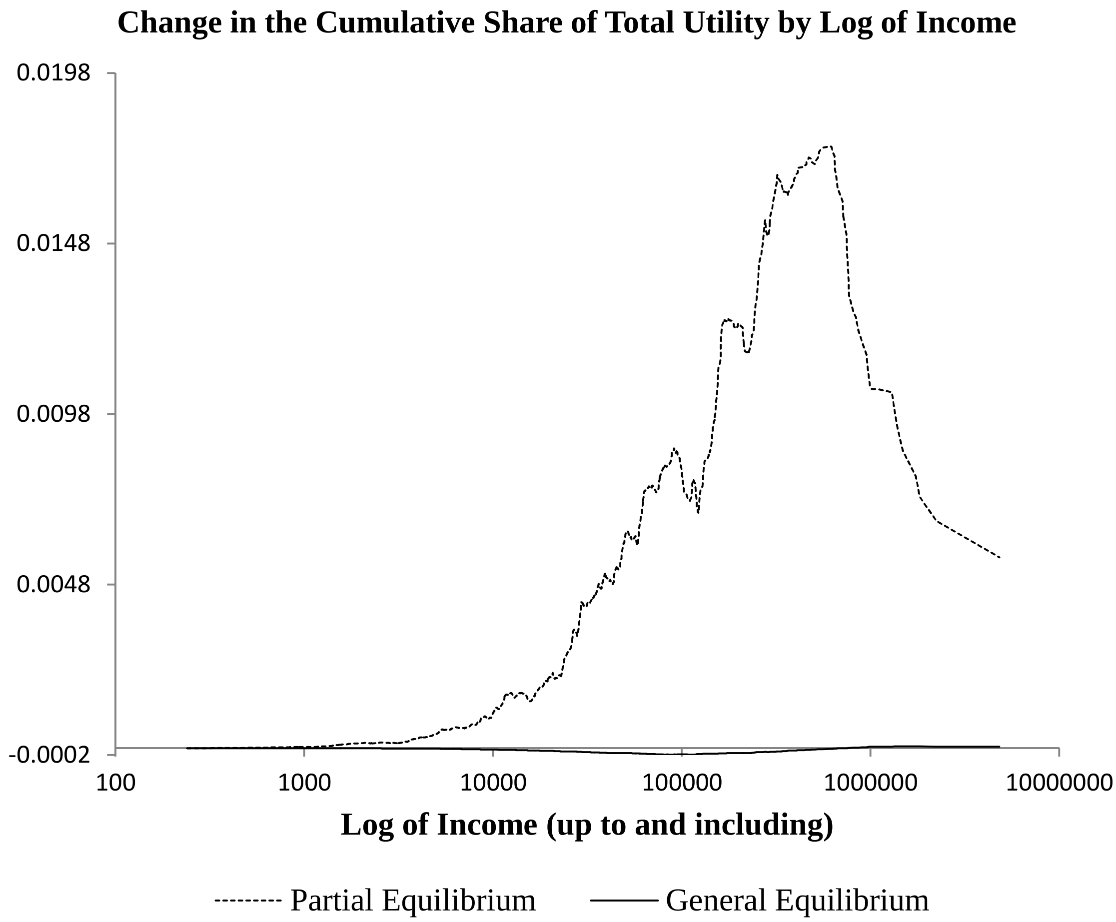

Figure 3.5a

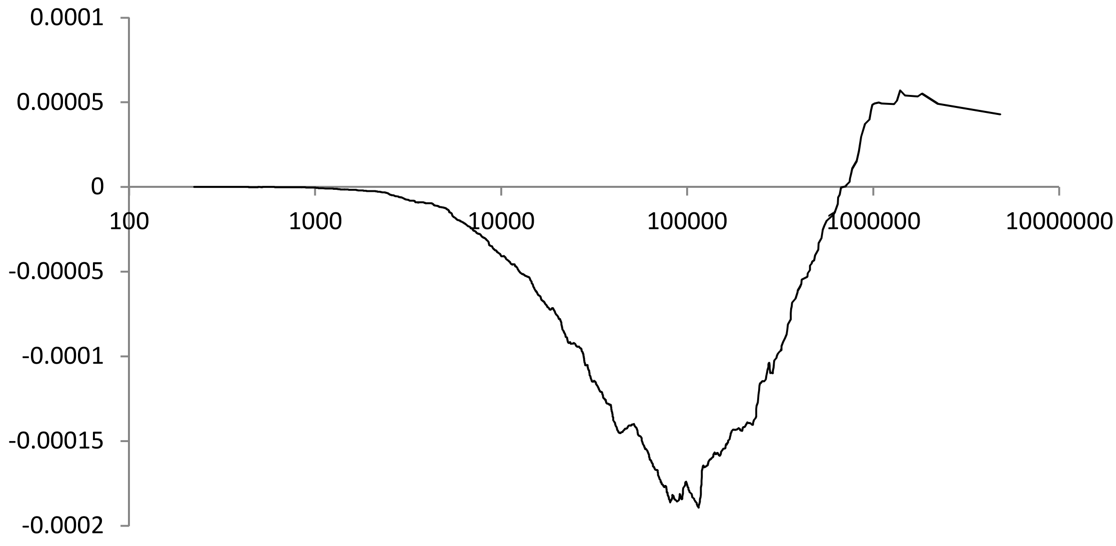

Figure 3.5b

Figure 3.5: The distribution of changes in the cumulative share of total utility by income in the Partial and General Equilibrium analyses. The bottom panel displays the general equilibrium plot at a finer scale

Turning attention to the PE (dashed) curve in Figure 3.5, the curve is positive and increasing (or at least predominantly increasing) up to around 600,000. This indicates that the PE analysis suggests the policy change is progressive and, through targeting environmental improvements at low-income households, alters the distribution of utility making it more equitable. In contrast, panel 7b shows a rescaled version of the GE curve. After re-scaling a clear distributional impact is evident and it is remarkably different to the one that is predicted by the PE analysis. The GE analysis suggests that once adjustments in location, property prices, rental revenues and compensation have been accounted for the direct environmental benefits are channelled away from poorer households. In fact, households at the lower end of the income distribution are left with a smaller share of the total utility. The benefits of the policy therefore are channelled to households with logged incomes between 100,000 and 1,250,000 who now receive a greater share of total utility. Overall then the GE analysis suggests that despite seemingly progressive targeting of the bypass, the overall welfare impacts are moderately regressive.

9.0 Concluding Remarks and Future Directions

In this paper I have sought to explore the magnitude and distribution of the welfare effects resulting from policies with localised environmental impacts through the development and calibration of an ESM with endogenous tenure choice. The mechanisms and feedbacks at work in this model have been illustrated through the presentation of a stylised two-neighbourhood policy analysis. The results from this example also enable us to draw some more general conclusions regarding the magnitude and distribution of welfare changes. In particular, the results clearly demonstrate that once relocation, capital gains, rental revenues, housing supply and compensation are taken into account the GE welfare changes predicted by the ESM deviate substantially both in magnitude and distribution from those anticipated using a conventional PE approach.

The modelling exercise highlights two key conclusions that may be drawn about welfare changes. First, considering the magnitude of the welfare impact, the GE analysis produces a substantially larger - by one order of magnitude - total willingness to pay value than the PE analysis, indicating that the general equilibrium adjustments that occur in response to the policy change, propagate and add to the initial environmental change and, ultimately, convey larger overall welfare gains to households. Second, with regards to the distribution of these welfare changes, the PE analysis suggests that targeting environmental policies directly towards poorer households could increase their welfare and act as a tool for redistribution. The GE analysis contradicts this assertion and suggests that the property market acts to channel these targeted benefits away from renters towards owners and away from lower income households towards wealthier ones. As a consequence of this mechanism, environmental policy is at best a convoluted tool for redistribution and at worst entirely misguided. Indeed, in the example presented here the overall amount of redistribution is small and, despite the policy being targeted, the redistribution that does take place is regressive in nature.

The GE analysis also has implications for the effectiveness and equity implications of initiatives like the 1973 Land Compensation Act. The motivation behind the payment of compensation is to reimburse households who are subjected to welfare losses as a result of the policy. Our results show that the payment of compensation to homeowners (and not renters) serves to broaden the inequality between the welfare effects experienced by homeowners and those experienced by renters. Moreover, the magnitude of compensation far exceeds the amount required to make these homeowners as well off as they were before the policy change. By neglecting to account for the medium term adjustments that take place, the compensation mechanism thus becomes a tool for redistributing welfare to homeowners.

In conclusion, policies that affect local environmental quality have complex distributional impacts. Although relatively young in terms of their theoretical development and range of applications, ESMs provide a framework which can provide insights into both the magnitude and distribution of the welfare effects that result from a policy change. The analysis presented in this paper has focused upon introducing an endogenous tenure choice in the ESM framework. However, there are a several other theoretical and empirical considerations that present continuing challenges. First, there are a number of additional complexities that have not been addressed in this paper, such as the role of moving costs (Kuminoff 2009), duality between residential and labour market decisions and a more complete development of housing supply with a buy-to-let property market. A comprehensive analysis would need to bring these complexities together into a single model. Second, from an empirical perspective, further work on estimating or calibrating the baseline model is needed. To achieve this, improvements are needed in terms of the availability of information relating to mortgage payments, ownership of the stock of rental housing and the nature of medium to long term housing supply. This is an active and rewarding area of research with many on-going developments and avenues for future work, which will have profound impacts on the way in which welfare analysis is conducted.

CHAPTER 4

Exploring Mortgage Interest Deduction Reforms: An equilibrium sorting model with endogenous tenure choice

1.0 Introduction

"The benefits of homeownership for families, communities and the nation are profound." - Elizabeth Dole

The promotion of homeownership has been a widespread and long-term focus of public policy (Andrews and Sánchez, 2011). Support for such policies derives both from political ideology and from a belief that homeownership delivers positive spillovers. Homeowners, it is argued, have greater incentives to invest in the physical and social capital of their communities, thus providing private and public benefits. There is a substantial body of empirical evidence that lends credence to this view. Homeownership is strongly correlated with property condition and maintenance (Galster, 1983), neighborhood stability (Rohe, 1996, Dietz and Haurin, 2003), child attainment (Green and White, 1997, Haurin et al., 2002, Bramley and Karley, 2007), citizenship (DiPasquale and Kahn, 1999) and lower crime rates (Glaeser and Sacerdote, 1996, Sacerdote, 1996, Sampson and Raudenbush, 1997)[40].

A wide variety of policy measures have been implemented to promote homeownership. Attempts have been made to encourage the supply of mortgage lending: for example, in the U.S. through the establishment of Government Sponsored Entities providing liquidity and security for mortgage lenders. Policies have also been implemented to encourage particular groups into homeownership: for example, in the U.K. through the Right to Buy scheme for social housing tenants. Homeownership has also been promoted through the tax system e.g. through exemptions from capital gains tax on property sales and mortgage interest deduction (MID). MID, the focus of this paper, allows taxpayers to subtract interest paid on a residential mortgage from their taxable income.

MID is present in the tax laws of many countries including the U.S., Belgium, Ireland, the Netherlands, Switzerland and Sweden and was previously offered in the U.K. and

[40] Of course, correlation is not causality. Doubts remain as to whether there is a direct causal link between homeownership and the observed positive spillovers or whether households who choose to own their homes are also more inclined to pro-social behavior.

Canada. It was introduced in the U.S. in 1913 when the homeownership rate was 45.9 percent. Under MID and numerous other initiatives, homeownership rose after the Second World War reaching a peak of 69 percent in 2004[41]. Currently, MID constitutes the second largest US tax expenditure[42] with the cost estimated to be some $104.5 billion dollars in foregone tax revenue in 2011 (Office of Management and Budget, 2011).

In the context of a large US fiscal deficit, MID has come under increased scrutiny. It has been argued that rather than encouraging homeownership the tax subsidy is simply capitalized into property values making properties no more and potentially less affordable than without the policy (Shapiro and Glaeser, 2003, Hilber and Turner, 2010). Furthermore, critics contend that MID most greatly benefits high-income taxpayers who would likely be homeowners irrespective of the tax incentives (Shapiro and Glaeser, 2003). Certainly higher income households are more likely to own their homes, hold larger mortgages and itemize mortgage interest payments on their tax returns (Poterba and Sinai, 2008). Of course, courtesy of their higher incomes, they also itemize at a higher rate (Glaeser and Shapiro, 2002). As a result, in 2004 the government paid an average $5,459 in MIDs to households earning over $250,000 compared to $91 for households earning below $40,000 (Poterba and Sinai, 2008).

In the face of strong opposition, particularly on the part of financial services interests and housing lobbyists, repeated efforts to reform MID in the U.S. have borne little fruit (Ventry Jr, 2010)[43]. Over the last three budget cycles the U.S. administration proposed reforms to MID, but on each occasion those initiatives have failed to pass into law. The key element of those proposals was to limit MID for households paying the top marginal rates of income tax. Other proposals for reform include: replacing MID with a system of tax credits (Follain et al., 1993, Dreier, 1997, Green and

[41] Source: U.S. Census Bureau.

[42] The largest being the exclusion of employer contributions for medical insurance and medical care.

[43] In March 2011 Moe Veissi, the president elect of the NRA, launched a call for action to Preserve, Protect and Defend the Mortgage Interest Deduction. http://www.realtor.org/government affairs/mortgage interest deduction.

Vandell, 1999), scrapping MID in order to fund cuts in federal income taxes (Stansel, 2011) and replacing MID with a fiscal incentive open only to first time buyers (Gale and Gruber, 2007).

The debate is fuelled by a lack of clarity with regards to how such reforms will play out. Clearly, eliminating the MID will increase the cost of borrowing for the purposes of buying property and, *ceteris paribus*, cause demand for owned properties to fall. This reasoning underpins the National Association of Realtors claim that "eliminating the MID will lower the homeownership rate in the U.S"[44]. Of course, it is recognized that the impact of eliminating the MID also depends on supply conditions in the property market. The extent to which falling demand translates into reductions in homeownership as opposed to falling prices depends on the price elasticity of housing supply. Bourassa and Yin (2008) estimate that for some groups the negative effect of losing MID may be more than outweighed by the positive effect of falling property prices. Indeed, homeownership amongst such groups could actually rise as a result of eliminating the MID.

What is less widely recognized is that changing market conditions in the property market will have ramifications in the closely associated rental market. Falling demand for homeownership can translate into rising demand for rental housing, increasing rental prices. More complex still is the interplay between homeownership and the desirability of residential locations. When homeownership contributes to the provision of various local public goods, reductions in homeownership in a neighborhood may reduce their desirability as a residential location. Since residential location choice is endogenous to the problem, eliminating MID might not only provoke the movement of individuals between ownership and rental but also the migration of households between neighborhoods.

While numerous attempts have been made to identify the impacts of eliminating the MID (Bourassa and Yin, 2008, Hilber and Turner, 2010, Toder, 2010) those studies

---

[44] Statement by NAR Chief Economist Lawrence Yun at the "Rethinking the Mortgage Interest Deduction" forum, Tax Policy Center, Washington, July 29, 2011.

have been based on a partial characterization of the problem. This paper develops a model that more completely describes the complex adjustments in spatially defined and interrelated property and rental markets and uses that model to explore some of the possible ramifications of MID reform.

The model developed in this paper is an equilibrium sorting model (ESM) (Kuminoff et al., 2010). ESMs provide a framework within which it is possible to examine how households choose their residential location from a set of discrete neighborhoods. As reviewed in Section 2, ESMs have been developed to examine a number of economic issues relating to choice of residential location. As far as I am aware, however, this model is the first to simultaneously model purchase and rental markets while endogenising tenure choice. In Section 3 the innovations of the model, particularly the specification of a neighborhood level of public good provision whose value depends in part on endogenous levels of homeownership and the development of an adjustment process to policy reform that accommodates capital gains, are outlined in detail.

To elucidate the pathways of adjustment that MID reform may initiate in property markets, Section 4 presents a simple two-jurisdiction calibration of the model based on 2000 census data for Boston, Massachusetts.[45] The calibrated model is used to

---

[45] In this chapter I have simplified the problem down to a model with two jurisdictions. This simplification serves to provide a clear illustration of the adjustment mechanisms in the model. However, in applications the choice of the number of jurisdictions is important as it defines the choice set faced by households. The appropriate number of neighbourhoods is defined by the nature of the problem being studied. In particular, each jurisdictions should provide a homogeneous level of public good provision; as a result jurisdictions are likely to be defined by geographical features such as valleys – which may influence the exposure of an area to pollutants, local authority boundaries – such as school authority boundaries – and the spatial extent of spatial interactions such as peer effects. There is no theoretical limit to the number of feasible jurisdictions, although in practice this is limited by the computing power available to the analyst and the resolution of the available data.

If the jurisdictions are incorrectly defined this is likely to cause bias in the willingness to pay values derived from the model Kuminoff (2009). This can arise through two channels: first simulate four different MID reform proposals: capping MID at a rate of 28%, replacing MID with refundable tax credits, scrapping MID and reducing income taxes and replacing MID with a lump sum payment to new owners. The simulations allow us to examine several important questions with regards to MID reform. In particular, to explore how reforms may impact purchase and rental prices, levels of homeownership, the distribution of welfare across income groups and the mixing of income groups within and across jurisdictions.

This analysis suggests that, contrary to existing claims, with the right policy design it may be possible to reform MID whilst maintaining the prevailing rates of homeownership, increasing public goods provision and contributing to a reduction in the federal deficit.

2.0 Equilibrium Sorting Models

In essence, equilibrium sorting models (ESMs) provide a stylized representation of the interactions of households, landlords and government within a property market. Originally developed to explain observed patterns of socio-economic stratification and segmentation in urban areas (Tiebout, 1956, Oates, 1969, Schelling, 1969, Ellickson, 1971, Epple, 1991), ESMs provide a formal account of the process whereby heterogeneous households sort themselves across the set of neighborhoods within a property market.

Neighborhoods, it is assumed, differ in quality according to the level of public goods each provides. Those public goods may reflect purely physical attributes of a location

---

if the parameters of the utility function are estimated from the model the definition of the jurisdictions will affect the assumed value of the public goods indices and influence the implied stratification of households. Second, incorrectly specifying the jurisdictions will affect the predictions that are made about how households respond to policy change. I envisage this having a greater impact in the exploration of targeted policy changes where omitting a jurisdiction removes a substitution option that is or would be available to households. It would be interesting to see future work exploring the sensitivity of WTP estimates to the definition of the choice set.

(for example, a neighborhood's proximity to commercial centers) or the levels of provision of local amenities (for example, the quality of local schools). An important distinguishing feature of ESMs is in allowing local amenity provision to be shaped by endogenous *peer effects*: that is to say, by the characteristics of the set of households that choose to locate in a neighborhood. Epple and Platt (1998), for example, present a model in which local taxes and lump sum payments are determined by the voting preferences of the residents in a neighborhood. Similarly, Ferreyra (2007) and Nesheim (2002), present models in which school quality is related to measures of the average income of households in a locality.

In an ESM, the mapping of households to quality-differentiated neighborhoods is mediated through property prices. Indeed, a solution to an ESM is taken to be a set of property prices that support a Nash equilibrium allocation of households to neighborhoods such that the supply and demand for properties are equated in all neighborhoods. While some simple ESMs have closed form solutions (Epple and Romer, 1991, Epple and Platt, 1998) equilibria for more complex models, perhaps including endogenous neighborhood quality, are usually calculated using techniques of numerical simulation (Bayer et al., 2004, Ferreyra, 2007).

Over the last decade ESMs have increased in popularity and complexity. Recent modeling extensions allow for moving costs (Bayer et al., 2009, Kuminoff, 2009, Ferreira, 2010), overlapping generations[46] (Epple et al., 2010, Bayer et al., 2011,

---

[46] Like many of the models in the literature, the ESMs developed in this chapter adopts a static framework, in reality households are more like forward looking agents who make location decisions in a dynamic setting. This issue is starting to receive attention in the literature, see Kuminoff et al. for a discussion of these developments. For example Epple et al. (2012) develop an over-lapping generations ESM in which household preferences are age dependent. In this two-period model there are young and old household types (with young households becoming old in the second period). Epple, Romano and Sieg show that introducing moving costs in their model leads to lower levels of segregation by income and age as households are forced to trade-off utility in the two periods. Similarly, Bayer et al (2011) develop a model with dynamic timing and location decisions using an infinite horizon life cycle approach in which household preferences change over the life cycle and

Epple et al., 2012) and simultaneous decisions in a parallel labor market (Kuminoff, 2007, Kuminoff et al., 2010). In addition, the ESM framework has been used to explore empirical data on the distribution of households and property prices in order to derive estimates of the value air pollution (Smith et al., 2004), school quality (Fernandez and Rogerson, 1998, Bayer et al., 2004) and the provision of open space (Walsh, 2007). ESMs have also been used to explore policy issues such as school vouchers (Ferreyra, 2007), open space conservation (Walsh, 2007, Klaiber and Phaneuf, 2010) and hazardous waste site clean ups (Klaiber and Smith, 2009). A comprehensive review can be found in Kuminoff et al. (2010).

One area that has received relatively little attention in the ESM literature is that of tenure. Indeed, the vast majority of ESM applications make the assumption that households rent their properties from absentee landlords. In those applications where tenure status has been considered it has been treated as a fixed household characteristic rather than a choice variable (Epple and Platt, 1998, Bayer et al., 2004). In reality, of course, households choose from a number of tenure options, with the key distinction being between ownership and renting. For a number of issues, such as the reform of MID policy, the choice of tenure is the central consideration of the policy debate.

Accordingly, one of the key contributions of this paper is to describe an ESM in which tenure choice is endogenised; in the model household choices to rent or purchase a property are a function of market conditions. Moreover, to capture the purported positive spillovers of homeownership, levels of homeownership in a neighborhood contribute positively to its residential desirability; in this model, rates of homeownership contribute to the provision of an endogenous local public good. Allowing for homeownership in an ESM has other ramifications, particularly with regards to simulating the outcomes of policy reform. When those policy reforms result

---

households anticipate changes in amenity values and prices. In the future, I can envisage the endogenous tenure ESM progressing into an overlapping generations framework as it provides a suitable means for exploring the issues of borrowing constraints, capital gains, inheritance and moving costs.

in price changes in the property market, homeowners and renters are impacted differently. In particular, homeowners will be faced by capital gains or losses that are not experienced by renters. The modeling framework in the next section outlines a method for incorporating that distinction.

3.0 The Model

3.1 The Economy

Consider a closed spatial economy consisting of a continuum of households. The model is closed insomuch as households may not migrate in or out of the economy. Households differ in their incomes, y, their preferences for housing, β, and preferences for homeownership, θ. Preferences for homeownership represent the private returns to homeownership that are not realized when renting. Such private returns are motivated by numerous considerations including i) freedom to modify housing, ii) satisfaction from homeownership status and iii) financial returns from capital gains. The distribution of household types in the population is defined by the joint multivariate density function $f(y, \beta, \theta)$.

The economy is divided into a set of spatially discrete neighborhoods, $j = 1, \ldots, J$ .In this model, each neighborhood is assumed to have its own local government. Henceforth, I describe these areas as jurisdictions. Each jurisdiction is characterized by a vector of local public goods, $g_j = \{z_{j,1} \ldots, z_{j,U}, q_{j,1}, \ldots, q_{j,V}\}$, comprised of $u = 1, \ldots, U$ exogenous elements, $z_{j,u}$, and $v = 1, \ldots, V$ endogenous elements, $q_{j,v}$, the level of provision of which is dependent upon the composition of the set of households that reside within the jurisdiction. The provision of public goods is assumed to be homogeneous within a jurisdiction.

3.2 The Demand Side

To reside in jurisdiction j a household must buy housing within it. The decision to rent, R, or own, O, housing is referred to as tenure choice. I describe the set of tenure options as $T = \{R, O\}$. Accordingly, our model is characterized by households choosing to participate in one of a number of property markets each defined by a jurisdiction and tenure bundle, $\{j, t\}$. The set of markets is given by $C = J \; x \; T$.

Households also choose a quantity of housing, a decision approximating real life choices over the size and quality of home to buy or rent.[47] Housing is defined as a homogeneous good that can be purchased at a constant unit price, p_j, within a jurisdiction[48] (Epple & Romer 1991, Epple & Sieg 1998, Epple & Platt 1998, Ferreyra 2007). Housing can also be rented from absentee landlords at the rental price, r_j.

The quantity of housing demanded by a household in market j, t is denoted $h_{j,t} = h(p_j, r_j, g_j;\; y, \beta, \theta)$. To become a homeowner a household must take out a mortgage[49] and pay mortgage interest, m, to the lender. Mortgage interest is paid only on the amount borrowed, given by the product of the loan-to-value ratio, δ_i, and the value of the housing purchased, $p_j h_{j,t}$. Differences in δ_i represent the varying abilities of households to make a down payment. Property taxes τ_p are paid on both rented and purchased housing.

Homeowners are permitted to itemize mortgage interest costs and property taxes, that is, to deduct these costs from their taxable income. Since the marginal rate of income tax increases with income, the implicit subsidy of itemization also increases with household income. However, not all households choose to itemize. I use the variable $item$ to denote whether a household itemizes. Empirically itemization rates are higher amongst high-income households. To account for this the model includes the probability that a household itemizes, which is expressed as a function of household income

[47] This simplification is made at the cost of assuming, somewhat unrealistically, that housing is continuously divisible and can be reconfigured without cost.

[48] In reality housing is not homogeneous, however, as Sieg, Smith, Banzhaf & Walsh (2002) illustrated, if housing enters the utility function through a sub-function that is homogeneous degree one, it is possible to construct a “housing quantity” index tantamount to an empirical analogue to the homogeneous housing unit, h.

[49] For simplicity, the model assumes that all households must take out a mortgage.

$$Prob(item = 1) = \Psi(y) \tag{4.1}$$

Where $\Psi(\ )$ denotes a cumulative distribution function and,

$$item = \begin{cases} 1 & if\ a\ household\ itemizes \\ 0 & otherwise \end{cases}$$

Accordingly, the implicit subsidy a household receives by itemizing mortgage interest payments and property tax payments on their tax return, $MID(p, h, \tau_p, t, y)$, is endogenous to the household's decision and depends upon the purchase price of property, p, the quantity of housing demanded, h, the property tax rate, τ_p, tenure choice, t, and household income, y (which also determines the loan-to-value ratio and the probability that the household itemizes).

Aggregate demand for housing in market j, t is calculated by integrating across households,

$$H_{j,t}^{D} = \int\int\int h_{j,t}.f(y, \beta, \theta)dyd\beta d\theta \tag{4.2}$$

3.3 The Supply Side

Housing supply is determined by purchase and rental property prices. Housing supply may differ between jurisdictions such that the housing supply function for a particular market j, t is denoted,

$$H_{j,t}^{S} = H^{S}(p_j, r_j) \tag{4.3}$$

Note that purchase prices may affect the supply of housing in the rental market and vice versa.

3.4 Government

Government operates at two levels, federal and local, serving the dual roles of redistributing income and providing local public goods. The federal government raises revenue through income taxes, charged at a series of marginal rates, $\boldsymbol{\tau_y}$, which are an increasing function of taxable income. The tax paid by a household is $tax_y = tax_y(y - MID, \boldsymbol{\tau_y})$ .The total federal tax revenue is,

$$T_y = \int\int\int tax_y . f(y, \beta, \theta) dy\, d\beta d\theta \tag{4.4}$$

Federal tax revenues are used to finance the provision of public goods. The revenue foregone to mortgage interest deductions is equal to the sum of the MID payments across all households.

$$TMID = \int\int\int MID. f(y, \beta, \theta) dy d\beta d\theta \tag{4.5}$$

It is assumed that the federal expenditure on local public goods provision is organized so as to allocate an equal amount of revenue per household,

$$E_j^F = S_j(T_y - TMID) \tag{4.6}$$

Where S_j is the share of the population locating in neighborhood j.

Local governments raise revenue through proportional property taxes, τ_p,[50] which are levied on the value of property. As such, the total property tax revenue of jurisdiction j is,

$$T_{p,j} = \tau_p p_j (H_{j,O}^D + H_{j,R}^D) \tag{4.7}$$

Local tax revenues are increasing in property prices and aggregate housing demand. Local tax revenues are used to finance local expenditure on public goods,

[50] Our model considers exogenous tax rates but easily extends to endogenous rates i.e. through a majority vote (Epple & Romer 1991, Epple & Platt 1998).

$$E_j^L = T_{p,j} \tag{4.8}$$

Total expenditure on local public good provision, therefore, is equal to the sum of federal and local expenditure,

$$E_j = E_j^F + E_j^L \tag{4.9}$$

3.5 Local Public Goods

Households derive utility from the combined provision of local public goods, represented by the index,

$$g_j = \Sigma_{k=1}^{U} \gamma_k z_{j,k} + \Sigma_{k=U+1}^{V+U} \gamma_k q_{j,k} \tag{4.10}$$

Where γ_k is the weight placed on the k^{th} element in g. For simplicity I consider the case where g_j consists of only one exogenous, z, and one endogenous public good, q;

$$g_j = z_j + \gamma q_j \tag{4.11}$$

Where γ is the weight that households place on q relative to z and is uniform across households and jurisdictions. Our specification implies, therefore, that households agree on the ranking of jurisdictions in terms of their provision of local public goods.

Endogenous public good provision within a jurisdiction is an increasing function of three inputs: government expenditure, E_j, homeownership rate, ρ_j and other characteristics of the community of households in that jurisdiction, x_j, such that,

$$q_j = q(E_j, \rho_j, x_j)$$

$$\frac{dq_j}{dE_j} \geq 0, \frac{dq_j}{d\rho_j} \geq 0 \; and \; \frac{dq_j}{dx_j} \geq 0 \tag{4.12}$$

Our specification makes two important assertions i) that homeownership provides positive spillovers and ii) that the proportion of homeowners in a community rather than the absolute number is important in delivering those external benefits. Those assertions are consistent with the majority of the literature on the social benefits of homeownership (Dietz and Haurin, 2003), which suggests that homeownership can have a positive impact upon crime rates, voting participation, property maintenance, neighbourhood stability, membership of community groups and the educational attainment of school children. The presence of x_j in the public good production function defines a peer effect whereby community characteristics, perhaps median household income, affect the provision of public goods. Such peer effects have considerable empirical support (Nechyba, 2003a) and have been incorporated in a number of existing ESM specifications (Nesheim, 2002, Ferreyra, 2007).

3.6 The Household Optimization Problem

Households derive utility from local public goods, g, consumption of housing, h, and other consumption, c. Preferences for local public goods are determined by the parameter α that is assumed to be constant across households while preferences for housing are determined by the parameter, β, which is allowed to vary across households.

The model also allows for the fact that households can derive more utility from housing when they own their home than when they rent it (or vice versa). Each household is characterized by a value for the parameter θ, which scales the utility derived from housing for homeownership. Values of θ greater (less) than one imply that a household gains more (less) utility from owning housing than from renting.

Household utility is defined by the function,

$$U_{i,j,t} = U\left(h, t, c; y_i, \alpha, \beta_i, \theta_i, g_j\right) \tag{4.13}$$

The household optimization problem can be decomposed into two stages. First, a household calculates its optimal housing and consumption choices for each market.

The conditional maximization problem is,

$$\max_{h,c|j,t} U_i\big(h,t,c;y_i,\alpha,\beta_i,\theta_i,g_j\big)$$
$$\text{s.t.}$$
$$y = \begin{cases} tax_y + r_j h_j + \tau_p p_j h_j + c & t = R \\ tax_y + (1+\tau_p + m\delta) p_j h_j & t = O. \end{cases} \tag{4.14}$$

Which yields the following conditional indirect utility functions,

$$V_{i,j,t} = \begin{cases} V(r_j, g_j; y_i, \alpha, \beta_i) & t = R \\ V(p_j, g_j; y_i, \alpha, \beta_i, \theta_i) & t = O. \end{cases} \tag{4.15}$$

Subsequently, households select the jurisdiction and tenure combination that provides them the greatest level of utility.

3.7 Equilibrium

An equilibrium of the model is defined by a set of jurisdictions, C , a one to one correspondence of households to jurisdictions and an associated set of rental and purchase prices for each jurisdiction, $p = \{p_1, \dots, p_j, r_1, \dots, r_j\}$, such that,

1. Each household resides in the jurisdiction that maximizes its utility given the equilibrium vector of prices and endogenous public good provision.
2. All housing markets clear, $H_{j,t}^D = H_{j,t}^S, \forall\, j, t$.
3. All local government budgets balance, $E_j = T_{p,j}, \forall\, j$.

3.8 Simulating Responses to Exogenous Policy Change

In reality policy changes occur in a world in which households already rent or own existing properties. That reality influences the outcome of a policy change in at least two ways. First, changes take place in the context of an existing housing stock whose quantity and location has been determined by households' initial choices. Second, a household's current tenure status determines whether their choices following the policy change are influenced by capital gains.[51] To see that more clearly, it is instructive to briefly contemplate how market changes impact differently on renters and owners.

Consider a change that leads to increased rental prices. When rental prices go up existing renters are unable to afford their current consumption bundle. Households can respond in a number of different ways. They can alter their tenure choice, for example, moving from owning to renting. They can also move to another location where property prices are lower and they can reduce their demand for housing and consumption. Moreover, they can do a combination of these. In contrast, consider a change that precipitates increased purchase prices. Owning a property outright prevents changes in prices from making the current consumption bundle unaffordable: homeowners are shielded against price fluctuations. Instead, a rise in prices presents homeowners with the opportunity to sell-up and use the capital gains to increase consumption or relocate to a jurisdiction that provides more desirable public goods.

To simulate the process of adjustment in the property market within the context of what is essentially a static model requires some careful consideration. I first assume that the market is in a state of long-term equilibrium, an equilibrium achieved under the baseline policy. Households have optimally chosen where to live, whether to rent or own and how much housing to consume. To reflect that state of the world, I imagine a property market in which all the housing units demanded under that baseline policy have been constructed and that these existing housing units cannot be demolished in the face of a policy change (though they can be repackaged and new units may be constructed).

The policy change is introduced into this world at a point after homeowners have paid for their current properties at the pre-change prices but before rent has changed hands,

[51] Other authors have examined the importance of moving costs in equilibrium sorting models (Bayer et al. 2009, Kuminoff, 2009). Like capital gains, moving costs can vary depending on the household's initial position and have the potential to alter the shape of the equilibrium that results from a policy change.

consumption goods have been bought and taxes and mortgage interest have been paid. As a result of the policy change, households reconsider their choices of housing units, location and tenure status and the model is solved for the set of property prices that bring the market back to equilibrium under the changed conditions. Households then buy properties or pay rent and make the tax and mortgage payments due according to their new housing decisions under these new prices.

For renters, the features of their choice problem following the policy change are little different from those characterizing their choice in the original long-term equilibrium. In contrast, the choices of homeowners under the new conditions will be influenced by the fact that any increase (decrease) in the price of their currently owned property will present them with capital gains (losses).

4.0 Simulating MID Reforms

The model developed above provides a rich environment in which to explore the general equilibrium consequences of reforming MID policy. Within that environment the impact on government expenditure, on patterns of community composition, on homeownership rates and on the levels and distribution of household welfare can be considered simultaneously. To undertake this exercise it is preferable to examine a model that replicates the real world. Such a model required reasonable but tractable functional forms that can be calibrated to produce a model that resembles a real world property market. Following the convention of Epple and Platt (1991) I specifically model Boston in 2000. To provide a clear and accessible illustration of the pathways of change that operate in light of a policy reform it is prudent to consider a simple two-jurisdiction version of the model. This simplification enables us to trace out clearly the chain of reactions that occur through these pathways. The model is coded in Matlab[52] and uses simulation and iterative numerical techniques to solve for market

[52] The Matlab code is available from the author upon request. I would like to thank Kerry Smith, Dennis Epple and Maria Ferreyra for providing data and copies of their code for solving other ESMs).

clearing prices and provision of endogenous public goods (Lagarias et al., 1998)[53].

4.1 The Proposed Policy Reforms

The current debate regarding reform of MID policy is motivated in part by the large U.S. deficit. Indeed, as part of plans to reduce that deficit, President Obama submitted federal budget proposals in 2011 and 2012 that advised capping itemized deductions, including MID, at 28 percent. Both times Congress has rejected the recommended tax reforms.[54] All the same, I take the proposal of capping MID at 28 percent as our first potential policy reform. In practice, this policy amounts to limiting the implicit tax subsidy to homeowners paying a marginal rate of income tax above 28 percent.

I also consider three alternative MID-reform policies: a refundable flat-rate tax credit, an income tax reduction and a new owner scheme. So as to compare the various proposed policies, I make the assumption that the central motivation for reform to the MID is reduction of the budget deficit. Accordingly, I calculate the reduction in deficit brought about by our baseline reform of a 28 percent cap on MID. I then tailor the three alternative MID-reform policies to ensure that they facilitate the exact same reduction in the budget deficit as the cap.[55]

Let us briefly review the alternative MID-reform policies. First, replacing MID with a refundable[56] flat-rate tax credit has been advocated by both the Center for American Progress, who propose a 15 percent refundable tax credit, and the National Commission on Fiscal Responsibility (2010), who propose a 12 percent non-refundable mortgage interest tax credit. For the purposes of our simulations, I model

[53] Due to endogeneity, the uniqueness of the equilibrium is not guaranteed. One way to explore this is to alter the initial values used in the code. In the simulations discussed below, this procedure had no influence on the outcomes, suggesting uniqueness of each equilibrium.

[54] The same proposal has been included in the 2013 budget proposals.

[55] Revenue equivalent policies were found using a search process.

[56] Here the term ‘refundable’ indicates that households whose income tax liability is lower than the value of the credit actually receive a payment from the Treasury covering that difference.

this reform as being a policy change in which MID is abandoned and, instead, all households who are owners can claim back a flat-rate percentage of their mortgage interest and property tax costs. As explained previously, that flat rate is chosen such that cost savings achieved by this policy are identical to capping MID at 28 percent.

Our second alternative MID-reform policy follows the proposal made by the Reason Foundation (Stansel, 2011) to scrap MID and instead introduce a revenue neutral reduction in federal income tax for all households. Here, I consider a policy in which MID is abandoned and a portion of the savings in government expenditure are used to fund an equal percentage reduction in income tax for all households. Again, the level of income tax reduction is chosen such that the policy achieves the same reduction in the federal government budget deficit as the other proposed reforms.

Our final alternative MID-reform policy takes motivation from the First Time Buyers scheme proposal made by Gale & Gruber (2007), which suggests scrapping MID and introducing a refundable payment to first-time buyers in the first year after a property is purchased. In the model this is achieved through a New Owner Scheme, which makes an equal lump sum payment to new homeowners. Again the level of payments to these first time buyers is chosen so as to ensure comparability in the reduction of the federal budget deficit across reforms.

4.2 Calibration

To carry out our simulations, specific functional forms have to be selected for the various structural equations of the model and parameter values for those functions must be determined. Following the example of Epple and Platt (1998), parameter values were chosen such that the model approximates the reality of the Boston Metropolitan (PSMA) area, though in our application I take data for Boston from 2000 and not 1980. Table 4.1 presents a summary of important statistics for Boston in 2000 and Table 4.2 summarizes the parameters obtained by calibrating the model to that reality. The assumptions and methods used in deriving those parameters are explained in the following.

$\bar{y}$	Mean Income (2000 USD)	74,119
y_{median}	Median Income (2000 USD)	55,183
ρ	Homeownership Rate	0.72
z	Mean Air Quality (NO_x)	3
q	Mean School Quality	420

Table 4.1: Empirical Statistics for Boston in 2000

4.2.1 Jurisdictions

To allow the pathways of response to MID reform to be studied with reasonable clarity, I explore a simple two-jurisdiction version of model. Extensions to multiple-jurisdiction models are relatively easy to implement, but greatly complicate interpretation. Again following Epple and Platt (1998), I achieve that by simply imagining that the Boston Metropolitan area is divided into two jurisdictions that I label A and B.

4.2.2 Households

Households in the model are characterized by three parameters: income, y, preferences for housing, β, and preferences for homeownership, θ. The first step in calibrating the model, therefore, is to establish the joint distribution of those parameters amongst the residents of Boston in 2000.

As made explicit shortly, a Cobb-Douglas utility function is assumed such that a household's preferences for housing, β_i, equate to the proportion of their income that they spend on housing. That data along with information on household income, y_i, is available from the census. To establish the joint distribution of y and β I fit a bivariate-normal distribution $f(lny, \beta) \sim N(\mu_f, \Sigma_f)$ to 2000 census data for Boston. Parameter values from that estimation are recorded in the first row of Table 4.2.

Notice that β is negatively correlated with y, indicating that high-income households spend a smaller proportion of their incomes on housing than low-income households.

Parameters	Description	Calibrated Value
$(\ln(y), \beta) \sim N(\boldsymbol{\mu_f}, \boldsymbol{\Sigma_f})$	Income and Preferences for housing	$\boldsymbol{\mu_f} = [10.604, 0.149]$ $\boldsymbol{\Sigma_f} = \begin{bmatrix} 1.045 & -0.0503 \\ -0.0503 & 0.007 \end{bmatrix}$
α	Preference for Local Public Goods	0.35
$\ln[\theta - 1] \sim N(\mu_\theta, \sigma_\theta^2)$	Preference for Homeownership	$(\mu_\theta, \sigma_\theta^2) = (-2.0, 9.0)$
m	Mortgage Interest Rate	0.1339
δ_0, δ_1	Parameters of the Loan to Value Ratio Function	$-0.3, \quad 6 \times 10^{-5}$
$\boldsymbol{\Psi}(y)$	Probability of Itemising	$< y <$: $Prob(item = 1)$ 38,000 : 0.234 72000 : 0.661 120,000 : 0.855 240,000 : 0.981 240,000+ : 0.999
y_{τ_y}	Income Tax Bracket (lower limits)	[0, 7350, 21925, 52975, 80725, 144,175]
τ_y	Marginal Income Tax Rate	[0, 0.15, 0.28, 0.31, 0.36, 0.396]
τ_p	Property Tax Rate	0.5
γ	Weight on School Quality in Public Goods Index	0.03
η	Elasticity of Housing Supply	3

Table 4.2: Calibrated Parameter Values

For simplicity, and due to a lack of existing empirical evidence, it is assumed that household preferences for homeownership, θ, are independent of income and preferences for housing. Accordingly, values were drawn from a log-normal distribution $ln\ (\theta - 1) \sim N(\mu_\theta, \sigma_\theta^2)$ with mean and variance chosen so as the baseline model predicted homeownership rates comparable to those observed in Boston in 2000[57]. The parameters selected through that procedure are also recorded in Table 4.2.

For the purposes of simulating the model, I create a simulated sample of 2,000 households, which I denote by $i = 1, \ldots 2{,}000$, with income, y_i, and preference parameters β_i and θ_i drawn from those estimated distributions.[58]

4.2.3 Taxes

In 2000, Federal income taxes were structured into six marginal tax brackets. Those tax brackets are defined by lower bound incomes, $\boldsymbol{y}_{\boldsymbol{\tau}_{\boldsymbol{y}}}$, at which the corresponding marginal tax rates, $\boldsymbol{\tau}_{\boldsymbol{y}}$, become payable. The first bracket ranging from income of \$0 to \$7,350 has a marginal tax rate of zero. Accordingly, \$7,350 is often referred to as the standard deduction. The tax brackets and associated marginal tax rates are recorded in Table 4.2. Table 4.3 illustrates how the income tax payable is calculated for households in each of the tax brackets.

The property tax rate, τ_p, was set at the average level for Boston in 2000 using data supplied by the Massachusetts State Government.

To capture the correlation between income and itemization rates, the probability of a

---

[57] Equilibria were also characterized for a range of alternative to explore the sensitivity of the results to the parameterization and to allow consideration of the range of permissible outcomes. The results remain qualitatively unchanged and are not reported here, however a full set of results is available from the author upon request.

[58] The baseline model was also run for population sizes of 500 and 10,000. This did not alter the results and conclusions that could be drawn.

household itemizing was calibrated using data on itemization rates by income in Poterba & Sinai (2008), which is reproduced within Table 4.2.

Bracket	**Income is greater than…**	**but less than …**	**Tax Payable**	
1st	0	7350	0	
2nd	7350	21925	0	+15% x amount over 7,350
3rd	21925	52975	2,186.25	+28% x amount over 21,925
4th	52975	80725	10,880.25	+31% x amount over 52,975
5th	80725	144175	19,482.75	+36% x amount over 80,725
6th	144175	-	42,324.75	+39.6% x amount over 144,175

Table 4.3: Income Tax Brackets 2000

4.2.4 Mortgages

In our model, the size of mortgage needed by a homeowner is determined by their loan-to-value ratio parameter, δ_i. For the purposes of the simulation, the relationship between loan-to-value ratio and household income was estimated empirically using data from the Survey of Consumer Finances reproduced in Poterba & Sinai (2008). Using that estimated relationship the parameter δ_i was calculated as,

$$ln\ \delta_i = \delta_0 - \delta_y y_i$$

Where δ_0 and δ_y are the estimated regression coefficients,

$$\begin{array}{l} ln\ \delta_i = -0.3 - (0.00006) y_i \\ \qquad\qquad (0.058) \ \ (0.000009) \end{array} \tag{4.16}$$

Since δ_y is positive, wealthier households face lower loan-to-value ratios and, as a consequence, lower marginal costs of purchasing housing.

The mortgage interest rate, *m*, was set to the average level for Boston in 2000 using data supplied by the Federal Housing Finance Association.

4.2.5 Housing Supply

Housing supply is specified using a Cobb-Douglas function following Epple and Romer (1991) and Epple and Platt (1998),

$$H_{j,t}^{S} = A_{j,t} p_{j,t}^{\eta} \tag{4.17}$$

Where $A_{j,t}$ is a constant reflecting property market factors such as local zoning restrictions, p is the price of a homogeneous unit of housing and η is the price elasticity of housing supply.

Following Epple and Platt (1998) η is set to three in all markets for the baseline simulation[59].

As I mentioned in the introduction, it has been noted that the degree to which reforming mortgage interest deduction will lead to falling homeownership rates or falling property prices depends on the price elasticity of housing supply. The housing supply elasticity adopted here was 3 which implies quite elastic housing supply. More recent work by Siaz (2010) suggests a price elasticity that is closer to 1 might be more suitable for the Boston metropolitan area in 2000. Whilst the pathways of adjustment would remain the same, as the elasticity of housing supply contracts I would anticipate that the results of the model simulations will also change: A tighter elasticity will work to offset the gains in utility derived from reductions in rental prices since the influx of previous owners into renting will cause greater pressure on prices the lower the elasticity of housing supply. I would expect this to also lead to fewer households moving out of owning, thus causing property prices to fall by a smaller amount than in the case with more elastic housing supply. As a result, it is likely that the less elastic housing supply is, the smaller the potential for gains in homeownership from reforming MID.[60]

---

[59] Alternatively, η could be set to 0 to produce a completely inelastic housing supply.

[60] Replicating the simulations with a price elasticity of housing supply equal to 1 confirms this intuition. The patterns of sorting observed in the baseline remain the same, although

4.2.6 Local Public Goods

For the purposes of simplification, the calibrated model considers only one exogenously determined local public good and one endogenously determined local public good. The extension to multiple local public goods is facile, but adds complexity to the interpretation of the simulation results.

I take air quality to act as a representative exogenous local public good. In our simulation, air quality is defined in units of nitrogen oxides concentration (measured in pphm) below the highest level observed in Boston in the Massachusetts Air Quality Report. Using that measure, the mean level for air quality in Boston in 2000 was 3. Accordingly I set air quality in jurisdiction B to that level but assume that jurisdiction A offers a slightly higher level of provision, 4.

Likewise, I take school quality to act as a representative endogenous local public. School quality is a natural choice in this regard since empirically it is correlated with many other measures of local public good provision (Black, 1999, Bayer et al., 2004, Bramley and Karley, 2007). Also there is an increasing body of evidence to suggest that school quality is determined, in part, by levels of local homeownership (Dietz, 2002, Dietz and Haurin, 2003).

Following Nechyba (2003a), Nechyba & Strauss (1994), Ferreyra (2007) and Fernandez & Rogerson (1998) school quality is determined by a production function. The functional form adopted in those papers is extended here to include a term relating to homeownership,

---

equilibrium property prices and rents are higher, reflecting the less elastic supply. Evaluating the 28% cap, the same patterns of adjustment are observed: The average owned property size falls as demand contracts; purchase prices fall; some lower income renters are encouraged into homeownership; some previous owners become renters. However, in this case the inflow of demand into the rental sectors causes overall increases in rental prices which remove some of the gains to renters.

$$q_j = AE_j^{\phi_1} {y_{median}}_j^{\phi_2} \rho_j^{1-\phi_1-\phi_2} \tag{4.18}$$

Where E is expenditure per pupil, y_{median} is median household income and ρ is the homeownership rate.

To calibrate the production function (19), I regress a state level measure of school quality (combined fourth grade mathematics and reading attainment score) against state level measures of median household income, homeownership rates (both taken from 2000 census data) and data from the National Assessment of Educational Progress (NAEP) on expenditure per pupil. The resulting regression equation was,[61]

$$\ln(q) = \underset{(0.428)}{2.17} \quad \underset{(0.044)}{+0.14\ln(E)} \quad \underset{(0.061)}{+0.26\ln(y_{median})} \quad \underset{(0.075)}{+0.6\ln(\rho)} \tag{4.19}$$

4.2.7 Household Preferences

The household utility function is specified as a Cobb-Douglas according to,

$$U_{j,t} = \begin{cases} g_j h^{\beta} c^{1-\alpha-\beta} & t = R \\ g_j \theta h^{\beta} c^{1-\alpha-\beta} & t = O. \end{cases} \tag{4.20}$$

Carbone and Smith (2008) show that when household preferences are assumed to be Cobb-Douglas, preference parameters for non-market goods can be simply retrieved using estimates of the implicit prices of those non-market goods taken from non-

---

[61] Standard errors are shown in parentheses. In the computed equilibria, income and expenditure are deflated to match the school production function, which was estimated in 2002 dollars.

This relationship suffers from potential endogeneity problems since median income, homeownership rates and expenditures are determined by the sorting of households. In future work I would collect data to facilitate a two stage least squares estimation or calibrate the school production function using values in the literature.

market valuation exercises. Here I take the implicit price of air quality, p_z, from the hedonic study by Harrison & Rubinfeld (1978)[62] and the implicit price of school quality, p_q, from the hedonic study by Bayer, Ferreira & McMillan (2007). Following Carbone and Smith, preferences for public goods, α, and the weighting parameter, γ, can then be calculated according to,

$$\alpha = \frac{p_z z_0 + p_q q_0}{y + p_z z_0 + p_q q_0} \tag{4.21}$$

$$\gamma = \frac{p_q}{p_z} \tag{4.22}$$

Where p_z and p_q are implicit prices for air quality and school quality respectively, and subscript 0 denotes a baseline value. The calibrated values from this procedure are 0.35 for α and 0.03 for γ.

4.3 Results

The long-run equilibrium under current policy conditions was calculated for a simulated sample of 2,000 households.[63] The impact of MID-reform was then investigated by finding the new equilibrium characterizing the property market when each of the four proposed policy reforms was instituted from that baseline.

Tables 4.4 and 4.5 describe important features of the equilibrium in the baseline and for each policy-reform scenario. Table 4.4 presents a characterization of those equilibria in terms of the composition and characteristics of the communities in each jurisdiction. Table 4.5 characterizes the equilibria from the perspective of households in each of the six tax brackets. Throughout our discussion of the results I will use the

---

62 The Harrison and Rubinfield (1978) study was chosen on account of the location of the study, this being the Boston SMSA.

63 In choosing a simulated sample size one faces a trade-off between small sample bias and computational efficiency. For the baseline scenario I experimented with larger population sizes up to 10,000, but found no significant changes in the characteristics of the equilibrium.

term “price” to refer to the price inclusive of property tax since this is the effective price faced by households.

4.3.1 Baseline with MID

Consider how the equilibrium evolves under the current system of MID. In the baseline A and B differ initially only in their exogenous provision of public goods. The better air quality in jurisdiction A shapes the resulting equilibrium. Households prefer a greater provision of public goods which increases demand for housing in A relative to B. Consequently, as shown in Table 4.4a, the population of A is higher than the population of B, with 82.3 percent of all households residing there. As the supply of housing in A is not infinitely elastic, relatively stronger demand in A drives the prices of housing in A above the prices in B. The purchase price of housing (including property tax) per unit of homogeneous housing is \$116.66 in A compared to \$75.09 in B. Likewise, the rental price of housing is \$108.25 in A and \$72.47 in B. The wedge between purchase and rental prices implies that on average preferences for homeownership more than offset mortgage costs.

Price differences between jurisdictions and tenure options precipitate the stratification of households. Column 4 of Table 4.4 confirms that households with relatively strong preferences for homeownership, θ, choose to purchase housing whilst those with relatively weak preferences for homeownership rent housing. Similarly, as can be seen from column 5 of Table 4.4a, households with who spend a relatively large proportion of their income on housing, high β, prefer lower housing prices and choose to reside in jurisdiction B. Since β is negatively correlated with income, this also introduces segregation by income. As shown in Table 4.5a, only 40 percent of households in the lowest income tax bracket (1st) choose to live in A compared to 100 percent in the highest tax bracket (6th). Consequently, the median income of households in A is almost 3 times that of B.

Within each jurisdiction some households rent whilst others own. Recall from the calibration that households with higher incomes face relatively lower loan-to-value ratios and, under the existing MID policy, can itemize their mortgage interest and property tax costs at a relatively higher marginal rate. Accordingly, the marginal cost of purchasing housing is lower for higher-income households and, ceteris paribus, households with high incomes are more likely to become homeowners. As shown in Table 4.5a, only 50 percent of households with incomes below the standard deduction choose to own compared to 72 percent of households in the highest income tax bracket. This result is consistent with observed homeownership rates in Boston in 2000. Returning to Table 4.4a, the concentration of higher income households in A leads the homeownership rate to be slightly higher than in B.

Recall from equation (4.5) that local property tax revenues depend on both purchase prices and the total quantity of housing demanded in a jurisdiction. In the baseline equilibrium, higher property prices in A more than offset larger property sizes in B such that tax revenues per household in A exceed those in B: $22,097 and $21,490 respectively. Larger local tax revenues translate directly into higher levels of local government expenditure on the endogenous public good. In addition, since median income is higher in A than B (column 7, Table 4.4a), jurisdiction A also benefits from relatively larger provision of the public good through a stronger peer effect. Overall, provision of the endogenous public good is higher in A, with a school quality score of 450 than it is in B, at 323. That difference in provision of the endogenous public good acts to exaggerate further the patterns of sorting sparked by the initial difference in public goods provision.

			Exogenous Public Good	Population Share	Price of Housing ($)	Mean preference for homeownership	Mean preference for housing	Mean Income ($)	Median Income ($)	Homeownership rate	Mean property size	Local expenditure ($)	Endogenous public good
			z	POP	p'	θ	β	$\bar{y}$	y_{median}	ρ	$\bar{h}$	e	q
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Baseline													
a	A	Purchase	4	0.51	166.66	1.1	0.2	86020	49223	0.62	114	22097	440
		Rental		0.31	108.25	1	0.2	67820			135		
	B	Purchase	3	0.1	75.09	1.1	0.4	25230	16546	0.57	153	21490	323
		Rental		0.08	72.47	1	0.4	22630			175		
28% Cap													
b	A	Purchase	4	0.53	113.25	1.1	0.2	83550	48988	0.64	111	22621	457
		Rental		0.3	107.39	1	0.2	71000			141		
	B	Purchase	3	0.1	73.49	1.1	0.4	24530	16613	0.58	154	22092	326
		Rental		0.07	71.72	1	0.4	24000			184		
19% Flat Rate Tax Credit													
c	A	Purchase	4	0.5	107.31	1.1	0.2	86980	48800	0.6	117	22461	442
		Rental		0.33	106.22	1	0.2	65540			138		
	B	Purchase	3	0.1	68.85	1.1	0.4	28180	16911	0.58	152	21786	326
		Rental		0.07	72.71	1	0.4	23670			179		
2.89% Income Tax Reduction													
d	A	Purchase	4	0.52	113.25	1.1	0.2	86000	47755	0.63	111	22683	457
		Rental		0.31	104.71	1	0.2	65540			138		
	B	Purchase	3	0.1	77.76	1.1	0.4	27040	16911	0.6	170	22592	334
		Rental		0.07	72.03	1	0.4	22660			179		
New Owner Scheme													
e	A	Purchase	4	0.6	117.2	1.1	0.2	79266	48837	0.73	100	21535	490
		Rental		0.23	98.17	1	0.2	78770			186		
	B	Purchase	3	0.14	81.45	1.1	0.4	22855	16033	0.76	147	21183	379
		Rental		0.03	70.36	1	0.4	29963			240		

* Where $p' = (1 + \tau_p)p$ for the purchase market and $p' = r + \tau_p p$ for the rental market. All prices are in 2000 USD.

Table 4.4: Characterising Equilibrium by Jurisdiction

		Share owning in A AO	Share renting in A AR	Share owning in B BO	Share renting in B BR	Mean property size h	Homeownership rate ρ	Mean endogenous public good q	Cost of policy C	Share gaining utility $\uparrow U$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Baseline										
a	1st Lowest	0.2	0.2	0.3	0.3	28.85	0.05	370	0	
	2nd	0.37	0.28	0.2	0.15	67.65	0.57	399	26	
	3rd	0.49	0.35	0.1	0.06	118.03	0.59	423	199	
	4th	0.62	0.34	0.03	0.02	174	0.64	436	868	
	5th	0.64	0.32	0.03	0.01	225.2	0.67	437	1484	
	6th	0.72	0.28	0	0	313.5	0.72	441	4477	
28% Cap										
b	1st Lowest	0.25	0.18	0.29	0.29	28.88	0.54	393	0	1.00
	2nd	0.39	0.26	0.21	0.15	67.59	0.60	420	124	1.00
	3rd	0.51	0.33	0.10	0.06	118.32	0.61	444	847	1.00
	4th	0.63	0.33	0.03	0.02	174.39	0.65	457	2284	1.00
	5th	0.64	0.32	0.03	0.01	227.01	0.67	458	3388	0.84
	6th	0.69	0.30	0.00	0.01	319.55	0.69	462	4978	0.83
19% Flat Rate Tax Credit										
c	1st Lowest	0.16	0.23	0.31	0.3	28.49	0.47	371	71	0.56
	2nd	0.34	0.31	0.19	0.15	66.88	0.53	402	204	0.38
	3rd	0.47	0.37	0.1	0.06	117.7	0.57	424	382	0
	4th	0.63	0.32	0.04	0.01	170.48	0.67	436	757	0
	5th	0.64	0.32	0.03	0.01	224.52	0.67	437	893	0
	6th	0.69	0.29	0	0.01	315.44	0.7	440	1093	0
2.89% Income Tax Reduction										
d	1st Lowest	0.27	0.2	0.26	0.27	29.82	0.53	390	0	0.97
	2nd	0.38	0.28	0.19	0.14	69.11	0.58	413	28.83	0.98
	3rd	0.5	0.34	0.1	0.05	120.65	0.6	434	155.48	0.97
	4th	0.63	0.31	0.04	0.02	176.3	0.69	446	386.1	0.85
	5th	0.63	0.33	0.03	0.01	234.44	0.64	448	768.52	0.87
	6th	0.73	0.26	0.01	0	320.93	0.74	451	2384.49	0.93
New Owner Scheme										
e	1st Lowest	0.37	0.06	0.5	0.06	30.12	0.87	427	1417	1
	2nd	0.52	0.11	0.29	0.08	66.4	0.81	449	1154	0.99
	3rd	0.6	0.25	0.1	0.05	122.96	0.7	474	674	0.97
	4th	0.64	0.3	0.04	0.02	183.48	0.69	484	495	0.86
	5th	0.66	0.3	0.03	0.01	240.92	0.68	485	384	0.88
	6th	0.72	0.27	0.01	0	336.08	0.72	488	180	0.93

Table 4.5: Characterising Equilibrium by Income Tax Bracket

4.3.2 28% Cap

Now consider how things change when MID is capped at a rate of 28 percent. Under the new policy those households in the top three tax brackets who had previously been able to itemize their expenditures on mortgage interest and property tax at 31, 36 and 39.6 percent respectively, would now be limited to itemizing at 28 percent. In the absence of other adjustments, the cap raises the per-unit cost of housing for the 85 percent of households in the top three income tax brackets that itemize mortgage on their tax returns in the baseline. Characteristics of the new equilibrium are presented in Tables 4.3b and 4.4b.

The immediate impact of the reform is a reduction in demand for housing amongst existing owners in the top three tax brackets. As shown in column 3 of Table 4.4b, contracting demand is accentuated by the assumption that housing units cannot be destroyed leading to downward inelasticity in the housing supply function. Accordingly purchase prices shrink by around 2.3 percent in both jurisdictions. Interestingly, this fall in purchase prices stimulates entry into homeownership by previous renters. Ultimately, the reform transforms the purchase market in A from a small number of large properties to a larger number of smaller properties. In contrast, average property sizes in B edge up as a result of an influx of high-income homeowners, forced out of A by the increased costs of a mortgage, who opt to buy large properties in jurisdiction B’s relatively cheaper property market.

These adjustments spill over into the rental market. The movement of a number of renters into the purchase market contracts demand for rental housing. To meet the rental supply constraint, rental prices fall (by 0.8 percent in A to $107.39 and by 1.1 percent in B to $71.72) inducing existing renters to expand their housing demand. Consequently, adjustments in the rental market are the mirror image of those in the purchase market; a smaller number of households now rent somewhat larger properties. Comparing column 9 of Table 4.4a and 4.4b confirms that the mean owned property size falls by 1.4 percent whilst the mean rental property size increases by 4.1

percent[64]. Despite some migration from B to A, the median incomes of both are almost entirely unchanged (column 7, Table 4.4b).

Further effects are triggered through the impact of those adjustments on the levels of provision of the endogenous public good: First, through changes in homeownership and, second, through changes in local property tax revenues. While homeownership rates fall amongst those in the very top tax brackets, this direct impact is more than offset by increased homeownership in the lower tax brackets occasioned by falling purchase prices. Overall, homeownership increases by 1.3 percent in A and 0.8 percent in B. Property tax revenues rise in both jurisdictions despite the fall in purchase prices because of a rise in the total quantity of housing demanded. The property tax revenues per household go up by 2.4 percent in A and 2.8 percent in B. The combined effect is a rise in the provision of public goods in both jurisdictions, with school quality rising by 1.6 percent to a score of 456 in A and by 0.9 percent to a score of 326 in B.

Perhaps unexpectedly, despite policy reform constituting a significant (23.2 percent) reduction in federal government spending, within the simulation the knock-on effects of a policy capping MID at 28% actually precipitates general welfare increases for households in our simulated population. The key driver of that finding is that a policy that increases mortgage costs for high-income households has the effect of reducing their demand for housing without causing those households to switch out of owning. Falling demand for housing reduces property prices, having positive impacts for low-income households and encouraging more of them into homeownership. Consider the

---

[64] This is partly due to the assumption that the housing stock is divisible and can be easily re-packaged, in reality this is like dividing a house into several flats etc. As discussed in chapter 3, although repackaging, also referred to as conversion, has been discussed in a number of papers (Capozza and Helsley, 1989, Pogharian, 1990, Rothenberg et al., 1991, Montgomery, 1992, Maddison, 2000) that discuss costs arising from bargaining, regulatory requirements and construction, the housing supply literature has focused on new construction and, as a result, it is difficult to find any empirical work quantifying the extent of these costs.

final column of Table 4.5b, the rise in public good provision and fall in rental and purchase prices leads to utility gains for all households in the lowest four income brackets. Moreover, despite being most directly and adversely affected by the cap, 84 and 83 percent of households in the 5^{th} and 6^{th} income tax brackets experience gains in utility, primarily through the increased levels of provision of local public goods.

4.3.3 Refundable Flat-Rate Tax Credit

A seemingly more progressive reform of MID would be to replace the current system with a refundable flat-rate tax credit. Under this policy, rather than being able to claim MID against income tax, the federal government reimburses all homeowners a certain percentage of their mortgage interest payments. To maintain comparability with the MID cap reform, I consider a refundable tax credit of 19 percent which leads to the same overall deficit reduction.

In contrast to capping MID, the introduction of a tax credit has immediate implications for all households. In the absence of any other adjustments, the marginal cost of purchasing housing reduces for households in the lowest two tax brackets and all non-itemizers. For itemizers in the top four tax brackets the marginal cost rises. For the top tax brackets, the MID cut is more severe than under the cap (down to 19% compared to 28%). Accordingly, as in the case of the cap, the reduction in MID leads to a contraction in housing demand amongst previous owners in the top tax brackets. Since those households are nearly exclusively located in A that demand contraction reveals itself as a fall in purchase prices in that jurisdiction to \$107.31 (column 2, Table 4.4c).

The pressures pushing down demand for property purchases in A are partially offset by further adjustments in the property market. First, a number of households who were renting in A choose to switch to owning in that jurisdiction as the price of purchasing falls. Second, low-income homeowners in B, who were not previously itemizing, are benefited by the tax credit and the reduction in property prices. Some of those households now choose to purchase property in A. This in-migration of relatively low-income households serves to reduce median income in jurisdiction A (column 7, Table

4.4c).

As previous owners become renters, the rental community in A grows putting upward pressure on rental prices. Those increases are more than offset, however, by falls in local government property taxes. Recall property taxes are based on the purchase price of housing and since purchase prices are falling in A, so are property tax bills. Overall, the rental price in A (inclusive of property tax) falls to $106.22. The mean property size in the rental market in A falls slightly as downsizing by existing renters outweighs the larger housing demand of previous owners switching to renting. Ultimately, falling homeownership and falling median household income in A are responsible for a reduction in public good provision in that jurisdiction (column 11, Table 4.4c).

Complex patterns of change are also observable in jurisdiction B. First, in the face of higher costs some high-income itemizing owners in B reduce their property sizes while others become renters. As a result, purchase prices in B fall to $68.85. Second, non-itemizing owners use the tax credit to purchase larger properties. Third, a number of renters now stand to benefit from the tax credit by becoming owners. In doing so, they opt for smaller properties. As is shown in columns 8 and 9 of Table 4.4c, in contrast to jurisdiction A, the homeownership rate in B rises slightly by 0.4 percent and the mean size of owned properties increases. Demand for rental housing increases as households switch from owning to renting. The upward pressure this exerts on rental prices is not completely offset by reductions in property tax (via lower purchase prices) and the rental price (including taxes) rises to $72.71. Changes in rental price and housing demand increase tax revenues, which coupled with higher median income (due to inward migration of higher-income households from A) and homeownership causes the endogenous provision of local public goods in B to increase by 0.9 percent to 326.

The flat-rate tax credit causes substantial changes across the economy, stimulating changes in the tenure and location choice of over 12 percent of the population. The progressive nature of the policy makes it unsurprising that the majority of the benefits are focused upon the lowest two tax brackets. What is surprising, however, is that a smaller proportion of households in the bottom two income tax brackets benefit from the tax credit in comparison to the cap, primarily as a result of the negative impact on public goods provision in A. In addition, more substantial increases in the costs of a mortgage and a reduction in public goods provision results in losses for all households in the top four income tax brackets.

4.3.4 Income Tax Reduction

Consider next a policy that removes MID and uses the resultant tax revenues to reduce federal expenditure (by 23.2 percent to maintain comparability with the cap policy) and reduce income taxes by cutting all positive marginal tax rates by 2.89 percent.[65]

For non-itemizers and renters, this reform is generally positive. Their lower income tax liability opens up the possibility of consuming larger properties or relocating to A to enjoy relatively higher levels of public good provision. For homeowners, the immediate impact of the reform depends on income. Households in the lowest tax bracket do not pay income tax and, as such, are not immediately affected by the policy reform. For homeowners in the 2nd tax bracket, housing expenditure represents a significant proportion of their income. The scrapping of MID is a significant loss and, for those with incomes below \$50,000, more than offsets their reduced income tax liability. Accordingly, the dominant patterns of change are for households in this group to switch to renting or to relocate as owners to B. In contrast, homeowners with income above \$50,000 tend to be advantaged by the policy reform. Accordingly, despite facing higher marginal costs of purchasing, many households in this group remain owners though some choose to switch out of owning and use their increased income to rent larger properties.

---

[65] In this discussion I abstract from changes in labour supply that may occur as a result of the reduction in income tax. Households receiving a lump sum payment, proportional to their income tax burden may increase or decrease their labour supply in response, which would have a further effect on the income and consumption choices made by these households. In future work it would be useful to incorporate the labour market to allow these interactions to be modelled. In the absence of a labour market model I am implicitly assuming that household labour supply decisions are not altered by a reduction in their income tax burden.

The characteristics of the new equilibrium are presented in Tables 4.4d and 4.5d. Overall, in A the population increases, mean property size falls as new owners purchase smaller properties than existing owners and the purchase price falls by 0.6 percent to $115.95 as low income owners exit the market. Rental prices also fall, by 3.3 percent, to $104.71 as demand declines (column 3, Table 4.4d). Higher homeownership and tax revenues serve to increase the provision of endogenous public goods, providing indirect benefits to some households at the bottom of the income distribution not benefitting directly from the tax reduction. In jurisdiction B, higher housing demand increases homeownership by 2.2 percent and raises purchase prices by 3.5 percent (column 8, Table 4.4d). Higher median income and tax revenues also contribute to a greater provision of endogenous public good.

Despite its seemingly regressive design, this policy increases utility for the majority of households. For higher income households, the utility benefits of the income tax cut tend to outweigh the loss in MID. Indeed, the proportion of households in the top two income tax brackets who gain from this policy reform is larger than under the 28 percent cap.[66] Even amongst households in the lowest tax bracket where the policy has no immediate impact, 97 percent experience gains in utility (column 9, Table 4.5d). Those gains are largely achieved through other market adjustments that increase public good provision.

4.3.5 New Owner Scheme

This final policy reform replaces MID with a New Owner Scheme that pays a lump sum of $2,250 to new homeowners. Again, this is revenue equivalent to the MID cap policy. The characteristics of the equilibrium under this policy appear in Tables 4.4e and 4.5e.

---

[66] Although administrative costs are not explicitly included in this model, future work may wish to consider the additional advantage of the tax reduction's lower administrative demands.

As with the other reforms, the removal of MID has the immediate effect of contracting housing demand amongst existing homeowners and encouraging substitution towards renting. Those responses put downward pressure on purchase prices and upward pressure on rental prices. The introduction of a new owner payment, however, stimulates entry into homeownership amongst previous renters. As households exit the rental market, pressure on rental prices is alleviated and the final result is that those prices fall significantly, by 9.1 percent in A and 2.9 percent in B, to meet the housing stock constraint. In contrast, the demand pressure created by these new homeowners results in increased purchase prices, rising by 0.4 percent in A and by 8.4 percent in B (column 3, Table 4.4e). While the group of new owners is large, each household demands a relatively small property. Mean owned property sizes fall and homeownership rises significantly, reaching 73 percent in A and 76 percent in B.

Despite higher purchase prices, tax revenues in both jurisdictions fall as a result of the lower total housing demand. While, migration between A and B reduces median incomes in both jurisdictions, the enhanced peer effect from increased homeownership causes the provision of endogenous public goods to rises by 8.8 percent in A and 17.3 percent in B (column 11, Table 4.4e). Accordingly, previous homeowners who lose the MID and gain no advantage from the new policy, are compensated in two ways. First, since property prices rise, they benefit from capital gains. Second, they benefit from increased levels of public good provision.

While focusing on new owners, this policy reform results in widespread welfare gains across the spectrum of households. In fact, these welfare gains Pareto-dominate those generated by the Income Tax Reduction[67]. The key pathway through which those gains are delivered is by encouraging a substantial movement of households into homeownership, a movement which increases the value of the properties of homeowners, reduces prices for those remaining in the rental sector and contributes to

[67] Again, it is worth remembering however that this model does not account for labour market decisions and could be improved by extending the analysis to a dual market model (Kuminoff, 2007).

a general increase in public good provision.

5.0 Discussion

This paper contributes methodologically to the existing literature by developing an ESM that incorporates an explicit endogenous tenure decision as well as endogenous local public goods that depend partly on homeownership. These innovations extend the range of policy problems to which ESMs can be applied to include those where tenure choice and the impact of policy reform on rates of homeownership are central. Moreover these innovations allow us to account for the influence of capital gains and housing stock constraints on the distribution of benefits.

A simplified model is calibrated to real world data and used to examine the possible consequences of reforms to the policy of MID in the U.S. This exploration begins to shed some light on the complex patterns of change that such reforms may precipitate in the property market and provides insights that help to inform some of the more acrimonious disputes surrounding the debate over MID reform. With regard to that debate, our calibrated simulations show that the impact of removing MID depends crucially on the nature of the policy that takes its place.

First, consider the argument that MID inflates property prices making homeownership less affordable (Glaeser & Shapiro 2002). Our results suggest that capping MID would indeed lead to a general fall in the purchase price of properties. This is also the case when I consider the introduction of a revenue equivalent flat-rate tax credit and a tax rebate. In contrast, the targeted New Owner Scheme increases purchase prices as demand for homeownership booms.

Second, supporters of MID argue that removing it would damage homeownership rates. Our simulations suggest that the impact of reform on homeownership may be positive or negative. For the Cap, Income Tax Reduction and New Owner Schemes I predict increased homeownership as purchase prices fall and new incentives for homeownership are introduced.

Third, critics of MID argue that it subsidizes excessive housing consumption amongst wealthy households, suggesting that the removal of MID would lead to a contraction in the average property size of owners in the top tax brackets. As with the homeownership rate, our simulations suggest that the nature of the policy reform has a strong influence on the mean property sizes demanded by households in each income tax bracket. Contrary to previous predictions, however, under each reform the mean property size demanded by households in the top income tax bracket increases as a result of lower purchase prices and, in the cases of the tax rebate and new owner payment, increases in disposable income.

Examining a range of alternative policy reforms also demonstrates the importance of policy design and the role of path dependency in shaping the outcome of those reforms. With the regard to the latter, there are three key mechanisms at work. First, owning a property shields high-income households from changes in property prices and subsequently enables them to channel benefits through capital gains. Second, housing stock constraints act to suppress prices and stabilize homeownership in the face of contracting demand. Third, endogenous public goods can act as a mechanism for compensating households. As a result, the complex patterns of change precipitated by policy reforms in the property market can have quite unanticipated results. Policies designed to be progressive, such as the tax credit reform, may do less to benefit poorer households than those that appear to be regressive, such as the income tax reduction reform. Likewise, policies that economists would normally assume to have excellent efficiency improving qualities, such as the income tax reduction reform, may be Pareto-dominated by others that bear none of those hallmarks, such as the new owner payment reform. Taken as a whole, our investigation suggests that several reforms to MID could maintain the prevailing levels of homeownership whilst delivering more public goods and contributing to a reduction in the federal deficit.

Of course, these results relate to the calibration of a simplified two-community problem. Given the results, it would be interesting to see future work directed towards the estimation of a large-scale model with more formally quantified social returns to homeownership. With these extensions it would be possible to simulate economy wide responses to the proposed reforms. Nonetheless, the results demonstrate the usefulness of the modeling framework and provide important insights into the broader implications of reforming MID.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

1.0 Concluding Remarks

The research undertaken in this thesis has been motivated by the challenges to economic analysis raised by the increasing importance of space. In particular, the thesis is divided into two parts that draw together four chapters, examining issues in estimation and prediction in the context of environmental economics.

Chapter 1 began by identifying a broad class of spatial issues that arise in the analysis of data that is spatially organised. I show that these spatial issues can be thought of as spatial data generating processes (DGPs) in the context of estimation. Furthermore, I demonstrate that this broad range of spatial-DGPs is encompassed in a general framework. Having formalised the spatial data issues chapter 1 considers the conventional spatial parametric estimation solution and contrasts this to a semi-parametric spatial smoothing approach. I derive a set of three identification conditions pertaining to a general spatial smoothing estimator and translate these into two sufficient identification conditions for the use of a local polynomial regression based spatial smoothing estimator. These conditions provide intuitive and concise guidance for applied analysts that can be referred to when considering the use of spatial smoothing in the analysis of spatial data. Finally, a series of Monte Carlo simulations are presented to illustrate the performance of the spatial smoothing estimator and contrast it to OLS and SEMs. Moreover, I demonstrate how the spatial smoothing estimator can be used in combination with a spatial econometric model to allow analysts to test for specific interaction effects.

Chapter 2 extended this work to develop a local adaptive spatial smoothing estimator (LASSE). I derive a comparable set of identification conditions for a LASSE and demonstrate the conditions under which the LASSE provides efficiency gains over a Fixed SSE alternative, which employs the same smoothing parameters at every location. The identification conditions reveal that the LASSE provides greater opportunities for identification and are always as efficient, if not more efficient, than an alternative Fixed SSE making the LASSE the preferable method of estimation. Chapter 2 also presents a series of Monte Carlo simulations that illustrate the performance of LASSE and contrast it to a Fixed SSE.

In Part II I progress to consider a number of issues associated with economic policy evaluation in spatial models. Chapter 3 sets out to contrast a conventional partial equilibrium approach to welfare analysis with a general equilibrium approach provided by the framework of equilibrium sorting models (ESMs) for evaluating policies that impact local environmental quality and give rise to adjustments in household location choices. The central contribution of Part II of the thesis is the development of an ESM with an endogenous tenure choice and refined middle to long term housing supply function. This model is capable of accounting explicitly for the myriad of complex responses and feedbacks that occur in response to a policy change, and therefore provides a pivotal tool for conducting medium to long term welfare analyses of policies with spatial impacts. Moreover, the model is able to differentiate between households in terms of their characteristics (income, tenure status and preferences) and, as a result, permits a detailed analysis of both the magnitude and distribution of welfare changes.

Chapter 3 develops the ESM with endogenous tenure choice and examines the intuition of the ESM model through a stylized two-neighbourhood model in which the key mechanisms of the model can be explored. The chapter then uses the model to explore how project evaluation using a conventional static analysis compares to an ESM based analysis that allows for property market adjustments using the Polegate bypass case study to calibrate the model and illustrate how the model might be used to provide input in a real policy context. Through allowing for rental and purchase markets the ESM model with endogenous tenure choice provides a far richer characterisation of the differences between the two approaches. Within the ESM framework I was easily able to adapt the model to explore how the magnitude and distribution of welfare are impacted when I allow for the possibility that households in the neighbourhood are also landlords and hence receive income from rental payments made by other households. The analysis highlights the complexity of the adjustments that follow a policy change and the distortionary impact that they can have on the distribution of welfare changes. Importantly, the model demonstrates the limitations of *ex ante* policy targeting and, furthermore, suggests that environmental policies are likely a poor tool for achieving redistribution as the property market channels gains from renters towards owners through increases in rental and property prices.

Chapter 4 adapted the ESM with endogenous tenure choice to examine a current policy question under debate in the US: Whether Mortgage Interest Deduction (MID) should be reformed. Chapter 4 contributes to the MID debate by exploring a number of potential revenue equivalent policy reforms using an ESM with simultaneous rental and purchase markets and an endogenous tenure choice. The model is extended by additionally endogenising local quality, such that the quality of each local jurisdiction is determined partly by its levels of homeownership and partly by local tax revenues collected from property taxes. The public policy relevance of the model is demonstrated through a calibration exercise for Boston, Massachusetts, which explores the impacts of various reforms of MID. The simulations confirm some of the arguments made about reforming MID but also demonstrate how the complex patterns of behavioural change induced by these reforms can lead to unanticipated effects. For example, the simulations suggest that it may be possible to reform MID whilst maintaining the prevailing rates of homeownership and reducing the federal budget deficit. The results also reiterate the conclusions of Chapter 3, namely that it is important to consider the interaction of policy design and the adjustment mechanisms at work in the property market, as these can alter the outcomes of policies that are designed to target specific groups.

2.0 Future Directions

This thesis presents a significant contribution to understanding and overcoming the challenges raised by spatial processes in the estimation of economic models and the evaluation of economic policy. The research presented within the thesis contributes to the frontiers of the topic and in doing so both highlights the need and paves the way for many new avenues of research.

In the context of estimation the thesis has made significant advances in developing a unified framework for conceptualising spatial data problems. The derivation of concise identification conditions and subsequent illustration of these through the use of Monte Carlo simulations provides an accessible introduction to the use of semi-parametric smoothing estimators for applied analysts. In the future it would be interesting to complement this research with a more formal examination of the asymptotic properties of the LASSE. It would also be of great benefit to see the work extended to deal with discrete data so that spatial smoothing techniques could be integrated into Logit and Probit models for analysing discrete choices. This would provide an opportunity to confront spatial data problems across a wider range of economic analyses, for example in the analysis of stated preferences through contingent valuation and choice experiments. This would be particularly useful in the analysis of recreational demand models where there is an obvious and overwhelming potential for omitted spatial covariates.

Similarly, the thesis presents the first applications of an ESM with an endogenous tenure choice. This methodological development broadens the range of policies that ESMs can be implemented to analyse. Through examining tenure choice and refining the housing supply function the model provides a useful tool for undertaking a general equilibrium analysis of policy change. The simulations presented in this thesis are based on calibrated models, it is essential for future research to develop an estimation procedure that can be adopted for policy analysis and through doing so identify the necessary data requirements for undertaking general equilibrium welfare analysis within this framework. In addition, I have focused on developing an endogenous tenure choice within the framework of the pure characteristics ESM developed by Epple and Platt (1998). In the future I would be interested in examining the extension of this to the random utility based ESM developed by Bayer et al (2004). This would also facilitate a transition from the homogeneous housing unit approach to a discrete treatment of housing. Housing supply itself presents a fruitful and challenging avenue for future research. In particular, the medium to long term housing supply function needs to be explored and the buy-to-let market could be integrated into the model to reflect housing supply decisions and the redistributive role of property ownership more accurately. More broadly, this research is part of a wider body of work at the forefront of the equilibrium sorting literature. As such, there are many developments taking place within this area, for example recent work has sought to explore moving costs (Bayer et al., 2009, Kuminoff, 2009), dynamics (Bayer et al., 2011), overlapping generations (Epple et al., 2010) and dual market models (Kuminoff, 2007, Kuminoff, 2009) which combine residential and labour decision making. Bringing together the innovations developed in this thesis with those evolving in the wider literature would provide a way to examine a broader range of the complexities and realisms of the economic problem.

APPENDIX A

The Polegate Bypass: Comparing Partial and General Equilibrium Welfare Measures

In this Appendix, I present the results of a more sophisticated ESM analysis designed to replicate the Polegate bypass case study more closely. The model specification is developed so as to imitate the situation in Polegate prior to the construction of the bypass. The introduction of the bypass is then simulated using information from the Post Opening Project Evaluation A27 Polegate Bypass report (POPE). As previously, results from both a partial equilibrium (PE) and general equilibrium (GE) ESM-based analysis are explored and contrasted.

A1 The A27 Polegate Bypass

Polegate is a small town in the East Sussex with 8,000 residents. Prior to the bypass Polegate had in the region of 18,000 vehicles per day passing through on the B2247 (Eastbourne News (19/06/2002)).

Figure A1: Road Map of Polegate (source Map data @ Google 2012)

Figure A1 provides a road map of Polegate. In 2002 the A27 Polegate Bypass was opened providing an alternative route to the original A27 (now the B2247) through Polegate between the A22 Cophall Roundabout and Golden Jubilee Way. A complete discussion of the objectives of the bypass and the project impacts is available in the POPE. This analysis focuses on the impact that the introduction of the bypass had on road noise levels across neighbourhoods. Under Part 1 of the

Land Compensation Act 1973, homeowners who were adversely affected by the change were able to claim compensation up to the value of the impact of the change on value of their properties. This compensation is included in the model.

A2 Calibrating the Model

A2.1 The Economy

Let us begin to develop a model of the Polegate area by dividing the area into neighbourhoods. Neighbourhoods are defined as areas within which the provision of local public goods is uniform. There are a number of features that could be used to define neighbourhoods. In this example I consider the age of the buildings in the area, the distance from Polegate town centre and the pre and post-bypass road noise exposure levels (provided in the Environmental Assessment Report). Information on the location of properties from OS Mastermap and the age of properties provided through the LandMap service was combined with information on the spatial distribution of changes in noise levels from the A27 Polegate Environmental Statement and assigned to a map of individual properties using ArcGIS. Property types were grouped into four categories. The intersection of these property types with changes in road noise level leads to the identification of 7 distinct neighbourhoods. The 7 neighbourhoods are displayed in the map in Figure A2. ArcGIS was then used to compute the average distance of properties within each neighbourhood to the centre of Polegate.

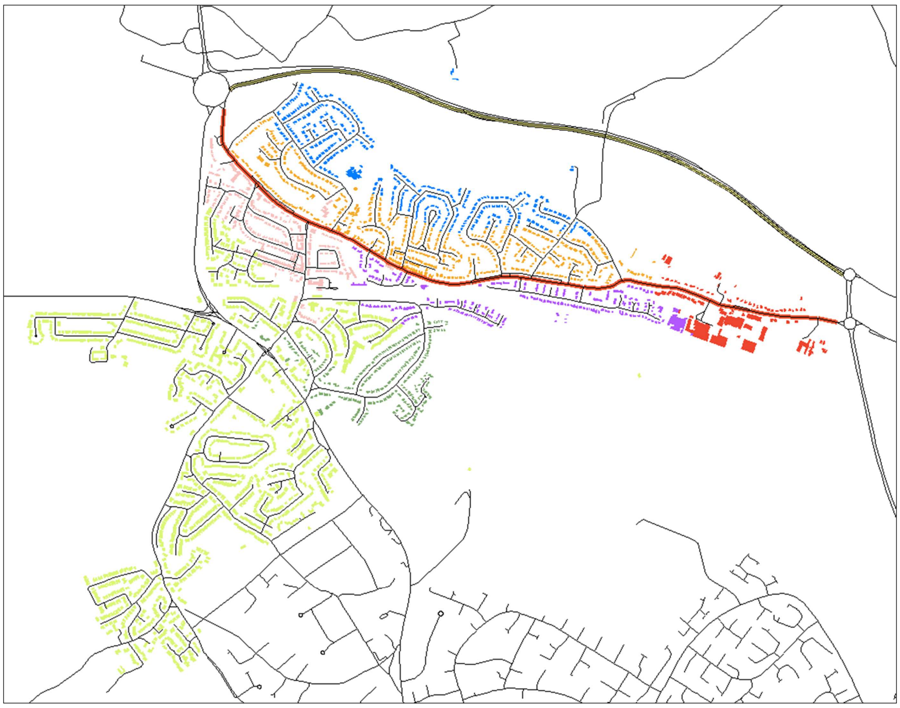

Figure A2: The 7 neighbourhoods identified by intersecting property type and road noise exposure before and after the bypass.

Data on property prices and characteristics, supplied by the Land Registry and Edina, along with census information on homeownership rates were added to the GIS layer. These were used to compute neighbourhood statistics on population shares and homeownership rates, which formed the basis of the calibration procedure for the parameters of the distribution of preferences for homeownership, $(\mu_\theta, \sigma_\theta^2)$, and the unobserved neighbourhood specific local public goods, $\boldsymbol{\xi}$. The data on property prices and characteristics were used to undertake a fixed effects hedonic regression to determine neighbourhood specific prices for a unit of homogeneous housing, $\boldsymbol{p^0}$.

The remainder of the model is calibrated in an identical manner to the two-neighbourhood models with rental revenue recycling that were discussed in the main body of this section. For compactness the details are not repeated here, however, the calibrated parameter values are detailed in Table A1 below.

Variables	Parameters	Values
Income and preference for housing	$(\mu_{\ln(y)}, \mu_\beta)$ $\Sigma_{\ln(y),\beta}$	$(9.83, 0.17)$ $\begin{pmatrix} 2.72 & -0.07 \\ -0.07 & 0.03 \end{pmatrix}$
Housing supply	a_j b	$\frac{10p_j^0}{H_j^{S\,0}}$ 0.1
Capacity	C_j	1.15
Preference for local public goods	α	0.11
Local public goods index weights	γ	0.02
Preferences for homeownership	$(\mu_\theta, \sigma_\theta^2)$	$(0.84, 0.02)$

Table A1: Parameters used in the calibration

A3 Welfare Effects of the Bypass

Taking the changes in noise levels predicted in the A27 Polegate Environmental Statement it was possible to replicate the environmental impact of the bypass. Table A3 characterises the seven neighbourhoods after the introduction of the bypass, taking into account household relocation, rental revenue recycling, housing supply constraints and compensation received under the Part I of the Land Compensation Act 1973.

For tractability it is prudent to separate our discussion of the results into three parts characterised by the direct environmental impact of the bypass. To simplify the discussion let us divide the neighbourhoods into three groups: ‘town centre’, ‘suburban’ and ‘bypass’. First, in the town centre neighbourhoods (1, 2 and 3) road noise falls by 4 dB as traffic is diverted away from the B2247. In these areas the environmental improvement leads to price increases as households with a relatively high preferences for environmental quality move there. Homeowners in these neighbourhoods enjoy capital gains while, in contrast, renters find these neighbourhoods less affordable. In fact, renters from town centre neighbourhoods 1 and 3 migrate to neighbourhood 6 near the bypass. The population movements are detailed in Table A4.

Second, suburban neighbourhoods (4 and 5) are south of the B2247 and experience no change in environmental quality. However, in the general equilibrium analysis they do experience changes in composition and price due to adjustments in the other neighbourhoods. As a result of the quality increases in the town centre neighbourhoods, and lower prices in the bypass neighbourhoods (6 and 7, to be discussed subsequently), prices in the suburban neighbourhoods fall as households migrate to neighbourhoods 1 in the town centre, and 6 and 7 near the bypass.

Third, bypass neighbourhoods (6 and 7) are located close to the new bypass and both experience an increase in road noise of 7dBs. This reduction in environmental quality prompts a series of relocation choices with some households seeking out higher environmental quality in the town centre (neighbourhood 3) and others being drawn to lower prices in the suburban area (neighbourhood 4). As housing demand contracts, prices in the bypass neighbourhoods fall. This attracts new demand from households who spend a relatively large share of their income on housing and provides benefits, in the form of lower rental costs, to renters. In addition, the payment of compensation provides large gains for owners.

Tables A5a and A5b summarises the welfare changes that are anticipated by both the partial and general equilibrium analyses respectively. The welfare changes include values for willingness to pay separated for renters and owners in each of the seven neighbourhoods. First, let us consider the partial equilibrium values. These are positive for renters and owners in the town centre neighbourhoods (1, 2 and 3), reflecting the environmental improvement, zero for households in the suburban neighbourhoods (4 and 5) since there is no change in the provision of local public goods here, and negative for households in the bypass neighbourhoods (6 and 7) who experience environmental degradation. In contrast, the general equilibrium values show that overall, the average willingness to pay value is positive in every neighbourhood. Moreover, for households initially located in the town centre neighbourhoods (1, 2 and 3) owners have higher average willingness to pay values than renters. In the suburban neighbourhoods (4 and 5) renters have higher average willingness to pay values than owners as a result of the benefits conveyed through falling rental prices. In the bypass neighbourhoods (6 and 7), homeowners have very high willingness to pay values as a result of the benefits provided by compensation. Renters have moderately high values, again attributable to the substantial fall in rental prices.

A4 Discussion

The application of our ESM to the Polegate bypass demonstrates the flexibility of the framework and its suitability for analysing these sorts of policy questions. The mechanics of the model have been discussed in detail in the main body of the section, here I focus upon the welfare changes induced by the A27 Polegate bypass and compare the key findings with those of the two-neighbourhood simulations.

Let us begin with the magnitude of welfare changes. As in the previous simulations, there is a large difference between the total willingness to pay values calculated using partial and general equilibrium analyses. In fact, for the Polegate bypass the difference is even greater, a result driven by the fact that the magnitude of the environmental changes in this analysis is greater than those examined in the two-neighbourhood simulations. In the partial analysis the total willingness to pay is negative as a result of the large losses in environmental quality that fall directly upon households initially located near the bypass route. The general equilibrium analysis sees these losses being avoided by households through relocation. As in the two-neighbourhood simulations environmental gentrification occurs with wealthier households from the suburbs relocating to the town centre and displacing poorer households, who in turn move to suburban or bypass neighbourhoods so as to benefit from lower property prices.

Turning now to the distribution of welfare changes, Figure A3 plots the changes in cumulative shares of utility for the current analysis. The pattern is, for the most part, equivalent to that presented in Figure 3.7 for the two-neighbourhood simulation. The partial analysis predicts a progressive re-distributional impact of the bypass with an increased share of total utility being provided to households with logged incomes less than 400,000. In contrast, under the general equilibrium analysis the policy has a small overall impact in terms of redistribution. In addition, the pattern of redistribution is almost the mirror image to that predicted by the partial analysis with the cumulative share of total utility falling amongst lower income households. Interestingly, in this policy analysis the greatest gains in the share of utility accrue to medium-to-high income households.

The results presented in this appendix demonstrate the application of the ESM with endogenous tenure choice to a policy analysis with several neighbourhoods. While the adjustments that take place are more complex to follow when the number of neighbourhoods increases, the results reiterate and reinforce those of the two-neighbourhood simulations. Comparing the PE and GE analyses reveals important differences in both the magnitude and the distribution of welfare changes. Furthermore, accounting for tenure choice and medium term property market adjustments reveals substantial differences in the welfare effects experienced by renters and owners, as well as across socio-economic groups. In conclusion, the partial equilibrium analysis provides misleading policy guidance. Moreover, the opportunities for harnessing environmental policies as a tool for redistribution is complicated by the complex adjustments that take place in the property market and interact with the policy.

	Town Centre						Suburban				Bypass			
	1		**2**		**3**		**4**		**5**		**6**		**7**	
Price	4449		5976		4576		4581		4176		4584		4331	
Population Share	0.07		0.11		0.06		0.45		0.05		0.16		0.10	
Homeownership Rate	0.83		0.93		0.89		0.89		0.77		0.91		0.79	
	Renters	Renters	Renters	Renters	Renters	Renters	Renters	Renters	Renters	Owners	Renters	Owners	Renters	Owners
Mean Income	24,053	45,832	36,238	75,674	29,067	49,817	28,459	82,596	7,980	44,537	41,517	111,460	22,074	47,163
Mean $\boldsymbol{\beta}$	0.39	0.39	0.03	0.02	0.35	0.35	0.22	0.21	0.63	0.61	0.08	0.08	0.47	0.46
Mean $\boldsymbol{\theta}$	1.01	1.12	1.00	1.11	1.00	1.12	1.00	1.11	1.02	1.13	1.00	1.11	1.02	1.13
Mean housing	2.35	4.47	0.16	0.33	2.45	4.19	1.48	4.01	1.30	7.02	0.87	2.14	2.78	5.64
Population	20	97	12	158	11	92	82	642	17	57	23	234	33	122

Table A2: Calibrated 7 neighbourhood baseline with rental revenue recycling

	Town Centre						Suburban				Bypass			
	1		2		3		4		5		6		7	
Price	4509		6516		5081		4537		4047		4283		4032	
Population Share	0.08		0.11		0.07		0.51		0.05		0.08		0.10	
Homeownership Rate	0.73		0.93		0.92		0.94		0.77		0.66		0.79	
	Renters	Owners	Renters	Owners	Renters	Owners	Renters	Owners	Renters	Owners	Renters	Owners	Renters	Owners
Mean Income	25,277	66,684	36,238	75,914	25,717	113,392	41,927	81,992	20,854	44,537	21,759	68,644	15,237	47,163
Mean $\boldsymbol{\beta}$	0.25	0.39	0.03	0.03	0.06	0.08	0.14	0.19	0.45	0.61	0.36	0.35	0.37	0.46
Mean $\boldsymbol{\theta}$	1.01	1.12	1.00	1.11	1.00	1.10	1.00	1.11	1.02	1.13	1.01	1.12	1.02	1.13
Mean housing	1.68	5.30	0.15	0.32	0.35	1.61	1.44	3.54	2.65	7.10	2.08	6.46	2.30	8.91
Population	36	97	12	160	9	109	48	771	17	57	44	86	32	122

Table A3: Neighbourhood characteristics after the bypass with revenue recycling, linear housing supply and compensation

	1		2		3		4		5		6		7	
	Renters	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**
From 1 Renters	0	0	0	0	0	0	0	0	0	0	20	0	0	0
From 1 Owners	0	0	95	0	0	0	0	0	0	0	0	2	0	0
From 2 Renters	0	0	0	12	0	0	0	0	0	0	0	0	0	0
From 2 Owners	0	0	0	0	158	0	0	0	0	0	0	0	0	0
From 3 Renters	0	0	0	0	0	0	0	0	0	0	11	0	0	0
From 3 Owners	0	0	0	0	0	0	9	0	0	0	0	83	0	0
From 4 Renters	0	36	0	0	0	0	0	34	0	0	12	0	0	0
From 4 Owners	0	0	2	0	0	0	0	0	639	0	0	1	0	0
From 5 Renters	0	0	0	0	0	0	0	0	0	0	0	0	17	0
From 5 Owners	0	0	0	0	0	0	0	0	0	0	0	0	0	0
From 6 Renters	0	0	0	0	0	9	0	14	0	0	0	0	0	0
From 6 Owners	0	0	0	0	2	0	100	0	132	0	0	0	0	0
From 7 Renters	17	0	0	0	0	0	0	0	0	17	1	0	15	0
From 7 Owners	0	0	0	0	0	0	0	0	0	0	0	0	0	122

Table A4: Population Movements

	Town Centre						Suburban				Bypass			
	1		2		3		4		5		6		7	
	Renters	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**
Mean WTP PE (£)	255	489	314	657	283	485	0	0	0	0	-803	-2,147	-531	-1,135
Total WTP PE (£)	-469,200													
Average WTP PE (£)	-293.25													
Total Compensation	861,250													
Change in rental revenues	0													
Change in mortgage payments	0													

Table A5: Partial equilibrium willingness to pay with rental revenue recycling, linear housing and compensation

	Town Centre						Suburban				Bypass			
	1		2		3		4		5		6		7	
	Renters	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**
Mean WTP GE (£)	238	619	246	770	260	2,195	113	98	175	41	128	5607	314	657
Total WTP GE (£)	4,593,600													
Average WTP GE (£)	2,871													
Change in rental revenues	860													
Change in mortgage Payments	38,580													
Total Compensation	861,250													
Net Benefits	3,771,790													

Table A5b: General equilibrium willingness to pay with rental revenue recycling, linear housing and compensation

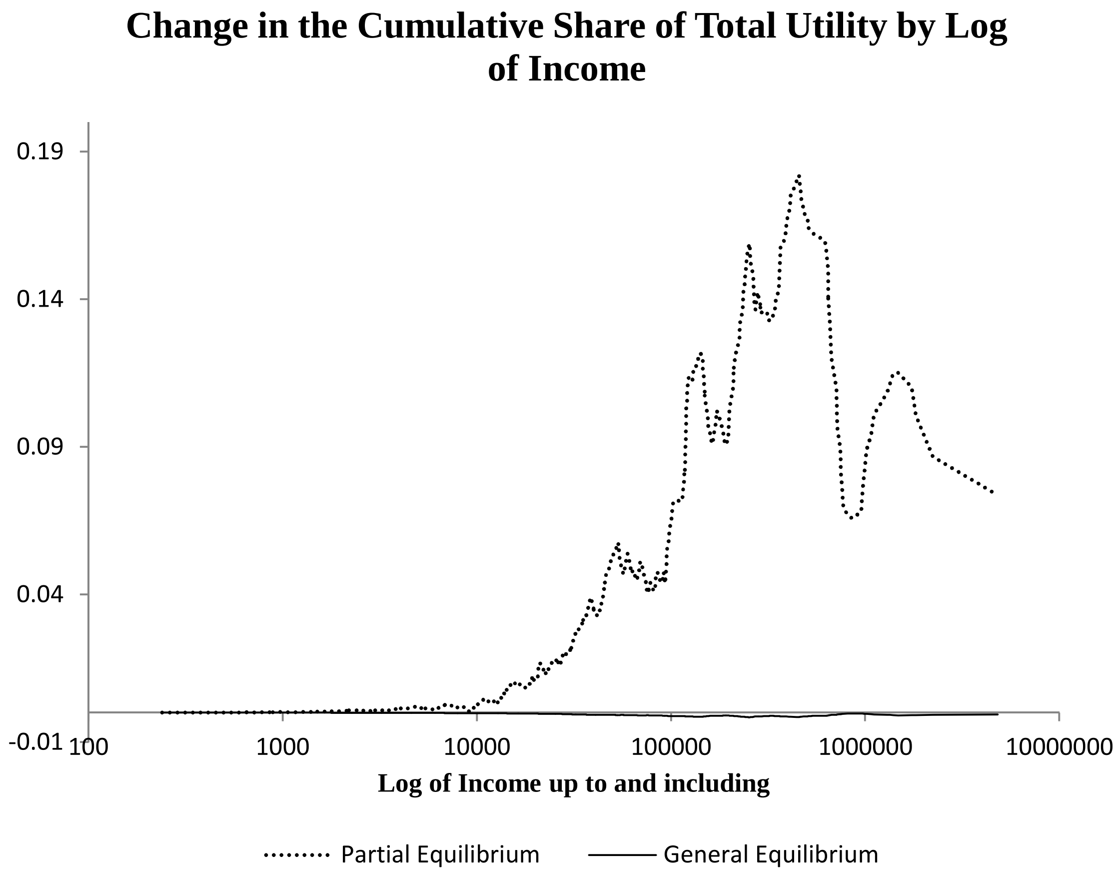

Figure A3a

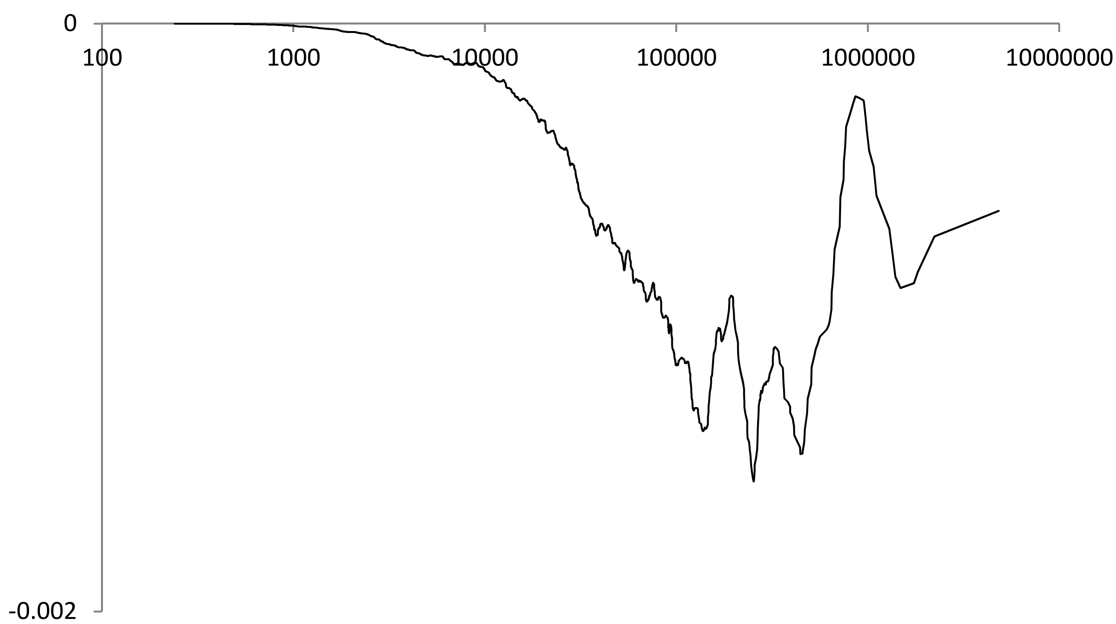

Figure A3b

Figure A3: The distribution of changes in the cumulative share of total utility by income in the partial and general equilibrium analyses. The bottom panel displays the general equilibrium plot at a finer scale.

REFERENCES

A SMITH, J. & E TODD, P. 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics,* 125**,** 305-353.

AKAIKE, H. 1974. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on,* 19**,** 716-723.

ANDREWS, D. & SÁNCHEZ, A. C. 2011. The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences. *OECD Journal: Economic Studies,* 2011.

ANSELIN, L. 1988. *Spatial econometrics: methods and models*, Springer.

ANSELIN, L. & GRIFFITH, D. A. 1988. DO SPATIAL EFFECTS REALLY MATTER IN REGRESSION ANALYSIS? *Papers in Regional Science,* 65**,** 11-34.

ANSELIN, L., COHEN, J., COOK, D., GORR, W. & TITA, G. 2000. Spatial analyses of crime. *Criminal justice,* 4**,** 213-262.

ANSELIN, L., FLORAX, R. & REY, S. J. 2004. *Advances in spatial econometrics: methodology, tools and applications*, Springer.

ANSELIN, L. 2010. Thirty years of spatial econometrics. *Papers in Regional Science,* 89**,** 3-25.

AUTOR, D. & HOUSEMAN, S. 2006. Temporary Agency Employment as a Way out of Poverty? *Working but poor: How economic and policy changes are affecting low-wage workers***,** 312-337.

BANZHAF, H. S. & WALSH, R. P. Testing for Environmental Gentrification: Migratory Responses to Changes in Environmental Quality. AERE Workshop, Estes Park, CO, May, 2004.

BARTELS, C. P. A. & KETELLAPPER, R. H. 1979. *Exploratory and explanatory statistical analysis of spatial data*, Springer.

BAUM-SNOW, N. 2007. Did Highways Cause Suburbanization?*. *Quarterly Journal of Economics,* 122**,** 775-805.

BAYER, P., MCMILLAN, R., MURPHY, A. & TIMMINS, C. 2011. A dynamic model of demand for houses and neighborhoods. National Bureau of Economic Research.

BAYER, P., MCMILLAN, R. & RUEBEN, K. 2004. An equilibrium model of sorting in an urban housing market. *NBER working paper*.

BAYER, P., KEOHANE, N. & TIMMINS, C. 2009. Migration and hedonic valuation The case of air quality. *Journal of Environmental Economics and Management,* 58**,** 1-14.

BAYER, P., FERREIRA, F. & MCMILLAN, R. 2007. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *The Journal of Political Economy,* 115**,** 588-638.

BELL, K. P. & BOCKSTAEL, N. E. 2000. Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Micro-Level Data. *Review of Economics and Statistics,* 82**,** 72-82.

BELSON, W. A. 1956. A technique for studying the effects of a television broadcast. *Applied Statistics***,** 195-202.

BENNETT, R. J. 1979. *Spatial time series: Analysis-forecasting-control,* Pion London.

BERGSTROM, T. C. & GOODMAN, R. P. 1973. Private demand for public goods. *The American Economic Review,* 63**,** 280-296.

BERRY, S. 1994. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*.

BINNER, A. & DAY, B. 2010. Spatial smoothing in hedonic price analysis: Identification of implicit prices in the presence of omitted spatial variables. *WCERE.* Montreal.

BINNER, A. & DAY, B. 2012. Exploring mortgage interest deduction reforms: An equilibrium sorting model with endogenous tenure choice. *Submitted to the Journal of Public Economics*.

BLACK, S. 1999. Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics,* 114**,** 577-599.

BOENTE, G. & RODRIGUEZ, D. 2008. Robust bandwidth selection in semiparametric partly linear regression models: Monte Carlo study and influential analysis. *Computational Statistics & Data Analysis,* 52**,** 2808-2828.

BOURASSA, S. C. & YIN, M. 2008. Tax Deductions, Tax Credits and the Homeownership Rate of Young Urban Adults in the United States. *Urban Studies,* 45**,** 1141-1161.

BRADY, M. & IRWIN, E. 2011. Accounting for spatial effects in economic models of land use: recent developments and challenges ahead. *Environmental and Resource Economics,* 48**,** 487-509.

BRAMLEY, G. & KARLEY, N. K. 2007. Homeownership, Poverty and Educational Achievement: School Effects as Neighbourhood Effects. *Housing Studies,* 22**,** 693-721.

BRODATY, T., CREÌ   PON, B., FOUGEÌ€RE, D. & CENTRE FOR ECONOMIC POLICY, R. 2000. *Using matching estimators to evaluate alternative youth employment programs : evidence from France, 1986-8,* London, Centre for Economic Policy Research.

CAPOZZA, D. R. & HELSLEY, R. W. 1989. The fundamentals of land prices and urban growth. *Journal of Urban Economics,* 26**,** 295-306.

CARBONE, J. & SMITH, V. 2008. Evaluating policy interventions with general equilibrium externalities. *Journal of Public Economics,* 92**,** 1254-1274.

CARD, D. & KRUEGER, B. 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review,* 84**,** 772-793.

CHOUMERT, J. & SALANIÉ, J. 2008. Provision of urban green spaces: some insights from economics. *Landscape Research,* 33**,** 331-345.

CORNEO, G. & JEANNE, O. 1999. Social Organization in an Endogenous Growth Model. *International Economic Review,* 40**,** 711-725.

CORRADO, L. & FINGLETON, B. 2012. WHERE IS THE ECONOMICS IN SPATIAL ECONOMETRICS?*. *Journal of Regional Science,* 52**,** 210-239.

DAVIS, P. 2006. Spatial Competition in Retail Markets: Movie Theaters. *The RAND Journal of Economics,* 37**,** 964-982.

DAY, B., BATEMAN, I. & LAKE, I. 2004. Omitted Locational Variates in Hedonic Analysis: A Semiparametric Approach using Spatial Statistics. *In:* CSERGE (ed.) *CSERGE Working Paper.*

DAY, B. 2005. *Hedonic Analysis of Property Markets: Theory and Applications to UK Cities.* Doctor of Philosophy in Economics, University College London (UCL).

DAY, B., BATEMAN, I. & LAKE, I. 2007. Beyond Implicit Prices: Recovering Theoretically Consistent and Transferable Values for Noise Avoidance from a Hedonic Property Price Model. *Environmental and Resource Economics,* 37**,** 211-232.

DIETZ, R. & HAURIN, D. 2003. The social and private micro-level consequences of homeownership. *Journal of Urban Economics,* 54**,** 401-450.

DIETZ, R. 2002. The estimation of neighborhood effects in the social sciences: An interdisciplinary approach* 1. *Social Science Research,* 31**,** 539-575.

DIPASQUALE, D. & KAHN, M. 1999. Measuring neighborhood investments: An examination of community choice. *Real Estate Economics,* 27**,** 389-424.

DREIER, P. 1997. The new politics of housing: How to rebuild the constituency for a progressive federal housing policy. *Journal of the American Planning Association*.

ELLICKSON, B. 1971. Jurisdictional fragmentation and residential choice. *The American Economic Review*.

EPIFANI, P. & GANCIA, G. 2006. Increasing Returns, Imperfect Competition, and Factor Prices. *Review of Economics and Statistics,* 88**,** 583-598.

EPPLE, D., ROMANO, R. & SIEG, H. 2012. The intergenerational conflict over the provision of public education. *Journal of Public Economics,* 96**,** 255-268.

EPPLE, D. 1991. Mobility and redistribution. *Journal of Political Economy*.

EPPLE, D. & PLATT, G. 1998. Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes. *Journal of Urban Economics,* 43**,** 23-51.

EPPLE, D., ROMER, T. & SIEG, H. 2001. Interjurisdictional sorting and majority rule: An empirical analysis. *Econometrica,* 69**,** 1437-1465.

EPPLE, D., ROMANO, R. & SIEG, H. 2010. The Intergenerational Conflict over the Provision of Public Education. *Working Paper*.

EPPLE, D. & ROMER, T. 1991. Mobility and redistribution. *The Journal of Political Economy,* 99**,** 828-858.

EPPLE, D. & FERREYRA, M. 2008. School finance reform: Assessing general equilibrium effects. *Journal of Public Economics,* 92**,** 1326-1351.

ESRC GLOBAL ENVIRONMENTAL CHANGE PROGRAMME 2001. Environmental justice: Rights and means to a healthy environment for all. *Special Briefing No 7*. ESRC.

FAN, J. & GIJBELS, I. 1995. Adaptive Order Polynomial Fitting: Bandwidth Robustification and Bias Reduction. *Journal of Computational and Graphical Statistics,* 4**,** 213-227.

FERNANDEZ, R. & ROGERSON, R. 1998. Public education and income distribution: a dynamic quantitative evaluation of education-finance reform. *The American Economic Review,* 88**,** 813-833.

FERREIRA, F. 2010. You can take it with you: Proposition 13 tax benefits, residential mobility, and willingness to pay for housing amenities. *Journal of Public Economics,* 94**,** 661-673.

FERREYRA, M. M. 2007. Estimating the effects of private school vouchers in multidistrict economies. *The American Economic Review,* 97**,** 789-817.

FLORAX, R. & FOLMER, H. 1992. Specification and estimation of spatial linear regression models: Monte Carlo evaluation of pre-test estimators. *Regional Science and Urban Economics,* 22**,** 405-432.

FOLLAIN, J., LING, D. & MCGILL, G. 1993. The preferential income tax treatment of owner‐occupied housing: Who really benefits? *Taylor & Francis*.

FOSTER, A. D. & ROSENZWEIG, M. R. 2003. Economic Growth And The Rise Of Forests*. *Quarterly Journal of Economics,* 118**,** 601-637.

FRANCISCO-FERNÁNDEZ, M. & VILAR-FERNÁNDEZ, J. 2005. Bandwidth selection for the local polynomial estimator under dependence: A simulation study. *Computational Statistics,* 20**,** 539-558.

FRIENDS OF THE EARTH 1999. Pollution Injustice: the geographic relation between household income and polluting factories. London: Friends of the Earth.

GALE, W. & GRUBER, J. 2007. Encouraging Homeownership Through the Tax Code. *Tax Notes*.

GALSTER, G. 1983. Empirical Evidence on Cross-Tenure Differences in Home Maintenance and Conditions. *Land Economics,* 59**,** 107-113.

GIBBONS, S. & MACHIN, S. 2003. Valuing English Primary Schools. *Journal of Urban Economics,* 53**,** 197-219.

GIBBONS, S. & MACHIN, S. 2006. Paying for Primary Schools: Admissions Constraints, School Popularity or Congestion. *Economic Journal* 116**,** C77-C92.

GIBBONS, S., LYYTIKÄINEN, T., OVERMAN, H. & SANCHIS-GUARNER, R. 2012. New Road Infrastructure: the Effects on Firms. *SERC Discussion Papers.* Spatial Economics Research Centre, LSE.

GIBBONS, S. & OVERMAN, H. G. 2012. MOSTLY POINTLESS SPATIAL ECONOMETRICS?*. *Journal of Regional Science*.

GLAESER, E. & SHAPIRO, J. 2002. The benefits of the home mortgage interest deduction.

GLAESER, E. L. & SACERDOTE, B. 1996. Why is there more crime in cities? : National Bureau of Economic Research.

GOULDER, L. H. & ROBERTON C. WILLIAMS, I. 2003. The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution. *Journal of Political Economy,* 111**,** 898-927.

GREEN, R. & VANDELL, K. 1999. Giving households credit: how changes in the US tax code could promote homeownership. *Regional Science and Urban Economics,* 29**,** 419-444.

GREEN, R. & WHITE, M. 1997. Measuring the Benefits of Homeowning: Effects on Children. *Journal of Urban Economics,* 41**,** 441-461.

GRIMES, A. & LIANG, Y. 2009. Spatial Determinants of Land Prices: Does Auckland’s Metropolitan Urban Limit Have an Effect? *Applied Spatial Analysis and Policy,* 2**,** 23-45.

HALL, P., MARRON, J. S. & PARK, B. U. 1992. Smoothed cross-validation. *Probability Theory and Related Fields,* 92**,** 1-20.

HARDLE, W. & MARRON, J. S. 1985. Asymptotic nonequivalence of some bandwidth selectors in nonparametric regression. *Biometrika,* 72**,** 481-484.

HARRIS, R. & JOHNSTON, R. 2008. Primary schools, markets and choice: studying polarization and the core catchment areas of schools. *Applied Spatial Analysis and Policy,* 1**,** 59-84.

HARRISON, D. & RUBINFELD, D. L. 1978. Hedonic Housing Prices and the Demand for Clean Air1. *Journal of Environmental Economics and Management,* 5**,** 81-102.

HAURIN, D., PARCEL, T. & HAURIN, R. J. 2002. Does homeownership affect child outcomes? *Real Estate Economics*.

HIGHWAYS AGENCY 2009. Post Opening Project Evaluation Five Year After Study: A27 Bypass. Department for Transport.

HILBER, C. & TURNER, T. 2010. The mortgage interest deduction and its impact on homeownership decisions.

HOFF, K. & SEN, A. 2005. *Homeownership, Community Interactions, and Segregation*.

HOSHINO, T. & KURIYAMA, K. 2010. Measuring the benefits of neighbourhood park amenities: application and comparison of spatial hedonic approaches. *Environmental and Resource Economics,* 45**,** 429-444.

ICHINO, A., MEALLI, F. & CENTRE FOR ECONOMIC POLICY, R. 2006. *From temporary help jobs to permanent employment : what can we learn from matching estimators and their sensitivity?,* London, Centre for Economic Policy Research.

JONES, C. I. 2005. The Shape of Production Functions and the Direction of Technical Change*. *Quarterly Journal of Economics,* 120**,** 517-549.

JUST, R. E., HUETH, D. L. & SCHMIDTZ, A. 2004. *The Welfare Economics of Public Policy: A Practical Approach to Project and Policy Evaluation,* Cheltenham, United Kingdom, Edward Elgar.

KELEJIAN, H. H. & PRUCHA, I. R. 1998. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *The Journal of Real Estate Finance and Economics,* 17**,** 99-121.

KELEJIAN, H. H. & PRUCHA, I. R. 1999. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review,* 40**,** 509-533.

KERRY SMITH, V., SIEG, H., SPENCER BANZHAF, H. & WALSH, R. P. 2004. General equilibrium benefits for environmental improvements: projected ozone reductions under EPA's Prospective Analysis for the Los Angeles air basin. *Journal of Environmental Economics and Management,* 47**,** 559-584.

KIM, J. & GOLDSMITH, P. 2008. A Spatial Hedonic Approach to Assess the Impact of Swine Production on Residential Property Values. *Environmental and Resource Economics*.

KLAIBER, A. H. & SMITH, K. V. 2010. General equilibrium benefit analyses for social programs. *Working Paper*.

KLAIBER, A. H. & PHANEUF, D. J. 2010. Valuing open space in a residential sorting model of the Twin Cities. *Journal of Environmental Economics and Management,* 60**,** 57-77.

KLAIBER, H. A. & SMITH, V. K. 2009. General Equilibrium Benefit Analyses for Social Programs. *paper commissioned by the Benefit-Cost Analysis Center, University of Washington*.

KNEIB, T., MÜLLER, J. & HOTHORN, T. 2006. Spatial Smoothing Techniques for the Assessment of Habitat Suitability. *In:* STATISTIK, I. F. (ed.) *Sonderforschungsbereich 386.* Institut für Medizininformatik, Biometrie und Epidemiologie

Freidrich-Alexander-Universitat Erlangen-Nürnberg.

KUMINOFF, N., SMITH, V. & TIMMINS, C. 2010. The New Economics of Equilibrium Sorting and its Transformational Role for Policy Evaluation. *NBER working paper*.

KUMINOFF, N. 2007. Recovering preferences from a dual-market locational equilibrium. *Manuscript Department of Agricultural and Applied Economics, Virginia Tech*.

KUMINOFF, N. V. 2009. Decomposing the structural identification of non-market values. *Journal of Environmental Economics and Management*, 57**,** 123-139.

LAGARIAS, J., REEDS, J. & WRIGHT, M. 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. *Siam journal of optimization*.

LEE, S. 2008. *Three essays on spatial autoregressive models and empirical organization.* PhD, Ohio State University.

LEE, T. C. M. & SOLO, V. 1999. Bandwidth selection for local linear regression: A simulation study. *Computational Statistics* 14**,** 515-532.

LEE, T. C. M. 2003. Smoothing parameter selection for smoothing splines: a simulation study. *Computational Statistics & Data Analysis,* 42**,** 139-148.

LEGGETT, C. G. & BOCKSTAEL, N. E. 2000. Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management,* 39**,** 121-144.

LIU, F. 2000. *Environmental justice analysis: Theories, methods, and practice*, CRC.

MADDISON, D. 2000. A hedonic analysis of agricultural land prices in England and Wales. *European Review of Agricultural Economics,* 27**,** 519-532.

MALLOWS, C. L. 1973. Some comments on C p. *Technometrics,* 15**,** 661-675.

MCINTOSH, C. 2008. Estimating Treatment Effects from Spatial Policy Experiments: An Application to Ugandan Microfinance. *Review of Economics and Statistics,* 90**,** 15-28.

MCMILLEN, D. P. 2010. Issues in spatial data analysis. *Journal of Regional Science,* 50**,** 119-141.

MCMILLEN, D. P. 2012. PERSPECTIVES ON SPATIAL ECONOMETRICS: LINEAR SMOOTHING WITH STRUCTURED MODELS. *Journal of Regional Science,* 52**,** 192-209.

MONTGOMERY, C. 1992. Explaining home improvement in the context of household investment in residential housing. *Journal of Urban Economics,* 32**,** 326-350.

MORAN, P. A. P. 1950. NOTES ON CONTINUOUS STOCHASTIC PHENOMENA. *Biometrika,* 37**,** 17-23.

MÜLLER, M. 2001. Estimation and testing in generalized partial linear models—A comparative study. *Statistics and Computing,* 11**,** 299-309.

NECHYBA, T. & STRAUSS, R. 1994. Community choice and local public services: A discrete choice approach. *Regional Science and Urban Economics,* 28**,** 51-73.

NECHYBA, T. 2003a. Introducing school choice into multidistrict public school systems. *National Bureau of Economic Research,* The Economics of School Choice.

NECHYBA, T. J. 2003b. Introducing school choice into multidistrict public school systems. *National Bureau of Economic Research.*

NESHEIM, L. 2002. Equilibrium sorting of heterogeneous consumers across locations: theory and empirical implications. *Cemmap working paper.*

OATES, W. E. 1969. The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis. *The Journal of Political Economy,* 77**,** 957-971.

OFFICE OF MANAGEMENT AND BUDGET 2011. Analytical Perspectives. *Budget of the U.S. Government, Fiscal Year 2011*.

PACE, R. K. & GILLEY, O. W. 1997. Using the spatial configuration of the data to improve estimation. *The Journal of Real Estate Finance and Economics,* 14**,** 333-340.

PACIOREK, C. J. 2010. The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical science: a review journal of the Institute of Mathematical Statistics,* 25**,** 107.

PAELINCK, J. & KLAASSEN, L. 1979. *Spatial Econometrics,* Farnborough, Saxon House.

PARTRIDGE, M. D., BOARNET, M., BRAKMAN, S. & OTTAVIANO, G. 2012. INTRODUCTION: WHITHER SPATIAL ECONOMETRICS? *Journal of Regional Science,* 52**,** 167-171.

PINKSE, J. & SLADE, M. E. 2010. THE FUTURE OF SPATIAL ECONOMETRICS*. *Journal of Regional Science,* 50**,** 103-117.

POGHARIAN, S. V. 1990. *Expanding the housing supply through conversions of the existing stock.* Massachusetts Institute of Technology.

POTERBA, J. & SINAI, T. 2008. Tax expenditures for owner-occupied Housing: Deductions for Property Taxes and mortgage interest and the exclusion of imputed rental income. *The American Economic Review*.

POUSTIE, M. 2004. Environmental justice in SEPA's environmental protection activities: a report for the Scottish Environment Protection Agency. University of Strathclyde Law School.

PUECH, F. 2004. How do criminals locate? Crime and spatial dependence in Minas Gerais. *Urban/Regional 0509018 Economics Working Paper Archive Econ WPA*.

RACINE, J. 2001. Bias-corrected kernel regression. *Journal of Quantitative Economics,* 17**,** 25-42.

REHDANZ, K. & MADDISON, D. 2008. Local environmental quality and life-satisfaction in Germany. *Ecological Economics,* 64**,** 787-797.

ROBINSON, P. M. 1988. Root-N-Consistent Semiparametric Regression. *Econometrica,* 56**,** 931-954.

ROHE, W. 1996. Homeownership and neighborhood stability. *Housing policy debate*.

ROSENBAUM, P. R. 1986. Dropping out of high school in the United States: An observational study. *Journal of Educational and Behavioral Statistics,* 11**,** 207-224.

ROTHENBERG, J., GALSTER, G. C., BUTLER, R. V. & PITKIN, J. R. 1991. *The maze of urban housing markets: Theory, evidence, and policy*, University of Chicago Press.

RUBIN, D. B. 1973. Matching to remove bias in observational studies. *Biometrics***,** 159-183.

RUPPERT, D. & WAND, M. P. 1994. Multivariate locally weighted least squares regression. *The Annals of Statistics***,** 1346-1370.

RUPPERT, D., SHEATHER, S. J. & WAND, M. P. 1995. An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association,* 90**,** 1257-1270.

SACERDOTE, B. 1996. Why is there more crime in cities? *NBER working paper*.

SAH, R. 1991. Social Osmosis and Patterns of Crime: A Dynamic Economic Analysis. *Journal of Political Economy,* 99.

SAMPSON, R. & RAUDENBUSH, S. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*.

SANCHIS-GUARNER, R. 2012. Driving Up Wages: The Effects of Road Construction in Great Britain. *SERC Discussion Papers.* Spatial Economics Research Centre, LSE.

SCHELLING, T. C. 1969. Models of Segregation. *The American Economic Review,* 59**,** 488-493.

SHAPIRO, J. & GLAESER, E. L. 2003. The Benefits of the Home Mortgage Interest Deduction. *Tax Policy and the Economy,* 17**,** 37.

SIEG, H., SMITH, K. V., BANZAHF, H. S. & WALSH, R. P. 2004. Estimating the general equilibrium benefits of large changes in spatially delineated public goods. *International Economic Review,* 45**,** 1047-1077.

SMITH, V. K. & CARBONE, J. C. 2007. Should Benefit–Cost Analyses Take Account of General Equilibrium Effects? *Research in Law and Economics,* 23**,** 247-272.

SMITH, V. K., SIEG, H., SPENCER BANZHAF, H. & WALSH, R. 2004. General equilibrium benefits for environmental improvements: Projected ozone reductions under EPA's prospective analysis for the Los Angeles air basin. *Journal of Environmental Economics and Management,* 47, 559-584.

STANSEL, D. 2011. Unmasking the Mortgage Interest Deduction: Who Benefits and by How Much? *heartland.org*.

THE NATIONAL COMMISSION FISCAL RESPONSIBILITY AND REFORM 2010. *The Moment of Truth*.

TIEBOUT, C. 1956. A pure theory of local expenditures. *The Journal of Political Economy,* 64, 416-424.

TODER, E. 2010. Reforming the Mortgage Interest Deduction. *urban.org*.

TRA, C. I. 2010. A discrete choice equilibrium approach to valuing large environmental changes. *Journal of Public Economics,* 94, 183-196.

VENTRY JR, D. 2010. The accidental deduction: A history and critique of the tax subsidy for mortgage interest. *Law & Contemp. Prob.*

WAGSTAFF, A., YU, S. & WORLD BANK. DEVELOPMENT RESEARCH GROUP. PUBLIC SERVICES, T. 2005. *Do health sector reforms have their intended impacts? : the World Bank's Health VIII Project in Gansu Province, China*, Washington, D.C., World Bank, Development Research Group, Public Services Team.

WALKER, G. 1998. Environmental justice and the politics of risk. *Town and Country Planning,* 67, 358-359.

WALKER, G. & BICKERSTAFF, K. 2000. Polluting the poor: an emerging environmental justice agenda for the UK? *Critical Urban Studies occasional papers, Centre for Urban and Community Research, Goldsmiths College, University of London, London*.

WALKER, G. Who bears environmental risk?: the challenges of the environmental equality agenda. Hazard Forum Synopsis, 2003.

WALSH, R. 2003. Analyzing Open Space Policies in a Locational Equilibrium Model with Endogenous Landscape Amenities. Boulder, CO: University of Colorado.

WALSH, R. 2007. Endogenous open space amenities in a locational equilibrium. *Journal of Urban Economics,* 61, 319-344.

WALTERT, F., SCHULZ, T. & SCHLÄPFER, F. 2011. The role of landscape amenities in regional development: evidence from Swiss municipality data. *Land Use Policy,* 28, 748-761.

WALTON, W. & SHAW, J. 2003. Applying the new appraisal approach to transport policy at the local level in the UK. *Journal of Transport Geography,* 11**,** 1-12.

ZABEL, J. E. & KIEL, K. A. 2000. Estimating the demand for air quality in four US cities. *Land Economics*, 76**,** 174-194.